



Università degli Studi di Salerno
Dipartimento di Studi Umanistici

Dottorato di ricerca
in
Studi Letterari, Linguistici e Storici

Ciclo XXXVIII

Coordinatrice del Corso di Dottorato: prof.ssa Maria Voghera

Curriculum di Studi Linguistici

Tesi di dottorato in Linguistica Inglese:

Appil: an Interface for Lexical and Linguistic Competence Improvement

Tutor: prof.ssa Siria Guzzo

Co-tutor: prof. Vittorio Scarano

Candidato:

Pasquale Esposito

Anno accademico 2025/2026

Alle anime gentili:
per Giulia, per nonno P.

Index

ABSTRACT	I
PREFACE	II
ACKNOWLEDGEMENTS	III
LIST OF FIGURES	IV
TABLES	VI
FORMULAS	VII
CODING BLOCKS	VIII
LIST OF ABBREVIATIONS	IX
<i>INTRODUCTION</i>	1
1 INVESTIGATING THE FOUNDATIONS OF SECOND LANGUAGE LEARNING	5
1.1 Introduction	5
1.2 Understanding Core Constructs and Competence in SLA	5
1.3 Krashen's Monitor Model and the Power of Input	7
1.4 Swain's Output Hypothesis: The Role of Comprehensible Output	9
1.5 Long's Interaction Hypothesis: Language Learning Through Exchange	11
1.6 Chomsky's Universal Grammar and its Role in SLA	14
1.7 From Instructed Learning to Task-Based Pedagogy: The Work of Rod Ellis	16
1.8 Paul Nation and the Central Role of Vocabulary in SLA	18
1.8.1 Measuring Vocabulary Growth with Statistical Methods	21
1.8.2 Lexical Development in Second Language Acquisition	27
1.8.3 Supported Approaches To Words Internalization	29
1.8.4 Theoretical Underpinnings Of Text Enhancement Effectiveness	32
1.9 Cognitive Dimensions Of Second Language Learning	39
1.9.1 Integrating Behavioral Studies Into The System	42
1.9.2 Evidence In Mental Lexicon Studies	49
1.10 The CEFR Framework: Linking Proficiency to Pedagogy	54
1.10.1 Curriculum Design	57
1.10.2 CEFR–Linguistic Branch Alignment	61
1.11 Conclusion	64
2 INNOVATIONS IN SECOND LANGUAGE LEARNING: COMPUTER-ASSISTED AND INTELLIGENT CALL	65

2.1	Introduction	65
2.2	Early CALL and Its Drill-Based Roots	65
2.3	The Move Toward Communicative CALL	67
2.4	Integrative CALL in the Age of Digital Learning	68
2.5	Intelligent CALL: From Static Programs to Adaptive Systems	70
2.6	Previous Works in the Field of CALL	72
2.6.1	SMILLE	74
2.6.2	Revita	74
2.6.3	Other Systems in CALL/ICALL	80
2.6.4	Comparative Discussion	81
2.7	Rethinking CALL and ICALL: Challenges and Critiques	85
2.7.1	Identified Gaps in Current CALL/ICALL	87
2.8	Implementing Digital Language Tools: Pedagogical and Technological Hurdles	89
2.8.1	Theoretical underpinnings	90
2.9	Looking Ahead: Opportunities in Digital Language Learning	92
2.10	Conclusion	94
3	METHODOLOGY: DIDACTICS AND COMPUTER SCIENCE	95
3.1	Introduction	95
3.2	Introduction to Didactic Methodology	95
3.2.1	Promoting Self-Regulated Learning	98
3.3	Introduction to Computer Science Methodology	130
3.3.1	Knowledge Graphs in Knowledge Representation and Reasoning	131
3.3.2	Ontologies as the Backbone of Linked Open Data in Linguistics	135
3.3.3	Network analysis in linguistic knowledge representation	137
3.3.4	Integrating Open Data and AI for Educational Systems	138
3.3.5	Systematic Literature Review: Linked Open Data	145
3.3.6	Meaning Representation	159
3.3.7	LLOD Accessibility	174
3.3.8	Practical Application of LLOD	179
3.3.9	Final evaluation and state of the art	194
3.4	Conclusion	196
4	APPIL: BRIDGING THE GAP IN ADAPTIVE LANGUAGE LEARNING	197
4.1	Introduction	197
4.2	The Design Philosophy Behind Appil	198
4.3	Core Functionalities of the Appil Platform	200
4.3.1	Adaptive Text Processing	200
4.3.2	Lexical Organization and Phraseology	201
4.3.3	Dynamic Personalization and Learner Profiles	201
4.3.4	Lexical Reinforcement through Gamified Interaction	204
4.3.5	Feedback and Micro-Negotiations of Meaning	205
4.3.6	Cognitive Load Management	207
4.3.7	CEFR-Aligned Progression	208
4.3.8	Automated Generation of Lesson-Based Activities	209
4.3.9	Contextualized Pronunciation and Interactive Reading	212

4.3.10	Spatial Involvement and Multimodal Cognition	212
4.3.11	Augmented Reality and Immersive Lexical Exploration	214
4.3.12	Network analysis and semantic connectivity	215
4.3.13	User-Centered Transparency and Interpretability	216
4.3.14	Overt and Covert teaching processes	217
4.3.15	Assessing Text Complexity	225
4.4	Cognitive and Pedagogical Dimensions of Appil	271
4.5	Conclusion	273
5	TECHNICAL INFRASTRUCTURE AND ARCHITECTURE	274
5.1	Introduction	274
5.2	Early Steps: From Static Web to Scalable Vision	275
5.3	Early Web Prototyping (HTML, CSS, JS)	275
5.3.1	Evaluating PHP: A Short-Lived Contender	275
5.3.2	Recognizing the Need for a Scalable Backend	276
5.3.3	Initial Web Deployment via JSP and Servlet Architecture	279
5.4	Architecture and Technology Stack	279
5.4.1	Microservices Architecture: Modularity and Scalability	279
5.4.2	Python Integration: NLP and Computational Linguistics	280
5.4.3	Modular, Scalable, and Extensible Design	281
5.5	Database Design and Data Management	284
5.5.1	Data Streams and Structures	284
5.5.2	Efficiency in Storage and Retrieval	285
5.5.3	Performance Considerations and Query Optimization	286
5.5.4	Futureproofing and Data Portability	287
5.6	Interface Design and User Experience	288
5.6.1	Clarity, Accessibility, and Cross-Platform Compatibility	288
5.6.2	Iterative Design Based on User Feedback	289
5.6.3	Usability Testing and Workflow Optimization	290
5.7	Pedagogical Alignment and Feature Integration	291
5.7.1	Mapping SLA Theory to Application Features	291
5.7.2	Adaptive Reading and Vocabulary Modules	292
5.7.3	Feedback Dashboards and Learner Profiling	292
5.8	Performance and Reliability Testing	293
5.8.1	Load Testing and Computational Demands	293
5.8.2	Ensuring Seamless Multi-Device Operation	294
5.8.3	Debugging Integration Bottlenecks	296
5.9	Technological and development limitations	297
5.10	Future Directions and Scalability	298
5.10.1	Planned Enhancements (AI, LMS Integration, Analytics)	298
5.10.2	Deliberate Constraints on Generative AI and LLM	299
5.10.3	Planned Integration of Generative AI	300
5.11	Conclusion	304
6	FIELDWORK: STUDY DESIGN, TESTING, AND ANALYSIS	306
6.1	Introduction	306

6.1.1	Project Details	306
6.2	Planning phase	307
6.3	Profiling Learners: Diagnostics for Learner-Centered Instruction	309
6.3.1	Demographic and Linguistic Profile	310
6.3.2	Perceived Needs and Self-Assessment	313
6.3.3	Vocabulary Section	316
6.3.4	Grammar Section	319
6.4	Platform setup	320
6.5	The Didactic Plan Behind Appil	322
6.5.1	Methodology and Didactic Approach	324
6.5.2	Lessons' plan	328
6.5.3	Lessons overview	329
6.6	Participant Response: Interpreting Feedback in Light of Access Patterns	334
6.6.1	Usage metrics	335
6.6.2	Usage Metrics: Behavioral Insights from Web Server Logs	337
6.6.3	Lexical Distribution and Transparency: Interpretation of Similarity on Frequency Metrics	341
6.6.4	Conclusion	347
7	CONCLUSION	349
7.1	Study Limitations	349
7.1.1	Ethical considerations	351
7.2	Future Research Directions	354
7.3	Concluding Reflections	358
	REFERENCES	368
	APPENDIX	421
A.	Profiling questionnaire	421
B.	Art. 5 GDPR Principles relating to processing of personal data	427
C.	Licensing	427
D.	Didactic plan overview	436
1.	Lesson 1: Introduction to English Phonetics and the Alphabet	436
2.	Lesson 2: The Verb <i>To Be</i>	438
3.	Lesson 3: Nouns and Articles in English	439
4.	Lesson 4: The Verb <i>To Have</i>	441
5.	Lesson 5: The Past Simple Tense	442
6.	Lesson 6: Vocabulary improvement and lessons recap	444
7.	Lesson 7: Subject Pronouns & Object pronouns	446
8.	Lesson 8: Prepositions of Place	448

Abstract

The ongoing evolution of digital education has opened new possibilities for second language learning, yet many existing platforms remain constrained by commercial interests and limited pedagogical depth. Despite the proliferation of language learning applications, few are explicitly grounded in robust Second Language Acquisition (SLA) theory (Ellis, 2008; Krashen, 1982; Long, 1996) or capable of offering authentic, learner-driven adaptation. This research emerges within that gap, responding to the growing need for open, scientifically informed, and pedagogically coherent digital tools that can personalize instruction while preserving methodological rigor.

This doctoral study explores the intersection of Second Language Acquisition (SLA), digital pedagogy, and computational linguistics through the design, development, and evaluation of Appil, an adaptive, web-based language learning platform created by me, under supervision, during my doctoral studies, and situated within the broader field of Intelligent Computer-Assisted Language Learning (ICALL) (Heift & Schulze, 2007). It integrates theoretical insights from SLA, psychology, and cognitive science (Evans, 2008; Kahneman, 2011; Vygotsky, 1986) with technological innovations drawn from Natural Language Processing (NLP) and Linguistic Linked Open Data (LLOD) (Chiarcos et al., 2013). The overarching goal is to demonstrate how open, data-driven systems can enhance learner autonomy and motivation (Deci & Ryan, 2014; Zimmerman, 2002) while maintaining instructional coherence within a structured learning framework.

Adopting a mixed-methods approach, the present research project combines theoretical synthesis, technical implementation, and empirical evaluation. It includes a systematic review of SLA theories and digital learning methodologies, the computational design of the Appil platform, and a field-based evaluation of its pedagogical effectiveness. The system's architecture merges CEFR-aligned progression (Council of Europe, 2020) with NLP-enabled personalization, providing learners with individualized flashcards, adaptive exercises, and real-time feedback tailored to their evolving proficiency levels.

Findings reveal that personalization grounded in SLA theory, rather than purely algorithmic adaptation, significantly enhances learner motivation, vocabulary retention, and engagement. Learners responded positively to flexible and user-driven content pathways, although sustaining participation and addressing digital literacy barriers remained ongoing challenges. The development process, conducted by a single researcher, underscored both the potential and the practical constraints of independent, interdisciplinary platform creation, while also highlighting the centrality of human engagement throughout design and implementation.

Preface

Open-source education and learning models should be promoted as accessible to all and free from commercial constraints. Over the past decade, numerous applications, both commercial and open access, have been developed to support second language learning. While some are grounded in robust pedagogical principles, others place greater emphasis on gamification. Nevertheless, these initiatives collectively highlight the potential of technology to facilitate language acquisition and skill development.

Despite the advent of artificial intelligence, maintaining the quality of learning remains paramount. A robust, scientifically grounded approach to education is indispensable for fostering the development of essential competencies. The current era is characterized by an abundance of data, including vast linguistic resources. However, raw data is not readily usable. To harness this potential and capitalize on individual learning preferences, it was deemed necessary to develop a tool that draws deeply on the theoretical frameworks of leading educational, psychological, and cognitive researchers.

At the same time, the project aimed to integrate best practices and findings from the field of digital learning technologies. This objective was pursued through the incorporation of knowledge from multiple domains, including second language acquisition, psychology, computer science, and cognitive science.

The primary goal was to design an interface that goes beyond traditional formal learning environments, offering a scientifically validated and progressively adaptive learning experience tailored to individual needs. To this end, an extensive review of the literature on second language learning was conducted to analyze and synthesize existing methods and techniques. Subsequently, a systematic review of open data was undertaken to identify key repositories and data structures suitable for integration into the digital interface.

The following chapters present the theoretical background, technical development, and both the pedagogical and technical evaluation of the proposed interface.

A handwritten signature in black ink, reading "Roxane Gpik". The signature is fluid and cursive, with the first name "Roxane" being more prominent and the last name "Gpik" following it.

Acknowledgements

This work was born from a desire to make language learning more accessible, efficient, and meaningful by bridging the gap between linguistics and computer science. My aim was not only to conduct research for its own sake, but to translate it into tangible results that could help learners and educators alike benefit from the potential of technology in education.

This research has also been an opportunity to broaden my experience and to move my horizon a step further. Engaging in an interdisciplinary project that connects language education, cognitive science, and digital learning has allowed me to explore new perspectives and deepen my understanding of how technology can transform human communication and learning. The process of inquiry, experimentation, and continuous refinement has been both challenging and rewarding, marking a meaningful step in my academic and personal development.

To my supervisors, I am profoundly grateful for their guidance, patience, and trust. Their insights challenged me to think critically and pushed me to refine this project into something both rigorous and practical.

My sincere thanks also go to the Municipality of Fisciano and the learners who participated in this study, their engagement and openness made it possible to transform theory into a living, interactive experience.

Finally, I dedicate this work to all those who believe in open, human-centered education and in the power of collaboration between technology and pedagogy. It is through their vision and shared commitment that language learning can continue to evolve, becoming not only more intelligent, but also more inclusive and transformative for everyone.

I would like to express my deepest gratitude to everyone who has accompanied and supported me throughout this research journey.

List of Figures

Figure 1 The Depth Individual Word Knowledge model Wolter's (2001).....	125
Figure 2 LLOD cloud September 2023 - https://linguistic-lod.org/lod-cloud	144
Figure 3 LLOD Cloud, January 2011	145
Figure 4 LLOD Cloud September 2023.....	145
Figure 5 PRISMA chart LLOD research.....	148
Figure 6 LLOD cloud resources accessibility for category	177
Figure 7 Amount of data in terms of RDF triples potentially accessible via a SPARQL endpoint.....	179
Figure 8 Appil Welcome Page	203
Figure 9 Appil's personalized content selection interface, presenting daily proposed texts alongside learner-selected materials sourced from Wikipedia and other CC BY-SA 4.0-compliant repositories. This design combines adaptive recommendation with open-access resources, supporting learner autonomy while maintaining pedagogical alignment and content transparency.....	204
Figure 10 Appil's personalized content search and selection interface, integrating daily recommended texts with learner-driven exploration of Wikipedia and other CC BY-SA-compliant materials	205
Figure 11 Appil's flashcard-based translation activity, illustrating how personalized lexical input is converted into adaptive, practice-oriented learning tasks.....	206
Figure 12 Appil's adaptive crossword activity, illustrating gamified retrieval practice for personalized lexical consolidation	207
Figure 13 Appil's adaptive feedback and semantic mapping interface, demonstrating micro-negotiations of meaning through contextualized network visualization (word: tools)	208
Figure 14 Appil's adaptive feedback and semantic mapping interface, demonstrating micro-negotiations of meaning through contextualized network visualization (word: Galaxy)	209
Figure 15 Appil's adaptive interface visualizing text input, lexical mapping, and interactive task components as mechanisms for managing cognitive load and sustaining engagement	210
Figure 16 Appil's lesson progression interface demonstrating CEFR-aligned sequencing and cumulative skill development across modules.....	211
Figure 17 Automatically generated Job Titles Quiz illustrating how Appil converts lesson content into adaptive, personalized exercises that reinforce contextual understanding.....	212
Figure 18 Appil's Irregular Plurals Game' illustrating the integration of abstract linguistic resources with interactive, applied learning practice.....	213
Figure 19 Appil's Verb Matching Memory Game, demonstrating adaptive difficulty scaling and gamified reinforcement of morphological competence.....	214
Figure 20 Appil's interactive reading and pronunciation activity, demonstrating contextualized listening and user-driven lexical selection for adaptive multimodal learning	215
Figure 21 Interactive map activity in Appil, demonstrating spatial involvement and multimodal grounding of lexical and cultural knowledge	216
Figure 22 Augmented-reality lexical exploration in Appil, where learners navigate a three-dimensional semantic network enriched with bilingual descriptions and multimedia references	217
Figure 23 Network visualization of the lexical item decoy demonstrating semantic associations and relational structure within Appil's adaptive lexicon analysis framework	219
Figure 24 Snapshot of the database with computed metrics	247
Figure 25 Distribution of lexical items across three distinct vocabularies: NVdB, NBC-FL, and AWL. Each circle represents one vocabulary, with the overlapping areas indicating the intersection of words present in multiple vocabularies	263
Figure 26 The correlation matrix illustrates the relationships between key numerical features in the dataset. Each cell represents the Pearson correlation coefficient between two variables. Values close to 1 indicate a strong positive relationship, values near -1 indicate a strong negative relationship, and values around 0 suggest little to no linear association.....	268
Figure 27 This histogram illustrates the distribution of the Combined Score across the dataset. The plot shows how frequently different values of the Combined Score occur, with the Kernel Density Estimate (KDE) curve overlaid to provide a smoothed representation of the data's distribution. The Combined Score is normalized and ranges from 0 to 1, with the histogram revealing the spread and concentration of scores. The plot suggests that most values of the Combined Score are clustered around the mid-range (0.5), with some higher and lower values forming the tails of the distribution.....	270
Figure 28 This plot presents the Kernel Density Estimate (KDE) of Stem Similarity for two subsets of the data: those with NVdB Presence (only) and those with NBC-FL presence (only). The KDE curves allow for the comparison of the distributions of Stem Similarity between these two groups, providing insights into how the presence of NVdB or Nation influences lexical similarity. The graph shows the smooth density of Stem Similarity for each group, with distinct peaks indicating the most	

common similarity values within each category. The plot may reveal that NBC-FL presence is associated with higher Stem Similarity on average compared to NVdB Presence, as indicated by the position and shape of the curves.....	273
Figure 29 Application essential workflow	283
Figure 30 Age statistics	314
Figure 31 Male/ Female statistics	314
Figure 32 Nationality statistics	314
Figure 33 Preferred everyday spoken language	315
Figure 34 When was the last time you studied English?.....	316
Figure 35 Which language do you find easiest to communicate in?.....	316
Figure 36 Speaking and listening are the main areas participants wish to improve.....	317
Figure 37 Would you enjoy learning through topics of personal interest in English?.....	318
Figure 38 Understanding English: How do you define your ability?	318
Figure 39 Writing in English: How do you define your ability?.....	319
Figure 40 Vocabulary Section Question 1.....	320
Figure 41 Vocabulary Section Question 2	320
Figure 42 Vocabulary Section Question 3.....	320
Figure 43 Vocabulary Section Question 4.....	320
Figure 44 Vocabulary Section Question 5.....	321
Figure 45 Vocabulary Section Question 6.....	321
Figure 46 Vocabulary Section Question 7.....	321
Figure 47 Vocabulary Section Question 8.....	321
Figure 48 Grammar test results	323
Figure 49 11-month usage-metrics heatmap.....	342
Figure 50 The heatmap provides an exploratory visualization of how lexical similarity behaves across frequency and developmental dimensions within the annotated lexicon. While the analysis is descriptive in nature, the observed patterns suggest systematic interactions between frequency-based and transparency-based measures of lexical complexity.	346
Figure 51 The heatmap reports the distribution of Italian-English word pairs according to their combined similarity score and their frequency band (Nation levels). The frequency bands are divided into high frequency (1-3,500), mid frequency (3,501-7,500), and low frequency (>7,500). Similarity is categorized into low, medium, and high based on the combined transparency score. The results show that high-frequency vocabulary accounts for the largest portion of the dataset, with 1201 items classified as low similarity, 489 as medium similarity, and 786 as high similarity. This indicates that although high-frequency words are abundant, a substantial subset lacks cross-linguistic transparency, which may contribute to processing difficulty and increased reliance on guessing strategies. Mid-frequency vocabulary is less represented, suggesting that learners interact less with items that are typically assumed to support progressive lexical development. In the low-frequency band, the proportion of high-similarity items (506) is comparable to that of low-similarity items (298), suggesting that transparency can compensate for low corpus frequency. Overall, the heatmap highlights that frequency alone is an insufficient predictor of lexical accessibility for Italian learners of English, and that transparency should be integrated into adaptive sequencing and instructional design.....	347

Tables

<i>Table 1 Overview of studies that have investigated the effects of textual enhancement on vocabulary and collocation learning in different contexts.....</i>	<i>35</i>
Table 2 Knowledge Graphs Key Network Metrics and Their Definitions	136
Table 3 Core resource components supporting adaptive language learning systems and their pedagogical functions.....	140
Table 4 Overview of lexical data categories and their descriptive scope in language resource design.	149
Table 5 Overview of linguistic information and lexical-semantic features reported by the analyzed resources.	149
Table 6 LLOD and non LLOD resources taken into account for the observation.....	171
Table 7 Key linguistic features characteristic of formal and academic language use.....	174
Table 8 Node types and relational structure in a lexical-semantic network representation.	192
Table 9 Example of semantic relations generated by the <code>disambiguate()</code> function. The table lists the target word, its identified sense (based on the Simple Lesk algorithm), and the corresponding semantic relations, synonyms, hypernyms, hyponyms, meronyms, holonyms, and antonyms, used to construct the relation graph.....	194
Table 10 Structured representation of the semantic relations extracted through the <code>disambiguate_endpoint()</code> API. Each entry lists the target word, its definition, and the related lexical items as returned by the Word Sense Disambiguation process.....	196
Table 11 Multilingual lexical alignment table capturing morphological, orthographic, and corpus-based correspondences between Italian terms and their English equivalents, enriched with similarity metrics and resource presence indicators	259
<i>Table 12 This table illustrates a multilingual alignment dataset consisting of Italian terms and their English equivalents, along with computed similarity measures and metadata flags. The variables include both lexical similarity metrics (e.g., levenshtein_normalized, stem_similarity) and database-specific attributes (e.g., presence in vocabulary databases and academic word lists). This combination of morphological, orthographic, and lexical information supports the evaluation of semantic alignment in bilingual terminology extraction and machine translation contexts</i>	<i>265</i>
Table 13 Limitations Introduced by the Java-Based Implementation.....	300
Table 14 Appil lessons' plan overview.....	329
Table 15 Roadmap for the conscientious integration of large language models (LLMs) into the platform. The roadmap outlines a staged approach to incorporating AI-driven capabilities while maintaining pedagogical validity, ethical compliance, and learner-centered design. Phase 1 focuses on strengthening the existing frequency-based lexical research framework and validating text evaluation metrics. Phase 2 introduces LLMs for content generation under strict constraints and human supervision. Phase 3 emphasizes pedagogical validation through teacher collaboration and curriculum alignment. Phase 4 addresses ethical risks, bias monitoring, and GDPR compliance. Phase 5 proposes the deployment of adaptive, personalized learning features, ensuring that AI supports learner agency and motivation rather than replacing them.....	360
Table 16 Profiling assessment questionnaire - This questionnaire has been created solely for organizational and educational purposes related to the English course promoted by the Municipality of Fisciano. The information collected will be used exclusively for educational purposes and will be processed in compliance with the General Data Protection Regulation (GDPR – EU 2016/679).....	429

Formulas

Formula 1 – Normalized frequency metrics representing the proportion of stop words, general vocabulary, and AWL terms within a given text corpus.....	119
Formula 2 Average semantic similarity between words w_i and stems s_i , using a general English vocabulary. Computed over general valid N matches; returns 0 if none found.....	119
Formula 3 Same as above but based on the Academic Word List. Uses $SAWL_{w_i, s_i}$ over $NAWL'$ pairs; returns 0 if no valid matches.....	120
Formula 4 Computes a weighted similarity score by combining general and AWL vocabulary similarities, scaled by their respective lexical densities, to reflect the prominence and impact of different vocabulary types within the text.....	120
Formula 5 Calculates textual density (D) as the ratio of total words to sentences, and lexical uniqueness (U) as the count of distinct word forms, capturing structural and lexical diversity in the text.	121
Formula 6 The formality score (F) as proposed by Heylighen and Dewaele (1999) quantifies linguistic formality by combining the relative frequencies of grammatical categories typically associated with formal versus informal registers. Higher values indicate greater formality in a text.	174
Formula 7 Example of average similarity score computation.....	252
Formula 8 Levenshtein distance measures the minimum number of edits, insertions, deletions, or substitutions, needed to transform one string into another.....	259
Formula 9 Normalizing edit distance to enable consistent similarity scoring between strings of different lengths.....	260
Formula 10 Sequence Matcher Formula: measures the similarity between two-word stems by comparing their longest matching subsequences. Here, M represents the number of matching characters in the optimal alignment, and $ a $ and $ b $ are the lengths of the Italian and English stems, respectively. This normalized ratio provides a value between 0 and 1, capturing partial matches and structural similarities, and is particularly useful for identifying morphologically related words across languages.....	261
Formula 11 This formula calculates the similarity score between two stemmed words, one in Italian and one in English, using a combination of Levenshtein distance (Formula 9) and sequence alignment ratio (Formula 10). The Levenshtein distance measures orthographic similarity, while the sequence alignment ratio captures structural similarity. The two components are weighted by parameters α and β , allowing flexibility in emphasizing one measure over the other. This score provides a comprehensive assessment of word similarity based on both character-level and sequence-level comparisons.....	264

Coding blocks

Code 1 Python function to obtain the stem of the lemma with NLTK Snowball Stemmer Library	242
Code 2 Implementation details of the morphological similarity algorithm. The function first stems each word using the appropriate Snowball stemmer for English and Italian, producing simplified root forms ('root1' and 'root2'). It then applies Python's 'SequenceMatcher' to compare the two stems and compute a similarity ratio based on shared character sequences. The resulting score ranges from 0 (no similarity) to 1 (identical morphology).	242
Code 3 Implementation of the Levenshtein distance algorithm with step tracking. This function, 'levenshtein_steps()', computes the minimum number of edit operations (insertions, deletions, substitutions) required to transform one string (s1) into another (s2). Beyond returning the numerical Levenshtein distance, it reconstructs the specific sequence of operations needed for the transformation. Each operation is described as a human-readable step, such as 'Remove', 'Add', or 'Replace', making the algorithm useful for visualizing string transformations or explaining text similarity computations in educational contexts	245
Code 4 Example of Levenshtein transformation steps between two words. This example demonstrates how the 'levenshtein_steps()' function identifies and lists the precise edit operations required to transform one string ("giocator") into another ("player"). Each step specifies the type of operation: replacement, insertion, or deletion , along with the affected character and its position in the string. This output provides an interpretable view of the morphological and orthographic differences between two lexical items.	246
Code 5 Text preprocessing function for lexical simplification and normalization. The 'preprocess_text()' function performs a series of text-cleaning operations to prepare raw input for linguistic analysis. It first converts text to lowercase and tokenizes it into individual words. Stopwords , common words that carry little semantic weight , are removed to reduce noise. The function then applies the Porter stemming algorithm, generating a list of tuples containing each original word and its corresponding stem. This preprocessing step helps standardize lexical forms and supports downstream tasks such as similarity analysis or vocabulary profiling	248
Code 6 Python function for retrieving semantic similarity from a MySQL database. This function, get_stem_similarity(), queries a MySQL database to obtain semantic similarity values for a given English word or its stem. The procedure first searches for direct matches of the full word in the metrics table. If no records are found, it performs a secondary lookup using the word stem. When multiple similarity values are returned, the function computes their arithmetic mean, providing a single representative similarity score. If no entries exist for either form, the function returns a default value of 0.0, ensuring robust handling of missing data.	249
Code 7 Python function for calculating text-level semantic similarity. This function, calculate_semantic_similarity(), computes the overall semantic similarity of a given text by integrating word-level similarity scores retrieved from a database. The process begins with text preprocessing and stemming to extract unique lexical items. Each word-stem pair is passed to the get_stem_similarity() function, which returns a similarity value based on stored semantic metrics. The cumulative similarity is then normalized by the number of unique stems, producing a score between 0 and 1. This approach provides a scalable metric for assessing the semantic cohesion and lexical relatedness of a text sample.....	250
Code 8 Function for calculating and displaying semantic similarity results in a GUI. The on_calculate_similarity() function serves as a user interface handler for the semantic similarity module. It retrieves user-input text from a graphical entry widget, processes it through the calculate_semantic_similarity() function, and displays the resulting normalized similarity score in a message box. If no text is provided, the function generates a warning prompt. This integration of computational processing with a graphical user interface enhances usability, enabling users to perform semantic evaluations interactively and view immediate feedback.....	251

List of abbreviations

AE — Aural Enhancement	L1 — First Language
AI — Artificial Intelligence	L2 — Second Language
AMTB — Attitude/Motivation Test Battery	L2MSS — L2 Motivational Self System
ANOVA — Analysis of Variance	LLM — Large Language Model(s)
AoA — Age of Acquisition	LLOD — Linguistic Linked Open Data
BNC — British National Corpus	LMM — Linear Mixed(-Effects) Model
CAF — Complexity, Accuracy, Fluency	LOD — Linked Open Data
CALL — Computer-Assisted Language Learning	LRE — Language-Related Episode
CALP — Cognitive Academic Language Proficiency	MALL — Mobile-Assisted Language Learning
CA — Communicative Approach	MTLD — Measure of Textual Lexical Diversity
CBI — Content-Based Instruction	MWE — Multiword Expression
CBLT — Content-Based Language Teaching	NLP — Natural Language Processing
CEFR — Common European Framework of Reference for Languages (<i>Italian: QCER</i>)	NNS — Non-Native Speaker
CFA — Confirmatory Factor Analysis	NS — Native Speaker
CLIL — Content and Language Integrated Learning	PBLT — Project-Based Language Teaching
CLT — Communicative Language Teaching	RQ — Research Question(s)
CMC — Computer-Mediated Communication	SCMC — Synchronous Computer-Mediated Communication
COCA — Corpus of Contemporary American English	SDT — Self-Determination Theory
CPH — Critical Period Hypothesis	SI — Structured Input
DA — Dynamic Assessment	SLA — Second Language Acquisition
DMM — Declarative/Procedural Model	SLL — Second Language Learning
EAP — English for Academic Purposes	SRL — Self-Regulated Learning
EFA — Exploratory Factor Analysis	TBL — Task-Based Learning
EFL — English as a Foreign Language	TBLT — Task-Based Language Teaching
ELT — English Language Teaching	TE — Textual Enhancement
EMI — English-Medium Instruction	TTR — Type-Token Ratio
ESL — English as a Second Language	UG — Universal Grammar
ESP — English for Specific Purposes	VKS — Vocabulary Knowledge Scale
FLA — Foreign Language Anxiety	WSD — Word Sense Disambiguation
FLE — Foreign Language Enjoyment	WTC — Willingness to Communicate
FoF — Focus on Form	ZPD — Zone of Proximal Development
FoFs — Focus on Forms	
GDPR — General Data Protection Regulation	
GJT — Grammaticality Judgment Test	
GLMM — Generalized Linear Mixed(-Effects) Model	
i+1 — Input Hypothesis notation	
ICALL — Intelligent Computer-Assisted Language Learning	
ID — Individual Differences	
IL — Interlanguage	
IPA — International Phonetic Alphabet	
IRT — Item Response Theory	
ISLA — Instructed Second Language Acquisition	
KG — Knowledge Graph	
KRR — Knowledge Representation and Reasoning	

Introduction

The quest for effective second language acquisition (SLA) has captivated educators, linguists, and technologists for decades, as mastering a new language has become an indispensable skill in today's globalized world. The ability to communicate in more than one language not only facilitates personal and professional growth but also enables individuals to engage with diverse cultures and ideas. The challenge of learning a second language, however, is not trivial. Learners are tasked with navigating the complexities of grammar, pronunciation, vocabulary acquisition, and cultural nuances, each of which presents unique obstacles. While language instruction has evolved over time, integrating new technologies, cognitive science insights, and linguistic datasets into second language learning methods has led to a transformative approach in the field. This thesis explores the convergence of these various disciplines to address the challenges and opportunities present in SLA, offering a comprehensive examination of academic literature, innovative tools, and applied strategies aimed at improving the effectiveness of language learning.

Language learning is both a cognitive and social phenomenon that demands a multifaceted approach. One of the major challenges in SLA is vocabulary acquisition, which plays a pivotal role in enabling learners to understand and produce language effectively. Building vocabulary is not simply about memorizing words but involves the complex process of integrating them into a mental lexicon that supports communication and comprehension. Similarly, learners must master the intricacies of grammar and syntax, develop phonetic awareness, and understand cultural contexts. The traditional methods of language teaching, such as rote memorization and teacher-directed instruction, have proven to be insufficient in addressing the diverse needs of modern language learners. Despite the wealth of research and instructional tools available, significant gaps persist in our understanding of how best to support SLA. This thesis seeks to bridge these gaps by synthesizing insights from SLA literature, curriculum design, and cutting-edge linguistic linked open data (LLOD), all while proposing a comprehensive didactic and technical solution through the development of a specific learning platform which will be labelled *Appil*. In this context, the following research questions (RQ) will guide the investigation:

RQ1: How can Natural Language Processing (NLP) and personalized content choices enhance the effectiveness of second language learning, allowing learners to have more control over their learning paths and content selection?

RQ2: Is it possible to create a learning environment that follows a traditional learning framework while still allowing learners the freedom to choose their own content, thereby balancing structure with personalization? The present research is structured around three interrelated scientific fields aimed at improving second language learning. The first focus lies in a comprehensive review of SLA

literature, which serves as the theoretical foundation for this study. This part of the thesis explores key topics in SLA research, including lexical acquisition, vocabulary size, and the cognitive processes involved in language learning. An in-depth examination of how learners acquire and process vocabulary provides the basis for understanding how language proficiency develops over time. Moreover, the discussion addresses the challenges inherent in curriculum design for SLA, particularly how instructional materials can be optimized for diverse learners. Topics such as text enhancements, grammar instruction, and the role of computer-assisted language learning (CALL) are examined, offering a thorough understanding of current SLA methodologies and highlighting potential areas for advancement.

A significant component of this research investigates the role of Linguistic Linked Open Data (LLOD) in language learning. LLOD represents an emerging domain that combines open data, linguistic resources, and computational methods to create interconnected datasets capable of enhancing digital language learning tools. The study examines key LLOD resources such as Princeton WordNet, *BabelNet*, and *Wiktionary* (Miller, 1995; Navigli & Ponzetto, 2012; Sérasset, 2015), evaluating their potential to support vocabulary acquisition, translation, and other language learning processes. Attention is also given to the OntoLex model, which structures lexical information in a machine-readable format, and to the VarTrans module, which addresses linguistic variation and translation across languages. These datasets and models are analyzed in terms of their contribution to second language acquisition (SLA), with specific emphasis on their capacity to foster adaptive and data-informed learning resources. Furthermore, the research identifies several challenges associated with integrating LLOD into language learning environments, including issues of accessibility, data interoperability, and the representation of linguistic diversity across different languages and dialects. Building on this theoretical and analytical foundation, the study also focuses on the development of Appil, a web-based application designed to integrate LLOD resources within a structured pedagogical framework. The platform developed for this study (hereafter, the learning platform) aims to enhance the language learning experience through dynamic exercises, real-time feedback, and adaptive learning pathways. Particular emphasis is placed on self-regulated learning, enabling users to advance at their own pace while receiving personalized recommendations derived from their performance data. The technical infrastructure incorporates linguistic resources from the LLOD ecosystem, allowing learners to engage with rich, data-driven content that mirrors authentic language use. In addition, mechanisms for text complexity analysis and automated exercise generation ensure that learners are continually challenged at appropriate levels of difficulty. Considerable attention is devoted to user interface design, ensuring accessibility, clarity, and engagement for a diverse range of learners. This thesis adopts a mixed-methods approach, integrating systematic literature review, computational analysis, and didactic design to address its central research questions. The work is

organized into six core chapters, each exploring a distinct yet interrelated dimension of Second Language Acquisition (SLA) and technology-enhanced learning.

The thesis unfolds through six interconnected chapters, each progressively building from theoretical foundations to practical implementation and empirical evaluation.

Chapter 1 establishes the theoretical foundation of the study, providing a comprehensive overview of SLA theories and models. It examines key frameworks, including Krashen's Input Hypothesis, Swain's Output Hypothesis, Long's Interaction Hypothesis, and Chomsky's Universal Grammar, alongside more contemporary perspectives such as Ellis's Task-Based Pedagogy and Nation's focus on vocabulary acquisition. The chapter concludes with a critical analysis of cognitive factors, CEFR-based frameworks, and challenges in curriculum design. (Chomsky, 1986; Council of Europe, 2020; Ellis, 1994; Krashen, 1982; Long, 1996)

Chapter 2 investigates innovations in technology-mediated language learning, tracing the evolution from traditional Computer-Assisted Language Learning (CALL) to Intelligent CALL (ICALL). It evaluates existing tools and systems, such as *SMILLE* and *Revita* (Katinskaia & Yangarber, 2018; Zilio et al., 2017), discussing their methodologies, strengths, and limitations. The chapter also explores current challenges in digital pedagogy and outlines opportunities for more adaptive, data-informed approaches to SLA.

Chapter 3 bridges the fields of didactics and computer science, establishing the methodological backbone of the research. It elaborates on self-regulated learning, adaptive algorithms, and multimodal integration, while also detailing the role of Linked Linguistic Open Data (LLOD), ontologies, and network analysis in supporting adaptive content generation. The chapter culminates in a systematic literature review of major linguistic databases such as *WordNet*, *BabelNet*, and *DBnary*, positioning them as crucial resources for scalable language learning systems (Miller, 1995; Navigli & Ponzetto, 2012; Sérasset, 2015).

Chapter 4 introduces Appil, the adaptive language learning platform developed within this research. It outlines its design philosophy, core features, and pedagogical strategies, integrating SLA principles with computational methods. The platform's complexity assessment index, dynamic text generation, and CEFR-aligned progression model are discussed in detail, illustrating how theory and technology converge to create personalized learning experiences.

Chapter 5 focuses on the technical infrastructure and system architecture underlying Appil. It describes the evolution from early prototypes to a modular, microservices-based architecture, emphasizing scalability, interoperability, and performance optimization. The chapter also reflects on pedagogical alignment, usability testing, and planned enhancements involving AI integration, learning analytics, and institutional scalability.

Chapter 6 presents the fieldwork and empirical evaluation of the Appil platform, developed and supervised within the framework of this PhD project. It outlines the study design, participant profiles, diagnostic assessments, and the didactic implementation of instructional sequences. The analysis of learner data, usage metrics, and feedback seeks to provide insights into the system's pedagogical effectiveness, while also identifying areas that may require further refinement. In doing so, the chapter aims to offer a balanced perspective on Appil's current potential and its developmental limitations. By synthesizing linguistic theory, computational modeling, and pedagogical innovation, this thesis contributes to the ongoing evolution of SLA methodologies. The integration of LLOD and adaptive technologies opens new avenues for developing data-driven, context-aware, and inclusive language learning environments. Ultimately, the newly created Appil platform exemplifies how research can be transformed into practical, learner-centered applications, bridging the gap between theory and practice. This work lays the groundwork for future explorations into intelligent, scalable, and culturally sensitive approaches to second language learning in the digital age.

1 Investigating the Foundations of Second Language Learning

1.1 Introduction

Learning a foreign language involves a wide range of cognitive processes, activating the learner on psychological, behavioral, and, even at times, physical levels.

This chapter reviews key theoretical perspectives and empirical research that inform the field of Second Language Acquisition (SLA). Drawing from linguistics, cognitive psychology, and applied pedagogy, the chapter is structured around foundational domains including implicit/explicit learning, grammatical competence, vocabulary acquisition, morpho-syntax, semantics and pragmatics, phonology, proficiency frameworks, and the integration of educational technology. Together, these frameworks support the development of an adaptive, learner-centered approach to SLA.

Building on this framework, the field of didactics has developed strategies and tools aimed at optimizing language acquisition and supporting learners in achieving communicative competence more effectively. The following sections examine lexical acquisition, with specific attention to how vocabulary size, depth, and frequency influence learning outcomes, and how these factors interact with cognitive processes such as memory and attention. Insights from behavioral studies are also reviewed to shed light on the psychological mechanisms that govern language processing and acquisition.

The chapter then turns to the technological and pedagogical challenges associated with Computer-Assisted Language Learning (CALL) systems and their extensions into Intelligent CALL (ICALL) (Chapelle, 2001; Levy, 2008; Schulze, 2008). A critical evaluation of current tools and infrastructures highlights their strengths and limitations, providing the rationale for the system proposed in this research. Overall, the literature review seeks to ensure that the technical solution put forward is not only innovative but also firmly grounded in empirical evidence and pedagogical theory.

1.2 Understanding Core Constructs and Competence in SLA

A central aim of SLA research is to translate theoretical constructs from cognitive psychology into adaptive learning environments (DeKeyser, 2007; VanPatten & Williams, 2014). Applied Linguistics shares this focus, emphasizing effectiveness and accessibility through technology. Digital tools now complement structured instruction by offering personalization, feedback, and flexible pathways.

However, while autonomy in informal learning broadens opportunities, it also risks issues such as misaligned content and cognitive overload, challenges less common in guided curricula.

Language acquisition is commonly divided into grammatical and lexical domains. Grammar encompasses syntax, morphology, and sentence structure (R. Ellis, 1994), while vocabulary concerns the breadth, depth, and retention of words (Nation, 2001). Though conceptually distinct, these domains are deeply interrelated. Additional fields contribute to SLA: phonetics and phonology address sound systems (Ladefoged & Johnson, 2014), morphology concerns words (Aronoff & Rees-Miller, 2003), syntax governs sentence structure (Chomsky, 1981; Searle, 1969), semantics deals with meaning (Cruse, 2000), and pragmatics focuses on context in communication (Searle, 1969). Collectively, these domains shape communicative competence.

Whereas native speakers acquire these areas largely implicitly (Chomsky, 1981; Pinker, 1994), second language learners often require explicit instruction (Ellis, 2006). Teaching is especially important for pronunciation (Celce-Murcia et al., 2010), morphology, syntax, semantics (Gass et al., 2008; Schmitt, 2008), and pragmatics, where cultural nuances are not intuitively grasped (Thomas, 1983). This distinction underscores the need for well-designed curricula that integrate these domains explicitly.

The Common European Framework of Reference for Languages (Council of Europe, 2001, 2020) links linguistic competence to five skills: listening, reading, writing, speaking, and spoken interaction. Phonetics and phonology support listening and speaking, morphology and syntax underpin written and spoken production, semantics contributes to all modalities, and pragmatics shapes interaction. The CEFR also emphasizes gradual progression (A1–C2), with learners evolving from basic patterns to complex structures and pragmatic strategies.

Krashen's hypotheses (1981, 1982, 1985) remain central to SLA. His Acquisition–Learning distinction differentiates subconscious acquisition from conscious learning, while the Input Hypothesis highlights the importance of comprehensible input ($i+1$). The Natural Order Hypothesis suggests grammatical structures are acquired in predictable sequences (Dulay & Burt, 1974), and the Monitor Hypothesis frames conscious knowledge as an editing tool. The Affective Filter Hypothesis (Krashen, 1982) stresses the role of motivation, confidence, and anxiety in enabling or blocking input. These principles collectively argue for meaningful communication, scaffolded challenge, and attention to affective factors in SLA.

Building on this foundation, the present research prioritizes lexical and grammatical acquisition, aligning pedagogical theory with adaptive digital solutions that foster both receptive (listening, reading) and productive (speaking, writing) skills. Vocabulary, as a cross-cutting competence, is central to this integration.

1.3 Krashen's Monitor Model and the Power of Input

Stephen Krashen's Monitor Model (1981, 1982, 1985) represents one of the most influential frameworks in the field of Second Language Acquisition (SLA), particularly in the latter part of the 20th century. Formulated as a collection of five interrelated hypotheses, the Acquisition–Learning Hypothesis, the Input Hypothesis, the Natural Order Hypothesis, the Monitor Hypothesis, and the Affective Filter Hypothesis, the model redefined theoretical assumptions about how languages are learned and provided a foundation for communicative and immersion-based language teaching approaches. At the core of the model is a strong emphasis on the role of naturalistic, meaning-focused input and the minimization of formal grammar instruction. Krashen's work emerged in response to behaviorist and structuralist models of language instruction, which had traditionally emphasized accuracy, grammatical correctness, and drill-based learning. In contrast, Krashen advocated for more learner-centered, communicative environments, where exposure to comprehensible input drives acquisition without the need for explicit linguistic analysis (Krashen, 1981).

The Acquisition–Learning Hypothesis serves as the foundation of the Monitor Model (Krashen, 1982), asserting that there are two distinct processes involved in developing language competence. Acquisition is described as a subconscious process akin to how children acquire their first language, relying on exposure to language in meaningful, communicative contexts. In contrast, learning is a conscious process involving explicit instruction and metalinguistic awareness of grammar rules (Krashen, 1981). Krashen posits that only acquisition leads to fluent and spontaneous language use, while learned knowledge may serve merely as a *monitor* to edit output in specific situations, namely when learners have sufficient time, are focused on form, and know the relevant rules (Krashen, 1982). This dichotomy introduced a fundamental shift in how SLA was conceptualized, encouraging educators to focus on immersive, input-rich environments rather than rule-based instruction.

Expanding on this foundation, the Input Hypothesis is arguably Krashen's most widely cited and pedagogically impactful idea. It proposes that language is acquired through exposure to input that is comprehensible to the learner, but slightly beyond their current level of competence, termed 'i+1' (Krashen, 1985). In this model, 'i' represents the learner's current interlanguage, and '+1' symbolizes the next step in their developmental sequence. The role of the teacher, then, is not to explain rules, but to provide input that is meaningful, contextualized, and slightly challenging. This emphasis on input rather than output significantly influenced the development of Communicative Language Teaching (CLT), content-based instruction, and extensive reading approaches, which prioritize comprehension over grammatical accuracy in early stages of learning.

Complementing the Input Hypothesis, the Natural Order Hypothesis posits that grammatical structures are acquired in a relatively fixed and predictable sequence, regardless of the order in which they are taught. This hypothesis draws on studies by Dulay and Burt (1974), who observed similar patterns in the acquisition of English morphemes among learners from different linguistic backgrounds. Krashen (1982) interpreted this consistency as evidence of an internal, biologically determined sequence of language development, thereby diminishing the efficacy of teaching grammatical rules out of order. While the Natural Order Hypothesis provided theoretical support for delaying explicit grammar instruction, it has also been criticized for overlooking individual variability and contextual factors that affect acquisition sequences.

The Monitor Hypothesis further develops the distinction between acquisition and learning by describing the limited role of learned knowledge in language performance. According to Krashen (1981), learned rules function as a conscious monitor, allowing speakers to make corrections under specific conditions. However, because the monitor requires time, attention to form, and explicit knowledge of rules, it is generally not useful in real-time communication. Krashen warned that overuse of the monitor could interfere with fluency, leading to hesitation and unnatural language production, particularly among anxious or perfectionist learners.

Finally, the Affective Filter Hypothesis introduces psychological and emotional variables into the SLA process. Krashen (1982) argues that affective factors such as motivation, self-confidence, and anxiety can act as a *filter* that either facilitates or inhibits the processing of comprehensible input. A low affective filter enables more input to be internalized, whereas a high filter blocks input from reaching the language acquisition device. This hypothesis helped popularize humanistic and learner-centered approaches to teaching, which seek to reduce anxiety and build supportive learning environments that foster risk-taking and engagement.

Although Krashen's model has been widely adopted in SLA pedagogy, it has also attracted significant criticism. McLaughlin (1987) argued that the acquisition–learning dichotomy is conceptually problematic, suggesting that the cognitive processes involved in language development are more integrated and dynamic than Krashen's binary framework allows. Swain (1985) challenged the Input Hypothesis by introducing the Output Hypothesis, which posits that language production is not merely the result of input, but a critical component of acquisition, particularly for developing grammatical accuracy. Long (1996) built on this critique by proposing the Interaction Hypothesis, which asserts that conversational interaction, especially when it involves negotiation of meaning, plays a central role in language development. These perspectives emphasize that comprehensible input, while necessary, is not sufficient, and that opportunities for output, feedback, and interaction are essential to SLA.

Additionally, critics have pointed out that the Natural Order Hypothesis (Krashen, 1982) fails to account for differences across learner populations, such as first language background, age, aptitude, and learning environment (Ellis, 1994). Others have noted that the concept of ‘i+1’ is difficult to measure and implement in classroom settings, making it less practical as a guiding instructional principle (Lightbown & Spada, 2006). Despite these critiques, Krashen’s Monitor Model has had an enduring influence on language teaching methodologies. Its legacy can be seen in the design of immersion programs, reading-based curricula, and task-based learning, all of which prioritize input and affective engagement over explicit grammar instruction.

In conclusion, while Krashen’s Monitor Model may not provide a comprehensive or empirically flawless account of SLA, it has been instrumental in shifting the field away from decontextualized grammar instruction toward a more communicative, learner-centered paradigm. The model’s emphasis on comprehensible input, emotional readiness, and naturalistic exposure to language continues to inform pedagogical practices and remains a key point of reference in both SLA research and classroom application.

1.4 Swain’s Output Hypothesis: The Role of Comprehensible Output

Building on Krashen’s input-centered account, which places comprehensible input at the core of acquisition, Merrill Swain offers a necessary complement by arguing that input alone cannot fully explain the development of advanced linguistic competence (Swain, 1995). Her Output Hypothesis, first articulated in the mid-1980s and refined in subsequent decades, positions language production, speaking and writing, as a causal driver of acquisition rather than a mere byproduct of it. Swain’s insight emerged from empirical observations of Canadian French immersion classrooms: despite extensive exposure to comprehensible input, learners often attained high receptive proficiency yet plateaued in productive accuracy and complexity, particularly in morphology and syntax (Swain, 1985, 1995). This discrepancy suggested that learners must also be pushed to produce comprehensible output, language that not only conveys meaning but is shaped by communicative pressure to be accurate and clear, so that gaps in their interlanguage become visible and revisable.

In Swain’s formulation, output serves three interlocking functions that collectively advance acquisition. First, the noticing or triggering function captures how attempts to express intended meanings expose the limits of current knowledge; when learners cannot say what they want, they become aware of specific holes in grammar, lexis, or discourse patterns, and this awareness sensitizes them to subsequent input that could resolve the gap (Swain, 1995). Second, the hypothesis-testing function frames production as a site where learners try out candidate forms under real communicative

conditions; interlocutor feedback, whether explicit correction or the subtler work of negotiation, confirms or disconfirms these hypotheses, prompting adjustments to the developing system. Third, the metalinguistic or reflective function highlights how collaborative speaking and writing encourage learners to talk about language itself, fostering analytic awareness that supports both explicit understanding and, indirectly, the growth of implicit knowledge. In this way, output deepens processing, recruits feedback, and organizes reflection, thereby transforming exposure into development.

Swain's proposal sits naturally alongside interactionist work, particularly Long's Interaction Hypothesis, which emphasizes the role of negotiation of meaning in making input comprehensible and directing attention to form. Output is the learner's contribution to that negotiation: attempts to clarify, repair, or elaborate utterances enact the very noticing, hypothesis testing, and feedback integration that propel change. The theoretical picture that emerges is not a contest between input and output, but a dynamic cycle, input affords data, output tests and reshapes it, and interaction orchestrates the conditions under which both become developmentally potent (Long, 1996).

Pedagogically, Swain's account argues against input-only designs that cultivate comprehension yet leave productive competence underdeveloped (Swain, 1985; Swain, 1995). Instruction should therefore engineer tasks that *push* output, information-gap and problem-solving activities, role-plays with clear communicative stakes, collaborative writing that requires joint drafting and revision, and peer discussion where accountability for both meaning and form is explicit (Swain & Lapkin, 1995; Long, 1996). Such tasks should be coupled with timely feedback, recasts, prompts, or guided self-repair, so that learners receive usable evidence about the adequacy of their hypotheses without derailing communication (Lyster & Ranta, 1997). In technology-mediated environments, these principles translate into interactive platforms that elicit spoken and written production, provide immediate, focused feedback, and support iterative cycles of attempt, reflection, and retry; done well, they extend the reach of comprehensible input with structured opportunities to transform understanding into use (Swain, 2005; Chapelle, 2003).

Critiques of the Output Hypothesis have refined rather than overturned it. Some argue that for strictly comprehension-oriented goals, output may be less central; others warn that unguided production risks stabilizing errors unless feedback is consistently available. These concerns underscore the conditional nature of output's benefits and reinforce the need to design tasks where communicative pressure is balanced by scaffolding and where feedback is both comprehensible and developmentally appropriate. The broader consensus in SLA now treats input, output, and interaction as mutually necessary conditions whose synergies, rather than any single mechanism, best account for robust development.

Swain's legacy, then, is to have rebalanced the field's explanatory toolkit. By demonstrating that production can trigger noticing, enable hypothesis testing, and cultivate metalinguistic reflection, she supplied a conceptual and empirical basis for integrating output into curricula in ways that complement input-rich approaches. In doing so, her work ensures that acquisition is understood not only as the internalization of comprehensible input, but as an intrinsically dialogic process in which learners actively reshape their interlanguage through purposeful communication and responsive feedback.

1.5 Long's Interaction Hypothesis: Language Learning Through Exchange

Michael H. Long's Interaction Hypothesis reframed second language acquisition (SLA) as a fundamentally dialogic process: learners do not merely absorb language from input; they co-construct it with interlocutors through conversational work that makes input comprehensible, draws attention to form, and elicits feedback that reshapes interlanguage. Long's early formulation (1983, 1985) argued that opportunities to negotiate meaning, through clarification requests, confirmation checks, and comprehension checks, are pivotal because they lead to modified input and modified interaction, increasing the likelihood that learners will process linguistically relevant features. In his later "version 2" (1991, 1996), Long integrated a sharper psycholinguistic rationale: interaction facilitates acquisition because it orchestrates attention, supplies both positive and negative evidence, and enables brief, need-driven focus on form within communicative activity. The result is a theory that reconciles the power of input (Krashen, 1982), the necessity of output (Swain, 1985, 1995), and the engine of interaction that binds them.

At its core, the Interaction Hypothesis makes three linked claims. First, conversational interaction creates learning opportunities that raw exposure rarely affords. When breakdowns or ambiguities arise, interlocutors slow down, rephrase, emphasize key words, and reconfigure syntax; learners ask for clarification or confirmation. These routines tailor input to immediate comprehension needs while preserving semantic content. Second, interaction provides contingent feedback, recasts, prompts, and signals of misunderstanding, that functions as negative evidence about what does not work and positive evidence about what does. Third, precisely because these moves are tied to the learner's current goals and problems, they render linguistic features psychologically real at the moment of need, increasing the likelihood of noticing (Schmidt, 1990, 2001) and subsequent restructuring.

The micro-mechanics of negotiation have been documented across native speaker–learner and learner–learner dyads (Gass & Varonis, 1985; Pica, 1994). Typical sequences begin with a trigger (a

problematic utterance or comprehension failure), followed by indicators (e.g., *Pardon?*, *Do you mean...?*), responses (repetitions, paraphrases, slower articulation), and resolutions that confirm mutual understanding. These turns do more than repair meaning they recalibrate linguistic form, spotlighting morphology (e.g., *tense marking*), lexical choice, and syntax (e.g., *word order*). Crucially, learners are not passive recipients. By attempting to restate, guess, or elaborate, they produce ‘pushed output’ (Swain, 1985), testing hypotheses about form under communicative pressure. Interaction thus yokes comprehensible input, pushed output, and feedback, the canonical triad that contemporary SLA treats as mutually necessary conditions for development.

Feedback within interaction varies in explicitness and effect. Recasts reformulate a learner’s error while preserving meaning (*He goes yesterday?* → *He went yesterday.*), offering a low-interruption path to revised form–meaning mappings. Prompts (elicitation, metalinguistic cues) push learners to self-repair, often deepening processing and retrieval. Classroom studies (Lyster & Ranta, 1997) and lab interaction research (Mackey, 1999; Ellis, Loewen, & Erlam, 2006) suggest both can facilitate acquisition, with outcomes moderated by the target feature’s salience, proficiency level, and task conditions. Recasts are efficient and unobtrusive but sometimes under-noticed (Philp, 2003), while prompts may yield more overt uptake at the cost of fluency if overused. The interactionist answer is not to canonize one feedback type, but to fit feedback to feature and context: morphology and phonology often benefit from salient recasts; syntax and lexis may respond well to prompts that require reformulation.

Long’s (1996) elaboration places attention at the center of the mechanism. Interaction draws attention to trouble spots at precisely the moment comprehension demands resolution, satisfying the conditions for Schmidt’s Noticing Hypothesis (Schmidt, 1990) and enabling memory processes that convert fleeting exposure into durable representation. Moreover, interaction supplies both input flood for target features (through repeated, tailored models) and contingent practice (through pushed output and feedback), aligning with skill-acquisition perspectives without reducing SLA to procedural skill alone. Importantly, Long differentiated focus on form, brief, reactive attention to linguistic elements arising from communicative need, from focus on forms (pre-taught, decontextualized grammar points). In interactive settings, form is not an agenda item; it is a resource recruited to solve communicative problems, which is why learners are more likely to process it.

Evidence for interaction’s benefits now spans modalities and methods. Early descriptive and quasi-experimental studies (Long, 1983; Pica, 1994) have been complemented by controlled experiments linking interactional conditions to gains on targeted structures (Mackey, 1999) and by meta-analytic syntheses indicating moderate positive effects of interaction on L2 development (Mackey & Goo, 2007). Gass and Mackey’s program (Gass & Mackey, 2000, 2006, 2015) has tied specific interactional

moves to later recognition and production gains, bolstering the causal claim. While effect sizes vary, by structure, proficiency, task, and feedback type, the cumulative finding is robust: who learners talk to, how they negotiate, and what feedback they receive matter for what they ultimately acquire.

The Interaction Hypothesis also generalizes beyond face-to-face talk to technology-mediated settings. In synchronous computer-mediated communication (CMC), whether text-based or voice-based, learners negotiate meaning through typed clarification requests and reformulations; interaction logs reveal dense clusters of focus on form with immediate or delayed uptake (Smith, 2003; Blake, 2000; Sauro, 2009). In asynchronous forums, the additional time affords deeper reflection and metalinguistic discussion, supporting the reflective function of output (Hampel & Stickler, 2012; Lee, 2008). These environments expand interaction opportunities, diversify interlocutors, and lower affective barriers, while enabling fine-grained analytics of learner behavior (e.g., *who noticed what, when, and which feedback types persist*). For Intelligent Computer-Assisted Language Learning (ICALL), the interactionist lens translates into design imperatives: detect communication breakdowns, surface just-in-time recasts or prompts, and recycle targets across tasks until performance stabilizes (Heift & Schulze, 2015; Chapelle, 2009).

critiques of the Interaction Hypothesis have pushed the theory toward greater precision rather than overturning it. Conversation-analytic work (Seedhouse, 2004) cautions against treating *negotiation* as a monolithic construct; interactional practices vary by activity type, for example, task-based exchanges versus IRF (Initiation–Response–Feedback) sequences, so theoretical claims must be context-sensitive. From a sociocultural perspective, Lantolf and Thorne (2006) contend that development is mediated by tools, artifacts, and social relations, emphasizing scaffolding and internalization over information-processing metaphors. Interactionist researchers have largely embraced these insights, broadening the framework to include mediation and the co-construction of opportunities for attention and feedback (Ohta, 2001; Mackey & Gass, 2015). A methodological critique notes that uptake does not necessarily equate to acquisition; in response, studies have incorporated delayed posttests, multiple measures (comprehension, production, controlled and free), and designs that assess the durability of change (Ellis, 2010; Loewen, 2005). Finally, not all linguistic structures are equally amenable to interactional treatment, low-salience morphology, for instance, may escape detection in spontaneous talk, prompting the use of input enhancement or planned focus on form alongside reactive feedback (Doughty & Williams, 1998; Long, 2015).

Pedagogically, Long's ideas crystallize in Task-Based Language Teaching (TBLT) as a delivery system for interaction. Tasks with clear communicative outcomes create authentic needs to understand and be understood; they naturally elicit negotiation and feedback. Designers can tune task complexity and grant planning time to manage the fluency–accuracy–complexity trade-offs, then

exploit post-task phases for concise, targeted focus on form. Instructors orchestrate interactional architecture: pairing patterns (NS–NNS, NNS–NNS), group sizes, and roles that maximize participation; training learners to use clarification requests and confirmation checks; and calibrating feedback to preserve communicative momentum while cueing form. Assessment follows suit, sampling task performance and tracking whether interaction-dependent features improve in accuracy and accessibility over time.

For an adaptive, web-based system such as the one envisioned in this dissertation, the Interaction Hypothesis provides concrete design requirements. The learning platform should engineer negotiation opportunities within listening and reading tasks, for instance, through interactive transcripts or glosses that invite reformulation, and within speaking and writing modules, where role-plays and problem-solving activities allow ASR and NLP-driven feedback to deliver salient recasts or prompts (Heift & Schulze, 2015; Li, Link, & Hegelheimer, 2015). The system should log interactional events, clarification requests, feedback types, and self-repairs, using this data to update a learner model that predicts readiness for specific form–meaning mappings (Chapelle, 2009; Wible & Tsao, 2010). Adaptive focus on form can then be scheduled dynamically: when a learner repeatedly encounters or produces a structure in negotiated sequences, the learning platform can insert micro-lessons or recycled tasks to consolidate that feature (Ellis, 2017; Long, 2015). Finally, fluency cycles with familiar materials ensure that newly noticed forms become automatized, closing the loop from interaction to stable competence, a process that embodies the interactionist ideal within an intelligent CALL environment.

In summary, Long's Interaction Hypothesis locates the engine of SLA in the give-and-take of communication. Interaction turns input into learnable data by aligning attention with need; it transforms output from mere performance to hypothesis testing; and it renders feedback developmentally potent by making it contingent and timely. The theory's enduring value is practical: it tells us not only that interaction matters, but which parts of interaction to cultivate and how to design classrooms and technologies so that negotiation, feedback, and focus on form do the heavy lifting of acquisition.

1.6 Chomsky's Universal Grammar and its Role in SLA

While Krashen emphasized the role of environmental input, affective variables, and subconscious learning in SLA, Chomsky advanced a nativist model centered on innate cognitive structures and Universal Grammar (UG).

Noam Chomsky's contributions to linguistics and language acquisition theory have had a transformative effect on the field, particularly through his proposal of Universal Grammar and the nativist framework it embodies. Emerging in the mid-20th century as a reaction against the behaviorist models dominant at the time, Chomsky's work introduced a fundamentally different understanding of how language is acquired, emphasizing the role of innate cognitive structures over environmental conditioning. Although his focus was initially on first language acquisition (FLA), Chomsky's theories have also shaped the direction of second language acquisition (SLA) research, particularly in debates around the extent to which language learning is biologically determined versus learned through interaction and input (Chomsky, 1965, 1986, 1995).

The so-called nativist turns in linguistics began with Chomsky's critique of behaviorism, most famously articulated in his review of B. F. Skinner's *Verbal Behavior* (Chomsky, 1959). In this seminal work, Chomsky challenged the notion that language could be explained as a system of learned habits formed through imitation and reinforcement. He argued that behaviorist theory failed to account for the creativity of language use, the rapid pace of language acquisition in children, and the ability of individuals to generate and understand sentences they had never heard before. Central to Chomsky's argument was the poverty of the stimulus, the idea that the linguistic input children receive is insufficient to explain the detailed and uniform grammatical knowledge they ultimately develop. From this, he concluded that children must be born with an innate linguistic capacity, which he later formalized as Universal Grammar (Chomsky, 1965).

Universal Grammar is defined as a set of structural principles and constraints that underline all human languages. According to Chomsky, UG constitutes a biologically determined system that enables children to acquire any natural language to which they are exposed, provided that input falls within a critical period of cognitive development (Chomsky, 1981, 1995). The language acquisition device (LAD), a hypothetical cognitive module within the human mind, is thought to activate UG principles and facilitate the formation of internal grammatical representations based on exposure to external linguistic input. This model helps explain how children can construct complex syntactic systems despite being exposed to fragmented, inconsistent, and even ungrammatical speech. The fact that language acquisition follows a relatively predictable trajectory across cultures and linguistic environments lends support to the universality of the mechanisms Chomsky describes.

In the context of second language acquisition, Chomsky's UG theory has inspired a substantial body of research investigating whether and how UG remains accessible beyond the critical period associated with first language development. Some researchers argue that adult SLA is constrained by maturational limits, suggesting that UG is only fully accessible during early childhood (Lenneberg, 1967; Johnson & Newport, 1989). Others posit that while UG may not operate identically in L2

acquisition, it still plays a role in shaping the way learners interpret and construct grammars in a second language (White, 1989; Schwartz & Sprouse, 1996). This line of inquiry has led to the development of generative approaches to SLA, which seek to identify how interlanguage grammars reflect principles and parameters derived from UG, and to what extent L2 learners can reset parameters in response to new linguistic input.

However, UG-based accounts of SLA have also been critiqued for underestimating the role of input, social interaction, and cognitive processing in language learning. Researchers working within usage-based (Tomasello, 2003), connectionist (Ellis, 2006), and sociocultural (Lantolf, 2000) paradigms argue that general cognitive mechanisms, such as pattern recognition, frequency sensitivity, and social scaffolding, are sufficient to explain language acquisition without invoking innate grammar. Critics have also noted that many of the assumptions underlying UG, such as the universality of specific syntactic structures or the strict modularity of the LAD, remain empirically difficult to verify. Nonetheless, Chomsky's theories continue to provide a foundational reference point in both FLA and SLA research, serving as a theoretical counterweight to empiricist and constructivist models.

In summary, Chomsky's theory of Universal Grammar marked a paradigm shift in the study of language acquisition by introducing the idea that language ability is hardwired into the human brain. While debates persist regarding the applicability of UG to second language learning, its conceptualization of language as an innate, rule-governed system continues to inform generative SLA research. Even in contexts where alternative frameworks are favored, Chomsky's influence remains unmistakable in ongoing discussions about the nature of linguistic competence, the structure of the mind, and the mechanisms by which language is learned.

1.7 From Instructed Learning to Task-Based Pedagogy: The Work of Rod Ellis

Krashen's hypotheses reoriented SLA toward meaning, arguing that progress is driven by exposure to comprehensible input under low-anxiety conditions and that conscious knowledge plays at best a monitoring role (Krashen, 1981, 1982, 1985). Yet precisely because Krashen cast acquisition as largely independent of explicit instruction, his model left open a practical question for classrooms and technology-mediated environments: what, exactly, can instruction do, and how should it be designed when time, syllabi, and assessment are inevitable constraints? This is where Rod Ellis's work becomes pivotal. While accepting that input and developmental readiness matter, as suggested by the Natural Order and supported by teachability considerations, Ellis systematically investigated how instruction can expedite acquisition, when attention to form is beneficial, and how tasks can be

engineered to create the right cognitive and interactional conditions for change (Ellis, 1994, 2001, 2003). In this sense, Ellis complements rather than contradicts Krashen: he translates input-centric principles into a practicable pedagogy that aligns classroom realities with psycholinguistic processes. Rod Ellis's contribution to Second Language Acquisition lies in showing that instruction, intelligently designed, can accelerate development without violating the internal logic of acquisition. His early synthesis established that learners proceed through relatively robust developmental sequences, but that these sequences are not immune to pedagogic influence; instruction can heighten the salience of target forms, strengthen form–meaning mappings, and provide opportunities for production and feedback that the ambient input often undersupplies (Ellis, 1994, 2008). Rather than assuming that more explanation yields more acquisition, Ellis reframed the teacher's task as the orchestration of conditions, input, interaction, output, and timely attention to form, under which implicit knowledge can grow.

A central thread in Ellis's scholarship is the distinction between focus on forms and focus on form. The former organizes teaching around isolated grammar points, typically detached from communicative purpose; the latter inserts brief, need-driven episodes of form attention into meaning-focused activity (Ellis, 2001). This shift is not semantic. In classrooms where meaning takes priority, moments of uncertainty, error, or communicative pressure create powerful opportunities for noticing. Ellis argues that such episodes, whether proactively seeded through input enhancement or reactively triggered through feedback, are most productive when they are contingent on learner need and embedded in genuine tasks. Here, 'form' is not an object of study for its own sake, but a resource marshalled to achieve communicative outcomes.

This logic culminates in Ellis's articulation of Task-Based Language Teaching (TBLT), where the curriculum is organized around tasks that approximate real-world goals and demand the negotiation of meaning (Ellis, 2003). Tasks are not thinly disguised grammar exercises; they are vehicles that mobilize comprehension, planning, and production in ways that expose learners to richer input and push them to stretch their interlanguage. By manipulating task complexity and granting planning time, teachers and designers can shape the balance between fluency, complexity, and accuracy (Ellis, 2005). Post-task phases then legitimize concise form-focused work, brief explanations, practice, or feedback, so that the linguistic payoffs of communication are consolidated without eclipsing it. In Ellis's hands, TBLT is not a rejection of grammar but a recontextualization: grammar support is delivered after and because communicative need has made it psychologically real.

Ellis's account of implicit and explicit knowledge provides the cognitive backbone for these design choices. He adopts a 'weak interface' position: explicit knowledge does not transmute directly into implicit competence, yet it can guide attention, hypothesis formation, and practice conditions under

which implicit knowledge is refined (Ellis, 2009). Crucially, he operationalizes these constructs with measurement tools, timed and untimed judgments, elicited imitation, metalinguistic tasks, so that claims about instructional effects can be tested rather than asserted. This methodological clarity has helped the field move beyond slogans toward cumulative evidence.

Within this architecture, corrective feedback becomes a principled tool rather than an annoyance. Ellis and colleagues show that both recasts and prompts can facilitate acquisition when they are comprehensible, immediate, and targeted, with their relative effectiveness mediated by the linguistic feature at stake and the learner's proficiency (Ellis, Loewen, & Erlam, 2006). The point is not to police every error but to leverage errors as gateways to noticing, keeping feedback light enough to preserve communicative momentum yet salient enough to trigger restructuring.

Ellis's work is attentive to classroom realities. He acknowledges institutional constraints that make a 'strong' version of TBLT difficult to implement wholesale and therefore outlines 'weak' implementations that retain task primacy while accommodating syllabi and tests. He also recognizes individual differences, aptitude, working memory, motivation, and shows how adjustable parameters such as planning time or task complexity can personalize challenge without abandoning communal tasks. The result is a pedagogy that is both principled and feasible: meaning drives the lesson; form receives timely, minimal, and consequential attention; and assessment samples performance on families of tasks rather than isolated items.

For the present project, an adaptive, technology-mediated environment with a specific emphasis on vocabulary and receptive skills, Ellis's framework is especially serviceable. Tasks can be designed around authentic listening and reading, with adaptive scaffolds that modulate complexity and supply just-in-time lexical support. Micro-interventions, brief glosses, pop-up explanations, or targeted practice, instantiate focus on form without derailing comprehension. Analytics derived from task performance can tune planning time and feedback type, implementing the weak-interface logic at scale. In short, Ellis offers a blueprint for turning Krashen's input-driven insight into an instructional design that is concrete, measurable, and aligned with communicative ends.

1.8 Paul Nation and the Central Role of Vocabulary in SLA

Building on the foundational understanding of curriculum design and its impact on language learning, it is essential to recognize vocabulary acquisition as one of the most significant dimensions of second language acquisition (SLA). Vocabulary is not merely a list of words to be memorized; it is the gateway to comprehension, communication, and participation in authentic language use. The process of acquiring vocabulary is inherently multifaceted, involving both cognitive mechanisms, such as

memory, attention, and retrieval, and pedagogical considerations, including input frequency, task design, and deliberate learning strategies. As SLA theory and research make clear, lexical acquisition goes far beyond teaching isolated items. It entails equipping learners with the tools and strategies needed to engage meaningfully with texts, conversations, and contexts, thereby granting them fuller access to the communicative potential of the target language.

Paul Nation's scholarship places vocabulary at the center of second language acquisition (SLA), arguing that lexical knowledge, its breadth, depth, and accessibility under time pressure, conditions what learners can comprehend, say, read, and write. Over four decades, his research has provided principled architecture for teaching and testing vocabulary, clear coverage targets for reading and listening, and practical tools for designing courses and technologies that actually move the lexical needle. For a project emphasizing receptive skills and adaptive delivery, Nation's work is indispensable: it offers quantifiable goals, teachable methods, and assessment instruments that align learning activities with communicative outcomes.

Nation's starting point is empirical: how much vocabulary do learners need to function independently? Drawing on corpus-based frequency data, he shows that word frequency is not a pedagogical curiosity but a design constraint. The most frequent two-to-three thousand-word families typically account for the bulk of running words in general texts; beyond that, coverage rises more slowly and depends on topic and register. From this observation flow two influential benchmarks: learners generally require around 95% lexical coverage to follow a text with support (glosses, dictionaries, teacher scaffolding) and approximately 98% coverage to read or listen with relative independence. Translating coverage into size estimates, Nation argues that extensive reading typically demands 8–9k word families for comfortable comprehension of unsimplified prose, while unsimplified listening tends to be feasible at 6–7k families, given the nature of spoken discourse and repetition. These are not iron laws, the topic, genre, and learner background matter, but they anchor curricular targets in measurable realities rather than hopes.

Two conceptual distinctions organize Nation's account of what it means to 'know a word.' First, breadth versus depth: breadth concerns how many word families a learner recognizes, whereas depth captures the richness of that knowledge, collocations, polysemy, register, morphological relations, and associations. Second, receptive versus productive knowledge: learners recognize more than they can retrieve and use under communicative pressure. A curriculum that pursues only breadth through recognition will fail to produce usable vocabulary; one that drills production without sufficient input will founder on frequency realities. Nation's answer is an integrated pedagogy codified in his Four Strands model: (1) meaning-focused input (extensive reading and listening at high coverage), (2) meaning-focused output (tasks that require message-focused speaking and writing), (3) language-

focused learning (deliberate study of form–meaning mappings, word parts, and collocations), and (4) fluency development (activities with familiar material that push speed and automaticity). Each strand should occupy roughly a quarter of course time. The simplicity of the prescription belies its depth: it ensures that lexical growth is driven by frequency-rich input, converted into accessible knowledge through practice and deliberate study, and made usable by fluency work where no new items are introduced.

Nation's commitment to deliberate learning is pragmatic, not doctrinaire. Because many mid- and low-frequency items are too rare to be acquired quickly from incidental exposure, learners benefit from targeted techniques: spaced retrieval of word–meaning pairs; word-part (morphological) strategies that exploit prefixes, suffixes, and roots to enlarge families economically; collocation training that builds phraseological depth; and generative use, where words are recycled across new contexts to strengthen form–meaning connections. None of this replaces input; rather, it accelerates movement through frequency bands that ordinary exposure would traverse too slowly to meet learners' goals or program timelines.

Assessment, for Nation, must track what instruction claims to build. He developed and refined tools such as the Vocabulary Levels Test (VLT) to estimate knowledge across frequency bands, the Vocabulary Size Test (VST) to gauge overall coverage potential, and productive versions that measure what learners can reliably retrieve. Importantly, these instruments separate form-recognition from controlled productive ability, keeping researchers and teachers honest about the gap between what learners 'know' and what they can deploy in real time. The tests also make curricular decisions auditable: if extensive reading is not yet feasible because coverage is too low, the remedy is not wishful thinking but a pivot toward graded input and language-focused learning to raise the floor.

Several further strands of Nation's work matter for design. He has guided the construction and use of frequency-based word lists, both general and domain-specific, and advocated lexical profiling to diagnose the coverage and dispersion of items in materials. He underscores the pedagogical value of multiword items, collocations, phrasal verbs, formulaic sequences, arguing that fluency and idiomaticity depend on chunk knowledge, not just single lemmas. He also promotes narrow reading and listening within topics to maximize repeated encounters with mid-frequency items, and he consistently recommends extensive reading/listening with texts pitched at or slightly above the learner's current coverage so that incidental learning remains probable rather than accidental.

Nation's approach is neither anti-grammar nor anti-communication; it is lexically realist. Grammar matters as a system for composing meanings, but without rapid access to words and chunks, grammatical knowledge has little to operate on. Communication matters as the point of language learning, but without planned encounters with high-value items and principled recycling,

communicative practice will stall. What Nation offers, then, is a way to calibrate instruction to the statistical structure of language while keeping tasks communicatively meaningful and time use balanced across the four strands.

For the present project, an adaptive web application prioritizing receptive skills with knock-on effects for production, Nation's principles translate directly into design requirements. Materials should be profiled for frequency coverage so that most texts fall at $\geq 95\%$ for scaffolded activities and approach $\geq 98\%$ for independent work. The system should maintain a learner-specific lexical profile (breadth and depth), updating it from behavior (reading/listening logs, response accuracy, speed) and using it to adapt text selection and gloss density. A built-in spaced retrieval engine should schedule micro-reviews of high-utility items just before predicted forgetting, while morphology modules teach productive affixes and word-family relations to compound gains. Collocation prompts and formulaic sequence practice can be surfaced after authentic encounters, operationalizing focus on form without derailing meaning. Finally, a fluency mode, re-reading, shadowing, timed retellings with familiar texts, ensures that recognition becomes rapid and available for output when needed.

Critiques of a vocabulary-led approach typically worry about neglecting grammar or reducing communication to counting words. Nation anticipates these objections by insisting on balance: the four strands prevent lexical study from crowding out interaction and meaning, while frequency awareness guards against the equally common error of immersing learners in texts far beyond their lexical reach. In short, Nation does not narrow SLA to vocabulary; he centers vocabulary so that the rest of SLA can succeed.

1.8.1 Measuring Vocabulary Growth with Statistical Methods

Vocabulary acquisition is often considered a qualitative process, but research in second language acquisition (SLA) has shown that it can also be understood through a statistical lens. The relationship between vocabulary size and language proficiency is not merely theoretical but is supported by measurable data. Studies have demonstrated that factors such as the frequency of word occurrence and the percentage of known vocabulary in a text play a critical role in how well learners can comprehend and retain new words. This section explores the statistical approaches that highlight the importance of vocabulary size in language acquisition, with a focus on key benchmarks and thresholds that influence learners' reading comprehension and overall language development.

On the other hand, the process of language development is often equated with a chaos-based system, emphasizing its dynamic, unpredictable, and self-organizing nature (Larsen-Freeman, 1997).

Nation's findings have profound implications for language teaching, suggesting that educators should prioritize vocabulary development through methods such as extensive reading to expose learners to

new words in context, intensive vocabulary instruction focusing on high-frequency and academic words, and teaching effective vocabulary learning strategies. By enhancing vocabulary instruction, teachers can significantly improve students' reading comprehension and overall language proficiency. The extensive reading strategy had been already in the focus during the debate between Krashen and Cobb (Cobb, 2008; McQuillan & Krashen, 2007) on the efficacy of extensive reading as a primary method for second language acquisition. Krashen argues that extensive reading is crucial for language development. His stance is based on several key points: first, that extensive reading provides learners with copious amounts of comprehensible input. Second, Krashen believes that through exposure to a wide range of texts, learners naturally expand their vocabulary. Third, extensive reading fosters fluency by allowing learners to process language at a natural pace. Finally, enjoyable reading experiences can enhance learners' motivation to continue learning the language (Krashen, 1989).

Conversely, Cobb challenges the sufficiency of extensive reading as the sole method for language acquisition. He acknowledges the benefits of extensive reading but argues that it may not provide sufficient exposure to low-frequency words that are essential for academic language proficiency. Additionally, Cobb points out that extensive reading alone may not develop explicit knowledge of grammar and syntax, and that individual differences among learners mean that some may benefit more from explicit instruction (Cobb, 2007).

The core of the debate lies in the contrast between implicit versus explicit learning. Krashen emphasizes the importance of implicit acquisition through exposure to language, while Cobb advocates for a balanced approach that combines extensive reading with explicit instruction. Supporting Krashen's stance, research has shown that extensive reading can lead to significant vocabulary growth, particularly in terms of incidental learning (Nagy et al., 1985), improve reading fluency (Alderson, 1984), and enhance learners' motivation and autonomy (Elley, 1991). However, supporting Cobb's position, studies suggest that while extensive reading increases vocabulary breadth, it may not adequately develop depth of vocabulary knowledge (Schmitt, 1998), and explicit grammar instruction can be beneficial for learners with specific needs (Norris & Ortega, 2000).

The Krashen-Cobb debate highlights significant pedagogical implications for language instruction, particularly in the context of vocabulary acquisition. Krashen advocates for the power of extensive reading, arguing that it provides ample comprehensible input that facilitates incidental vocabulary learning and fluency development. Extensive reading helps learners acquire vocabulary in a natural, enjoyable way, thus enhancing their motivation and engagement (Krashen, 1985, 1989). However, Cobb challenges the sufficiency of extensive reading as the sole method for vocabulary development. He emphasizes the limitations of this approach, noting that it may not always provide adequate exposure to low-frequency words or develop explicit knowledge of grammar and syntax. Cobb argues

for a more balanced approach that integrates explicit vocabulary instruction alongside extensive reading (Cobb, 2007).

In light of this debate, it becomes evident that while extensive reading is crucial for vocabulary acquisition, it should be complemented by strategic learning activities. Research by Nation and others underscores the importance of effective vocabulary learning strategies. Techniques such as Spaced Repetition and Chunking, which involve reviewing vocabulary at increasing intervals and grouping related words, respectively, can further optimize vocabulary retention and understanding (Nation, 2001; Schmitt, 2000). Chunking involves grouping related words together to aid in recall and understanding. For example, learning vocabulary related to a specific topic or theme can strengthen memory connections. Chunking has been demonstrated to improve vocabulary acquisition by organizing information into manageable units (Baddeley, 1992). Research indicates that active recall helps consolidate memory and improve long-term retention of vocabulary (Nagy et al., 1985). Spaced Repetition entails reviewing vocabulary items at increasing intervals, optimizing learning efficiency. This technique is supported by evidence suggesting that spacing out review sessions enhances memory retention compared to massed practice (Cepeda et al., 2006). Mnemonics create memorable associations or images for new words, which can improve retention. Studies show that mnemonic devices, such as imagery and associations, are effective in enhancing vocabulary learning (Bellezza, 1981). Collocation Learning focuses on learning words in combination with other words, which helps learners acquire natural language patterns. Research highlights that learning collocations, or common word combinations, improves fluency and language comprehension (Lewis, 1997).

Context plays a vital role in vocabulary acquisition. When learners encounter new words in meaningful contexts, they are more likely to understand and retain them. Nation emphasized the importance of reading authentic texts, as they provide rich contextual information (Nation, 2006). Moreover, the type of context matters. Learners benefit more from encountering words in multiple contexts, as this reinforces meaning and flexibility of use. Additionally, the depth of processing involved in encountering a word in context is crucial. For example, inferring the meaning of a word based on surrounding information is more effective than simply looking up the definition. The distinction between explicit and implicit learning is a subject of ongoing debate (Hulstijn, 2005; Rieder, 2003), often overlapping with other concepts such as declarative/procedural knowledge and explicit/implicit memory (Hulstijn, 2005; Muñoz, 2012). This confusion arises from the fact that while incidental and intentional learning focus on motivation and methods, the explicit/implicit divide concerns the cognitive processing of language, with explicit learning involving conscious, effortful processing and implicit learning operating more automatically (Hulstijn, 2005, 2013).

Ellis (1994a) provides definitions for both terms, describing implicit learning as the natural acquisition of knowledge without conscious effort, whereas explicit learning involves a more conscious, hypothesis-testing approach. However, this distinction is not always clear, as learning is a dynamic process, and some experts (e.g., DeKeyser, 2003) argue that the explicit/implicit divide is better represented on a continuum, with learning outcomes shaping fluid ‘knowledge’ rather than fixed categories (Gass & Selinker, 2008).

Research on vocabulary acquisition suggests that perceptual and articulation aspects of new words are best learned implicitly, while meaning and mapping meaning onto form should be learned explicitly (Ellis, 1994). However, it is important to treat these conclusions cautiously, as the theoretical underpinnings of explicit/implicit learning remain unclear. Factors such as individual learner differences and the salience of language features input further complicate the feasibility of one mode over the other (Hulstijn, 2005).

Being vocabulary acquisition, and so its retention, a fundamental aspect of language learning, and tailoring vocabulary learning activities to different proficiency levels is crucial for effective instruction. Adapting activities to the learner’s stage ensures that instruction is both appropriate and effective, facilitating better comprehension and more efficient vocabulary retention and use. Moreover, the real challenge lies in integrating and embedding the wide range of techniques proven effective by scientific literature into technological tools to enhance learning effectively. By leveraging these evidence-based strategies within technological frameworks, educators can create more engaging and impactful learning experiences that cater to diverse learner needs and proficiency levels. Vocabulary acquisition and retention are fundamental to language learning, requiring instruction tailored to different proficiency levels. A major challenge is integrating evidence-based techniques into technological tools to optimize learning outcomes. By embedding scientifically validated strategies within digital frameworks, educators can create engaging and effective learning experiences that support diverse learner needs.

At the beginner level (CEFR A1–A2), learners focus on acquiring 500–2,000 high-frequency words essential for basic communication. Instruction emphasizes concrete nouns, core verbs, adjectives, and function words, using visual aids, flashcards, and simple sentence-building exercises. Gradual exposure to contextualized input through dialogues and basic reading materials strengthens foundational vocabulary and communicative competence.

At the intermediate level (CEFR B1–B2), learners expand their lexicon to 2,500–6,000 words, including abstract vocabulary, collocations, and idiomatic expressions. Instruction prioritizes semantic depth, word relationships, and flexible use through structured reading, discussions, and

writing tasks. Contextualized exposure to phrasal verbs and discipline-specific vocabulary enhances fluency and comprehension, supported by noticing strategies that analyze word use in authentic texts. At the advanced level (CEFR C1–C2), learners develop near-native proficiency with 8,000–15,000+ words, incorporating academic, technical, and specialized terminology. Instruction involves morphological analysis, lexical inferencing, and discourse comprehension, fostering nuanced expression through academic reading, analytical writing, and professional discussions. Techniques such as corpus-based learning, extensive reading, and exposure to complex texts refine lexical sophistication and precise language use.

Aligning vocabulary instruction with proficiency levels ensures systematic language development, preparing learners for advanced communication in academic, professional, and social settings.

Our attention, however, is primarily focused on the Lexical Approach, a method introduced by Michael Lewis in the 1990s (Lewis, 1993a). The lexical approach to second language acquisition underscores the centrality of vocabulary and fixed lexical phrases as foundational elements of language learning and communication. This approach posits that language is fundamentally composed of grammaticalized lexis rather than lexicalized grammar. This paradigm shift emphasizes the importance of 'chunks' of language, such as collocations, fixed and semi-fixed expressions, and idioms, over traditional grammar rules (Lewis, 1993b). The lexical approach argues that mastery of these chunks facilitates fluency more effectively than a primary focus on grammatical structures.

The Lexical Approach emphasizes several core principles that redefine language instruction. First, it prioritizes the acquisition of lexical chunks as holistic units of meaning. This perspective asserts that language learning should focus on these prefabricated units rather than dissecting language into individual words and grammatical rules (Lewis, 1993). By doing so, learners are better equipped to use language naturally and fluently, reflecting how native speakers employ these chunks in real-world communication.

Moreover, the Lexical Approach advocates for corpus-informed instruction. This method relies on authentic language data sourced from corpora to design teaching materials and activities. Such an approach ensures that learners engage with real language use, thereby enhancing the relevance and applicability of their learning experiences (Willis, 1990). By analyzing and practicing with language that mirrors actual usage, learners can more effectively acquire and utilize lexical chunks.

A key aspect of the Lexical Approach is its focus on fluency development. It posits that fluency is achieved more effectively through the acquisition and use of lexical chunks rather than through traditional grammar-focused instruction. Lewis (1993) argues that mastering these chunks allows learners to produce language more spontaneously and coherently, which is crucial for effective communication.

While grammar is not disregarded, the Lexical Approach assigns it a secondary role compared to lexical learning. This approach posits that grammatical understanding will naturally develop as learners encounter and use lexical chunks in context. It emphasizes that proficiency in a language is most effectively achieved through practical application rather than isolated rule-based exercises (Lewis, 1993).

Supporting this perspective, the study *Incidental Development of Implicit and Explicit Grammar Knowledge from Reading: The Case of Lower-Intermediate Learners of English* demonstrated that reading short texts with repeated exposure to target structures can foster the development of implicit grammar knowledge, as assessed by receptive instruments (Ward & Lo, 2023). The findings revealed that these benefits extend beyond younger learners. Although cognitive maturation and sensitive periods suggest a transition from implicit to explicit learning capabilities during puberty (DeKeyser, 2000), the study found that implicit knowledge development can still occur among senior secondary learners (Ward & Lo, 2023).

This evidence aligns with the Lexical Approach's emphasis on context-driven learning, suggesting that exposure to language in meaningful contexts, such as through extensive reading, can support the natural evolution of grammatical understanding. While the study does not advocate for replacing current grammar instruction entirely with reading, it provides a justification for placing greater emphasis on reading and extensive reading as a separate and valuable component of the language learning curriculum.

Additional considerations in applying the Lexical Approach include the role of context in language use. Lexical chunks are highly context-dependent, and learners must develop the ability to use these chunks appropriately across different situations. Contextual awareness is essential for applying lexical chunks effectively and for understanding their nuanced meanings (Willis, 1990).

The approach also promotes learner autonomy by encouraging students to seek out and learn lexical chunks independently. This empowerment supports self-directed learning and enhances learners' ability to engage with authentic language materials outside of the classroom setting (Lewis, 1993).

Furthermore, the integration of technology in language learning has expanded the possibilities for implementing the Lexical Approach. Digital tools and online corpora provide learners with access to authentic language data and interactive exercises, facilitating the acquisition and practice of lexical chunks in diverse contexts (Collins COBUILD, 1987).

In summary, the Lexical Approach offers a compelling framework for language teaching by shifting the focus from isolated grammatical rules to the practical use of lexical chunks. Through corpus-informed instruction, emphasis on fluency, and the promotion of learner autonomy, this approach

aligns language learning more closely with natural language use and develops the effectiveness of communication.

According to the lexical approach, a substantial portion of language use involves the production and comprehension of these pre-formed chunks. Lewis (1997) suggested that the mind stores these chunks for rapid retrieval and contextual use, making language processing more efficient than constructing phrases from scratch using grammatical rules. This process is enhanced by exposure to natural language contexts, which helps learners internalize chunks as functional units rather than isolated grammatical elements.

Empirical research supports the lexical approach by highlighting the frequency and utility of lexical chunks in natural language use. Nattinger and DeCarrico (1992) stress the importance of these phrases in fluent language production and advocate for their central role in language teaching. Their findings suggest that instruction focusing on lexical chunks leads to more natural and fluent language use, similar to that of native speakers.

In digital environments, the lexical approach informs teaching practices by emphasizing activities that involve recognizing, categorizing, and utilizing lexical chunks in realistic contexts. Data-driven learning, where learners analyze corpora to identify and practice common chunks, and task-based learning that incorporates these phrases into communicative tasks, are practical applications of this approach (Willis, 1990). However, critics such as Swan (2006) argue that while the lexical approach enhances language production, it may insufficiently address the underlying grammatical structures necessary for creative language use. Some also contend that a focus on chunks might result in formulaic language production without a deeper understanding of flexible grammatical usage.

Despite these critiques, the lexical approach has significantly impacted modern language teaching by emphasizing practical communication skills and real-life usage. It aligns with a broader trend toward communicative and task-based teaching methods, which are believed to be more effective for developing functional language proficiency.

1.8.2 Lexical Development in Second Language Acquisition

At the core of all communication, whether spoken or written, lies the use of words, which act as the essential building blocks of language. The process of acquiring and using vocabulary is fundamental to both fluency and comprehension, as it shapes how ideas are expressed and understood. This section will explore the various dimensions of word knowledge, form, meaning, and usage, and examine how these components contribute to effective language learning and communication. The understanding of a word and its meaning is a process that looks at the overall content but is always based on access

to the unit itself, highlighting the importance of grasping individual words to build deeper comprehension.

Understanding what it truly means to know a word is a foundational concept in the study of language acquisition, particularly in second language learning. While many learners may recognize and use words in context, the depth of their knowledge can vary significantly. Lexical knowledge is not merely about memorizing a word's meaning; it encompasses a range of cognitive, linguistic, and contextual factors. In this chapter, we will explore the multiple dimensions of word knowledge, including form, meaning, and use, and how these elements play a crucial role in second language acquisition. By examining how learners internalize and activate vocabulary, we can better understand the complexities of lexical acquisition and its impact on fluency and comprehension.

The process of discovering a word can be likened to a journey, where learners traverse through various stages of awareness, understanding, and mastery. This journey is not linear; rather, it is marked by multiple encounters with the word in diverse contexts, each encounter contributing to a deeper and more nuanced understanding of its form, meaning, and use. At the beginning of this journey, learners may only have a superficial recognition of the word, perhaps encountering it in a list or hearing it used in conversation. At this early stage, the word exists primarily in isolation, lacking significant connections to the learner's existing knowledge or linguistic repertoire.

The acquisition of vocabulary and the clear definition of the process of acquiring a word remain complex and challenging in second language acquisition. Richards (1976) identified several key features of word knowledge, such as frequency, register, form, associations, meaning-concept, and meaning-associations, which contribute to understanding how learners acquire and use words. Nation (1990, 2001) expanded on Richards' framework, specifying nine distinct types of word knowledge, and made an important distinction between receptive and productive knowledge for each feature. Nation's model also presents form, meaning, and use as the primary facets of word knowledge, each with three subcategories, such as spoken and written forms. This distinction underscores the multifactorial, incremental nature of vocabulary acquisition, highlighting that learning vocabulary is a dynamic, ongoing process rather than a discrete event. As Klapper (2006) points out, vocabulary knowledge is subject to constant change, influenced not only by the learner's evolving understanding but also by individual perception and external factors. The denotation and connotation of a word, shaped by context and experience, contribute to how learners internalize and apply vocabulary. Schmitt (1998) further argues that there is no single moment when one can definitively claim to fully 'know' a word, reinforcing the notion that vocabulary acquisition is progressive. In parallel, Paul Meara's (1996) three-dimensional model of vocabulary knowledge introduces size, organization, and accessibility as key dimensions. The size dimension refers to the number of words a learner knows,

which correlates with overall language proficiency. The organization dimension highlights how words are interrelated in the learner's mental lexicon, including relationships such as synonyms, antonyms, and collocations. Finally, the accessibility dimension concerns the speed and ease with which learners can retrieve words, which influences both comprehension and production. Meara's model, together with Nation's framework, provides a comprehensive view of how vocabulary acquisition unfolds, shaped by both internal cognitive processes and external influences, as learners' lexicons evolve over time.

As the journey unfolds, the process of noticing plays a crucial role. Sharwood Smith (Smith, 1991, 1993) emphasized that for a word to become part of a learner's active vocabulary, it must first be noticed, both in the input and in the learner's own linguistic output. This noticing is akin to the first steps on the path, where the learner acknowledges the presence of the word and begins to recognize its potential relevance. Through repeated exposure and conscious attention, the word begins to move beyond mere recognition toward a more active role in the learner's lexicon.

With continued exposure, the learner enters a phase of increased awareness, which reflects a more sophisticated stage in the journey. Awareness involves not only recognizing the word but also reflecting on its multiple meanings, its syntactic properties, and its possible collocational patterns. This stage may involve learners' engagement with the word in different contexts, deepening their understanding of its usage and extending its utility in both comprehension and production. The word thus evolves from a fleeting encounter to a more integrated part of the learner's language system, ready for activation when needed.

By the final stages of this journey, the word has become fully internalized, embedded in the learner's vocabulary, and capable of being used with greater fluency and confidence. At this point, the learner's cognitive understanding of the word aligns with its communicative potential. The journey of discovering a word, therefore, is not merely about memorizing its definition; it is about navigating through a dynamic, ongoing process of awareness, reflection, and integration, ultimately enhancing both comprehension and expressive capacity in the target language.

Complete mastery of a word demands understanding its syntactic, semantic, pragmatic, phonological, orthographic, and morphological dimensions, a level of comprehension rarely attained, even by native speakers. This underscores the inherent *fuzziness* of word knowledge (Thornbury, 2002), a complexity that learners, researchers, and teachers must acknowledge.

1.8.3 Supported Approaches To Words Internalization

Vocabulary acquisition stands as a complex process requiring strategies that promote deep retention and internalization. Traditional methods, while foundational, often fail to cater to individual learning

needs and optimize long-term retention. The advent of computer-assisted language learning (CALL) has introduced innovative approaches that personalize instruction, offer immediate feedback, and leverage adaptive technologies to enhance vocabulary acquisition. CALL applications, such as spaced repetition software (SRS), have been shown to significantly improve retention rates by optimizing review intervals based on cognitive science principles (Nakata, 2011).

Beyond exposure and repetition, assessment plays a crucial role in interiorizing vocabulary. Effective vocabulary assessment extends beyond word recognition to include contextual understanding, word collocations, and nuanced meanings. Dynamic assessment (DA), grounded in Vygotsky's sociocultural theory, offers a formative and interactive approach where learners receive feedback during assessment, allowing for real-time lexical development. This method not only evaluates existing knowledge but also scaffolds further learning, reinforcing deeper lexical acquisition.

A comprehensive vocabulary acquisition strategy must integrate both incidental and intentional learning, balance receptive and productive skills, and foster strong form-meaning associations. The advent of technology has significantly transformed language learning, including the acquisition of vocabulary. Digital tools and resources provide diverse and innovative ways to enhance vocabulary development, making the learning process more engaging, efficient, and tailored to individual needs. The integration of technology into vocabulary instruction has expanded opportunities for interactive engagement, personalized learning, and exposure to authentic language use, ultimately improving both short-term retention and long-term acquisition.

One of the most notable contributions of technology to vocabulary learning is interactive learning. Digital platforms offer a variety of exercises, games, and quizzes that encourage active participation. Unlike traditional rote memorization, these activities immerse learners in dynamic learning experiences that facilitate deeper cognitive processing. Many interactive tools incorporate elements such as progress tracking, adaptive difficulty levels, and contextual learning, which enhance motivation and sustain learners' interest over time. Moreover, digital environments can simulate real-life contexts where vocabulary is naturally used, reinforcing both comprehension and practical application. These tools not only make vocabulary acquisition enjoyable but also promote learner autonomy, allowing individuals to engage with language in ways that best suit their learning preferences.

Immediate feedback is another crucial advantage of technology in vocabulary instruction. Traditional learning methods often require learners to wait for feedback from teachers or peers, which can slow down progress and make it harder to correct mistakes in real time. Digital learning tools, however, provide instant feedback, allowing learners to assess their understanding and make corrections immediately. This immediate response mechanism refines learning efficiency, reinforces correct

usage, and prevents the reinforcement of errors. Additionally, it fosters a sense of self-efficacy as learners gain confidence in their ability to recognize and address their mistakes. Research suggests that prompt and constructive feedback plays a key role in identifying knowledge gaps and accelerating vocabulary acquisition.

Personalized learning is another transformative aspect of technology in vocabulary development. Adaptive learning technologies analyze learners' performance, preferences, and progress, tailoring content and difficulty levels to match their individual needs. This customization ensures that learners receive appropriate challenges without becoming overwhelmed or disengaged. Many vocabulary-learning platforms integrate intelligent systems that adjust to a learner's strengths and weaknesses, ensuring targeted reinforcement of difficult words while minimizing unnecessary repetition of already mastered vocabulary. Personalized learning environments contribute to higher engagement and improved learning outcomes, as learners feel more connected to the material and can progress at their own pace.

Technology also grants learners access to a wealth of authentic materials, including digital books, articles, podcasts, videos, and interactive media. Exposure to authentic language use in real-world contexts helps learners acquire vocabulary more naturally and meaningfully. Unlike textbook-based learning, which often presents vocabulary in isolated lists, digital resources provide words and phrases within meaningful contexts, aiding comprehension and retention. Authentic materials also expose learners to diverse linguistic registers, regional variations, and colloquial expressions, helping them develop a more nuanced understanding of language. Studies have shown that encountering vocabulary in multiple contexts strengthens retention and promotes deeper learning. Finally, vocabulary management tools such as digital flashcards, vocabulary notebooks, and spaced repetition systems further enhance vocabulary retention. Spaced repetition algorithms optimize memorization by scheduling review sessions at strategic intervals. This method is based on cognitive research indicating that repeated exposure to vocabulary at increasing time intervals strengthens long-term retention. Digital vocabulary notebooks allow learners to create personalized word lists, add multimedia annotations, and organize words according to themes or difficulty levels. These tools provide structured and systematic approaches to vocabulary learning, making it easier for learners to track their progress and reinforce their knowledge over time (Godwin-Jones, 2018; Lin & Lin, 2019). In conclusion, the integration of technology into vocabulary learning has revolutionized the way learners engage with new words and phrases. Through interactive learning, immediate feedback, personalized instruction, authentic exposure, and structured vocabulary management, digital tools offer an effective and engaging means of vocabulary acquisition (Stockwell, 2010; Nation, 2013; Webb & Nation, 2017; Reinders & White, 2016). As technology continues to evolve, its role in

language learning will likely expand, offering even more innovative approaches to mastering vocabulary in the digital age.

1.8.4 Theoretical Underpinnings Of Text Enhancement Effectiveness

Effective vocabulary acquisition strategies are crucial for language learners as they progress beyond the initial stages. Text enhancements, techniques that increase the salience of target vocabulary within written texts (e.g., *highlighting*, *glossing*, *annotation*), play a dynamic role in this development (Laufer & Hulstijn, 2001; Watanabe, 1997; Han, Park, & Combs, 2008). While these methods initially support basic word recognition and contextual understanding, they become increasingly valuable in fostering deeper learning. Specifically, text enhancements facilitate the grasp of nuanced semantic relationships, the integration of new vocabulary with existing lexical knowledge, and the construction of a richer mental lexicon, ultimately moving learners beyond rote memorization toward genuine vocabulary mastery (Schmidt, 1990; Lee & Huang, 2008).

The role of text enhancements in vocabulary acquisition has become increasingly prominent in educational research, especially in the context of second language (L2) learning. Text enhancements refer to the array of strategies used to make target vocabulary more noticeable within written texts, such as highlighting, annotating, or visually modifying specific words or phrases. These techniques are designed to support the initial syntagmatic learning phase of non-native speakers, where vocabulary is often acquired through exposure to word sequences and contextual cues. However, as learners progress, text enhancements can play a crucial role in facilitating the transition to a more paradigmatic organization of the mental lexicon. By drawing attention to semantic relationships, synonyms, and connections to previously learned words, text enhancements help learners integrate new vocabulary into their existing cognitive framework.

Moreover, the incorporation of personalized and adaptive text enhancements has shown promise in tailoring the learning experience to individual needs. These adaptive strategies adjust the presentation of text based on a learner's current vocabulary knowledge, learning pace, and cognitive style. By dynamically modifying the salience of new words in relation to a learner's existing lexicon, personalized enhancements ensure that each learner receives targeted support that optimizes their acquisition process. This chapter delves into the effectiveness of these techniques, examining how they not only aid in the immediate recognition and retention of new words but also support the development of deeper, more interconnected lexical knowledge, ultimately enhancing long-term language proficiency.

Drawing on a range of empirical studies and theoretical frameworks, we will examine how different forms of text enhancement, such as glosses, annotations, and textual modifications, impact

vocabulary acquisition. We will also consider the cognitive mechanisms underpinning these enhancements and their effectiveness in diverse learning contexts. Through this exploration, the chapter seeks to elucidate how strategic text modifications can support vocabulary development and contribute to more effective language instruction.

Many researchers support Schmidt's (2001) view that the allocation of attentional resources is essential for learning. However, learners' own mechanisms for drawing attention may not always be sufficient. For example, L2 readers might skip over words they don't consider important for understanding the text (Hulstijn, 1993), making word learning through reading somewhat unpredictable. As a result, lexical enhancements and intervention tasks that direct the reader's attention toward specific words can substantially increase the likelihood of vocabulary acquisition. According to Laufer and Hulstijn (2001), motivational factors such as the learner's perceived need, along with cognitive factors including search and evaluation, are central to the process of word learning. Rott (2007) further explains that the components of the Involvement Load Hypothesis predict that a higher degree of involvement leads to greater incidental vocabulary learning. Learners reach the highest level of involvement when they experience an intrinsic motivation to learn rather than responding to external demands, when they actively seek out meanings instead of passively receiving them, and when they decide how a new word can be integrated with other words in an original sentence instead of merely identifying distinctions among words in controlled exercises like gap-fill tasks.

Text enhancements can be implemented using various techniques, one of which is text glossing. However, Roth's (Rott & Williams, 2003) arguments, as discussed by scholars such as Hulstijn (Hulstijn, 1992), Laufer (Laufer, 1992), challenge the effectiveness of glossing. Specifically, it has been argued that meanings inferred from context are more likely to be retained than those provided directly by glosses. This perspective is supported by the 'mental effort' hypothesis, which posits that (1) inferring meanings requires cognitive effort, and (2) increased mental effort enhances a learner's ability to recall and retain information acquired through such effort (Hulstijn, 1992, 2001).

In Rott's (2007) experimental study, glosses, defined as explanatory notes or word definitions, were found to support text comprehension. However, the research did not clarify how the number or distribution of glosses influences idea recall or individual readers' comprehension. Future studies should examine different frequencies and placements of glosses to determine the optimal ratio of glossed to non-glossed words that maximizes both word learning and comprehension. The study also left unexplained why word retrieval tasks did not enhance retention as expected, or whether varying the number of retrieval prompts might produce differential effects on learning outcomes. Moreover, Rott's investigation focused on a narrow range of word-knowledge dimensions, leaving broader

aspects, such as collocational or pragmatic knowledge, unexplored. Finally, although the study sought to measure incidental vocabulary acquisition, it did not account for the possibility that participants might have shifted to intentional learning in subsequent readings, potentially confounding the results. Arani and Yazdanimoghaddam (Arani & Yazdanimoghaddam, 2016) investigated the effects of input flooding and textual enhancement on the syntactic development of Iranian EFL (English as a Foreign Language) learners. Their results indicated that both techniques positively influenced learners' ability to recognize and produce syntactic structures. When used together, these methods led to higher scores in both production and recognition tasks. The study's findings offer valuable insights for language teachers, suggesting that incorporating both input flooding and textual enhancement can effectively support syntactic development in language learners.

Table 1 below provides a summary of various studies that have investigated the effects of textual enhancement on vocabulary and collocation learning in different contexts. The research spans a range of participant groups, from elementary EFL learners to university students, and explores different aspects of vocabulary acquisition, including single words and collocations. The methodologies employed in these studies vary, with some comparing textual enhancement to other instructional techniques, while others focus solely on the impact of textual enhancement itself. The findings offer valuable insights into the effectiveness of textual enhancement in different learning scenarios.

Study	Participants	Focus	Methodology	Findings
Wang & Wang (2009)	Not specified	Vocabulary acquisition	Compared output task, glossing, and textual enhancement	Output task significantly aids vocabulary learning; textual enhancement and glossing have no general facilitative effect.
Nahavandi & Mukundan (2014)	Iranian elementary EFL learners	Vocabulary intake	Experimental group with highlighted words vs. control group with unenhanced text	Experimental group outperformed control group in vocabulary intake.
Ghafouri & Masoomi (2016)	Iranian university students	Vocabulary acquisition	Textual enhancement	Textual enhancement promotes vocabulary acquisition.
Zhou & Lu (2010)	Not specified	Learning of English collocation	Textual enhancement	Significant effect on acquisition of collocation form, but not on meaning learning.

Szudarski & Carter (2016)	Not specified	Acquisition of verb-noun and adjective-noun collocations	Input flood only vs. input flood plus textual enhancement	Input flood plus textual enhancement improves form recall and form recognition.
Fazlali & Shahini (2019)	Iranian EFL learners	Grammatical and lexical collocation	Textual enhancement	No effect on grammatical collocation knowledge, significant effect on lexical collocation knowledge.
Fang & Jiang (2020)	Intermediate learners	Acquisition of collocations	Textual enhancement	Textual enhancement improves collocation acquisition; effectiveness does not vary with semantic transparency.

Table 1 Overview of studies that have investigated the effects of textual enhancement on vocabulary and collocation learning in different contexts

In their study, Christian Jones and Daniel Waller (Jones & Waller, 2017) examined the effects of Textual Enhancement (TE) and Auditory Enhancement (AE) on vocabulary learning. Their results indicated that both TE and AE produced more durable vocabulary learning outcomes compared to explicit teaching alone. The experimental group, which received enhanced input, showed significant short-term gains in productive knowledge, and overall, the treatment group outperformed the control group in both receptive and productive knowledge.

The authors emphasized the potential benefits of enhanced input, particularly for beginner learners, as it helps to signal which language elements are more useful or noteworthy. Koprowski (2005) supports this idea, noting that language learning materials often present various language elements without indicating their relative importance. By focusing on high-frequency and highly useful vocabulary, enhanced input can help learners prioritize their attention and guide teachers in emphasizing key language elements.

Additionally, the study suggests that TE (or TIE) and AE can be valuable tools not only in the classroom but also for independent study. Learners can use these strategies informally, such as highlighting useful vocabulary after reading, or more formally, in conjunction with tools like Coxhead's Academic Wordlist or the *Compleat Lexical Tutor*. Auditory methods, like recording and repeating key vocabulary, can also reinforce learning.

Cheng and He (2023) conducted a comprehensive review of empirical research on input enhancement and foreign language vocabulary acquisition, covering studies published from 2009 to 2023. Their

findings are consistent with the broader trend observed in earlier studies, affirming that textual enhancement generally has a positive impact on vocabulary acquisition.

The review highlights the variety of approaches to input enhancement that have been explored over the years, including techniques such as highlighting, glossing, and the use of contextual clues. Notably, much of the research has concentrated on reading materials, with comparatively less attention given to other modes of language input, such as listening.

Key takeaways from Cheng and He's review include the significant increase in the number of studies on input enhancement in recent years, reflecting a growing academic interest in this area. Additionally, there is an emerging focus on exploring the interactive and comparative effects of different input enhancement techniques. Cheng and He suggest that future research should expand to investigate the effectiveness of input enhancement across various modes of language learning, such as listening and speaking, and in different contexts, including online learning environments and traditional classroom settings.

As highlighted by Sauer (2017) the explicit/implicit learning debate ties into the broader theory of input enhancement, which looks at how different forms of processing lead to more effective learning. The concept of 'depth of processing,' developed in the 1970s within cognitive psychology, plays a key role in understanding the mechanisms behind input enhancement and the impact of explicit versus implicit learning modes.

While there is little empirical evidence regarding the superiority of one learning mode over the other (DeKeyser, 2003), these debates contribute to ongoing research on how best to approach language acquisition. Some research suggests that language features with high frequency and salience may lend themselves more easily to implicit learning, while others may require more conscious processing.

1.8.4.1 How Text Enhancements Influence Learning

Several enhancement techniques that can make the process more effective. Three key methods, textual enhancement (TE), aural enhancement (AE), and input flooding, offer distinct benefits and limitations in vocabulary acquisition. Understanding their impact can help optimize language learning strategies.

1.8.4.2 Textual Enhancement (TE)

Textual enhancement involves using visual cues, such as bolding, underlining, italicization, and color-coding, to highlight specific linguistic features. This method increases the visibility of target

vocabulary and structures, aiding in retention. For example, bolding a new word like serendipity or using a different color for all verbs in a passage can enhance learner focus. Research has shown that TE is particularly effective when combined with glossing, where definitions or explanations accompany highlighted terms (Han, Park, & Combs, 2008; Lee & Huang, 2008; Watanabe, 1997). Additionally, TE draws attention to word forms and grammatical structures, promoting a deeper understanding of language mechanics (Schmidt, 1990; Sharwood Smith, 1991). However, the effectiveness of TE can vary depending on factors such as learner engagement and the type of vocabulary being targeted. Not all studies show significant improvements, and excessive highlighting can be counterproductive, overwhelming learners and distracting from overall comprehension (Izumi, 2002).

1.8.4.3 Aural Enhancement (AE)

Aural enhancement emphasizes auditory input to direct learners' attention to specific linguistic features. Techniques include volume adjustments, strategic repetition, and emphasis on key words or phrases. For instance, repeating a difficult word like squirrel multiple times in an audio exercise can help learners distinguish subtle sounds and improve pronunciation. Although promising, research on AE remains limited compared to TE. Studies have explored how auditory salience can aid attention and pronunciation but show mixed outcomes (Leow, 2009; Wong, 2001; Winke, Gass, & Sydorenko, 2010). While AE may enhance auditory attention and pronunciation skills, its overall effectiveness in vocabulary acquisition is not yet well established, suggesting the need for further empirical investigation.

1.8.4.4 Input Flooding

Input flooding exposes learners repeatedly to a specific linguistic feature, such as a grammatical structure or vocabulary item, to reinforce recognition and retention. For example, a reading passage that repeatedly uses the past perfect tense can help learners internalize its structure. This method can increase noticing and facilitate vocabulary acquisition (Trahey & White, 1993; Reinders & Ellis, 2009; Lee & Huang, 2008), yet its effectiveness varies. Some studies indicate that repeated exposure improves retention, while others suggest it does not significantly outperform non-enhanced input (Hernandez, 2008). Additionally, excessive repetition can lead to cognitive overload, making it harder for learners to process and retain information.

1.8.4.5 Cognitive Mechanisms Underlying Text Enhancement Effects

The effectiveness of text enhancement depends not only on the design of instructional materials but also on the cognitive mechanisms that mediate how learners process input. Central among these mechanisms are attention, noticing, and working memory capacity, each playing a pivotal role in determining whether an enhanced feature becomes integrated into long-term linguistic knowledge. According to Schmidt's Noticing Hypothesis (1990, 2001), learners must consciously register a linguistic form for it to be acquired. Textual enhancement (TE), by making features perceptually salient through typographic cues such as bolding or underlining, increases the likelihood that learners will allocate attentional resources to the target structure (Sharwood Smith, 1991; Leow, 2015).

However, noticing alone does not guarantee acquisition. Learners must also engage in depth of processing, linking the highlighted form to meaning and context in ways that strengthen encoding (Craik & Lockhart, 1972). Working memory further conditions the outcome of TE: learners with higher capacity can hold enhanced forms in active focus while processing meaning, reducing cognitive overload (Baddeley, 2003; Mackey et al., 2002). In contrast, learners with lower working memory may struggle to integrate form and meaning when too many features are highlighted, leading to superficial processing.

Theories of input processing (VanPatten, 1996, 2015) emphasize that learners naturally prioritize meaning over form; thus, enhancement is most effective when it supports, rather than competes with, comprehension. Finally, frequency effects interact with attention: repeated exposure to enhanced forms facilitates automatization, but excessive redundancy can cause habituation, reducing salience over time (Ellis, 2002). In sum, TE is most effective when it aligns with cognitive principles, respecting limited attentional resources, fostering meaningful connections, and leveraging working memory to mediate the path from input to intake.

The impact of text enhancement varies across learning contexts, shaped by modality, proficiency level, instructional goals, and cultural expectations about language learning. In formal classroom settings, TE can be systematically integrated into reading passages, vocabulary lists, or grammar exercises, often complemented by teacher-guided tasks that reinforce noticing. Research suggests that enhancement is particularly beneficial for intermediate learners, who possess enough linguistic knowledge to recognize contrasts but still require support to notice less salient forms (Han, Park, & Combs, 2008; Lee & Huang, 2008). Beginners may experience cognitive overload when multiple forms are enhanced simultaneously, whereas advanced learners may perceive enhancement as redundant unless it targets subtle pragmatic or discourse-level features.

In computer-assisted and mobile learning environments, TE can be dynamically adapted: learners may toggle glosses, highlight difficult words, or receive adaptive prompts based on performance

(Godwin-Jones, 2018; Lin & Lin, 2019). The strength of TE in such contexts lies in its customizability, allowing individualization according to proficiency and cognitive style. In informal learning, for example, during extensive reading, pleasure-heavy-handed enhancement can disrupt flow and reduce motivation. Lighter forms of support, such as pop-up glosses or gradually fading color coding, may better balance notice with fluency (Nation, 2001; Webb & Nation, 2017).

Cross-cultural factors also influence learner response: students from traditions emphasizing explicit grammar instruction often respond positively to overt enhancement, while those accustomed to communicative approaches may prefer subtler cues (Ellis, 2008; Doughty & Williams, 1998). Overall, textual enhancement is not a universal solution but a flexible pedagogical tool. Optimal results are achieved when enhancement aligns with learner profile, proficiency, and modality, and when it is integrated with complementary techniques such as glossing, feedback, and interaction.

1.9 Cognitive Dimensions Of Second Language Learning

While instructional techniques such as textual enhancement, aural enhancement, and input flooding provide external means of directing attention to language features, their ultimate effectiveness depends on the internal cognitive processes of learners. In other words, enhancement makes linguistic input more salient, but it is the learner's perceptual, memory, and attentional systems that determine whether noticing occurs and whether exposure results in durable acquisition. For this reason, exploring the cognitive mechanisms underlying text enhancement and the broader cognitive factors in SLA is essential. These perspectives reveal why the same instructional technique may lead to significant gains for one learner but minimal effects for another, and they highlight the importance of aligning pedagogical strategies with individual cognitive profiles.

Humans exhibit a high sensitivity to cognitive styles, which influence how we perceive, process, and interact with the world around us. These cognitive processes shape our ability to learn, solve problems, and communicate, making them central to education, psychology, and everyday decision-making. Understanding the nuances of cognition and cognitive styles provides valuable insight into how individuals approach intellectual tasks, interpret information, and adapt to different environments.

Cognition is a broad and multifaceted concept encompassing the mental processes involved in acquiring, processing, and utilizing knowledge. According to Te Molder and Potter (2005), cognition involves the accumulation and interpretation of information gained through learning or experience, which is then transformed into knowledge. This process enables individuals to understand and navigate the world around them. In essence, cognition functions as the intermediary between external

stimuli and internal understanding, allowing humans to interpret their environment, solve problems, and make informed decisions. The study of cognition spans various domains, including perception, memory, language, and reasoning, each contributing to a comprehensive understanding of how individuals process information and apply it to their daily lives. Rather than being a passive reception of information, cognition requires active engagement, where individuals interact with their surroundings, interpret signals, and construct meaning. This engagement is crucial for the development of higher-order thinking skills, such as critical analysis and abstract reasoning, which support complex problem-solving and decision-making.

Cognitive style refers to the preferred manner in which individuals perceive, think, and remember information. It represents the distinctive ways in which people approach intellectual and perceptual activities, serving as a link between personality and cognition (Witkin et al., 1977; Keefe, 1979; Riding & Rayner, 2013). Cognitive style is not merely a reflection of intellectual ability but rather an expression of how individuals engage with and process information. It encompasses a range of cognitive processes, including how people organize thoughts, solve problems, and make decisions. Cognitive style is highly individualized, shaped by a combination of genetic predispositions, cultural influences, educational experiences, and personal preferences. Some individuals may exhibit an analytical cognitive style, characterized by a preference for logical reasoning and structured problem-solving, while others may have a more holistic style, favoring intuition and a broad perspective. Recognizing these differences in cognitive style is essential for educators and psychologists, as they can influence teaching methods, learning outcomes, and the overall educational experience.

The way in which learners approach problem-solving is influenced by their cognitive style. Different cognitive styles lead to varied strategies for tackling challenges. Some learners, particularly those with a field-independent (FI) cognitive style, prefer to deconstruct tasks into smaller, manageable components. This analytical approach involves breaking down a problem into its constituent parts, analyzing each element individually, and then synthesizing a solution. This method is particularly effective in structured, well-defined problems where a step-by-step approach is advantageous. In contrast, learners with a field-dependent (FD) cognitive style may prefer a holistic approach, perceiving problems in their entirety and focusing on the relationships and patterns that emerge from the broader context. These learners might rely on intuition and global thinking, making connections between seemingly disparate pieces of information to arrive at a solution. This approach is often more effective in complex, ill-defined problems where understanding the overarching theme is key to finding a resolution. The variability in problem-solving approaches underscores the importance of recognizing individual cognitive styles in educational settings and highlights the need for diverse

teaching strategies that accommodate both analytical and holistic approaches to enhance learning outcomes.

Cognitive style is distinct from intelligence, as the two constructs, although related, fulfill different functions. Cognitive style refers to the characteristic way in which an individual processes information, while intelligence generally denotes the capacity for learning, reasoning, problem-solving, and adaptation to novel situations. Empirical research indicates that cognitive styles are not directly correlated with intelligence levels; a higher IQ does not necessarily correspond to superior cognitive performance across all contexts. An individual with a high IQ may demonstrate strong analytical and logical reasoning skills but may not perform equally well in tasks that require holistic thinking, creativity, or divergent problem-solving. Conversely, a person with a lower IQ but a strong holistic thinking ability might perform exceptionally well in situations that require pattern recognition or intuitive problem-solving. This distinction highlights the importance of understanding that diverse cognitive approaches, rather than intelligence alone, can contribute to success in various domains.

Cognitive styles are not fixed traits but can vary among individuals and even within the same individual depending on the context. This variability suggests that cognitive style is a dynamic characteristic that changes in response to different tasks, environments, or educational experiences. As noted by Witkin et al. (1977) and Messick (1984), cognitive styles offer a measure of differences among learners, providing valuable insights into how they interact with and interpret their surroundings. This has implications for education and psychology, suggesting that learners may benefit from flexible instructional strategies tailored to their cognitive style. A learner who typically adopts an analytical approach may require a more holistic strategy when dealing with complex, unstructured problems. Similarly, educators who understand the cognitive styles of their students can adapt their teaching methods to support both field-dependent (FD) and field-independent (FI) learners, fostering a more inclusive and effective learning environment.

Cognitive styles can be broadly categorized based on how individuals perceive and interact with their surrounding context, particularly in terms of field dependence and independence. Field-dependent learners tend to perceive the whole field or context rather than focusing on individual components. They often exhibit a more social orientation, relying heavily on external cues and social aspects of communication and interaction. FD learners may prefer physical proximity to others and utilize social referents to guide their understanding and interpretation of information. In language learning, this reliance can manifest as an emphasis on contextual clues and social interactions to derive meaning from text or speech. In contrast, field-independent learners are more analytical and focus on specific components of a task, often analyzing items separately from their surrounding context. They typically set their own learning goals and prefer using internal cues rather than depending on external or social

information. FI learners are often more self-reliant and capable of decontextualizing information, which can be advantageous in tasks requiring attention to detail or pattern identification within complex contexts.

The susceptibility of learners to various types of cues differentiates FD and FI learners. Field-independent learners are generally less influenced by external cues and social aspects of the environment, allowing them to restructure the field independently. This cognitive style enables them to detach information from its original context and reanalyze it based on internal logic and patterns. Consequently, FI learners are particularly adept at problem-solving, critical thinking, and identifying hidden patterns or inconsistencies. They often operate autonomously, relying on their internal frameworks and cognitive strategies rather than external guidance or peer support. Conversely, field-dependent learners are more conformist to the existing organization of the field and are less likely to restructure information. They tend to accept material as it is presented, relying on external cues, which makes them more influenced by the social and environmental context in which learning occurs. This characteristic often makes FD learners more receptive to collaborative learning environments, where social interaction and external feedback play a significant role in cognitive processing.

In the context of vocabulary learning, understanding distinctions between cognitive styles and their implications for task performance can enhance instructional strategies. Field-dependent learners tend to approach tasks holistically, relying on contextual and social cues. They are less likely to deconstruct the information and often digest material within its given structure. This makes FD learners particularly adept at tasks requiring integration of information within a broader context or involving social interaction. However, their performance might suffer in tasks requiring detailed analysis or restructuring of information. Field-independent learners excel in identifying patterns and details within complex contexts. Their analytical approach allows them to separate components from their context, enabling them to detect relevant information and make connections that might be overlooked by FD learners. This cognitive flexibility makes FI learners effective in problem-solving, critical thinking, and identifying hidden patterns. However, their focus on analytical detail might hinder their ability to perceive the broader context or collaborate effectively in highly social environments. Understanding these differences in task performance highlights the importance of differentiated instruction and instructional design tailored to diverse cognitive styles.

1.9.1 Integrating Behavioral Studies Into The System

Although phenomena such as Dual Process Theory (DPT) and neuroscientific principles are often explored in cognitive psychology and decision-making studies, they are not always directly linked to

language acquisition. This chapter seeks to bridge this gap by integrating these cognitive concepts to enrich our understanding of language learning processes.

Dual Process Theory, as articulated by Kahneman in *Thinking, Fast and Slow* (2013), differentiates between System 1 (automatic, subconscious processing) and System 2 (effortful, analytical processing). Research in second language acquisition (SLA) suggests that implicit learning, often associated with System 1, plays a crucial role in acquiring grammar and phonology without conscious effort (Ellis, 2005). Conversely, explicit learning, linked to System 2, supports metalinguistic awareness and deliberate rule application (DeKeyser, 2000). By applying this framework to language acquisition, we aim to examine how these distinct processing schemas influence both implicit and explicit learning mechanisms, reinforcing findings that learners benefit from a balance of both approaches (Ullman, 2016).

The concept of personal semantics (Coronel & Federmeier, 2016), self-knowledge about personal preferences, emerges as distinct from general semantics, which pertains to knowledge shared across individuals. Personal semantics has been shown to modulate the N400 component of the ERP, a key marker of semantic processing, in a manner similar to general semantic knowledge. Interestingly, the N400 responses elicited by personal preferences mirror those observed for general knowledge in terms of timing and scalp distributions. This finding suggests that, despite the personal relevance of these preferences, they engage similar neural mechanisms to those involved in processing general semantic information. This challenges the notion that personal and general semantic knowledge are processed entirely separately or rely on distinct cognitive systems. For instance, the processing of self-relevant information, such as preferences, seems to tap into the same neural networks as general semantic processing, rather than forming a separate, isolated domain of knowledge.

However, some differences do exist in the processing of personal versus general semantics. A notable distinction is the somewhat smaller congruency effects observed in personal preference sentences, which suggests that the strength of expectancy for personal information might differ from that of general semantic information. Unlike general preferences, which are often assessed using cloze probability measures, personal preferences are typically evaluated based on individual self-assessments. This may lead to weaker expectancies for the congruent items in personal preference sentences, highlighting a potential divergence in how expectancy operates within personal semantics compared to general semantics. The subtle differences in N400 amplitude could point to varying degrees of predictability or salience between personal and general knowledge, with personal preferences potentially invoking more individual, context-dependent processing.

In addition to these N400 modulations, differences between personal and general semantics also emerge later in the processing stream, particularly in the form of larger late positive complex (LPC)

responses to personal preference sentences. This suggests that personal semantics might also invoke distinct, episodic memory processes, particularly when preferences are linked to autobiographical experiences. The LPC, often associated with the retrieval of contextual details, may indicate that personal preferences involve both semantic and episodic memory components. This is consistent with findings from studies on self-relevant information, where larger LPC responses have been observed for information tied to personal relevance or autobiographical memories.

Ultimately, these findings suggest that personal semantics, while clearly self-relevant, shares important functional and neurobiological features with general semantic knowledge. The overlap in processing mechanisms for personal and general semantics underscores the idea that self-knowledge does not operate in isolation but rather draws on the same cognitive and neural resources that support general semantic processing. Despite these similarities, the subtle differences in processing speed, expectancy, and the role of episodic memory highlight the unique nature of personal semantics, which integrates both semantic and autobiographical elements. This contrast with general semantic knowledge opens up new avenues for exploring how different types of knowledge are organized and processed in the brain, and how they contribute to our understanding of self-related information.

Additionally, neuroscientific phenomena such as the N400 effect, which reveals how repetition and contextual information can reduce cognitive load (Rugg, 1985), provide neurobiological support for cognitive fluency. The reduction in N400 amplitude with repeated exposure to linguistic stimuli suggests that comprehensible input, as emphasized by Krashen (2002), facilitates efficient lexical access and meaning processing. Moreover, studies on the P600 effect indicate that grammatical violations activate effortful syntactic reanalysis, further aligning with the dual-processing distinction in language comprehension (Osterhout & Holcomb, 1992).

Furthermore, by drawing on Birdsong's (1994) application of DPT in decision-making within SLA, we explore how cognitive processes impact hypothesis formation and knowledge transfer from L1 to L2. Empirical studies show that L2 learners rely on System 1 for intuitive judgments of grammaticality when fluency increases, while less proficient learners engage System 2 for rule-based analysis (Dabrowska, 2019). This shift in processing mode mirrors findings in other cognitive domains, reinforcing DPT's applicability to language learning.

Expanding on these ideas, recent research on intelligent machines and their display of Fast and Slow thinking introduces a compelling dimension to this discussion. As artificial intelligence (AI) systems become increasingly sophisticated, their decision-making processes often mirror human cognitive patterns. Behavioral studies indicate that humans perceive AI-driven decisions as either intuitive (e.g., neural networks recognizing speech patterns) or deliberative (e.g., *rule-based translation models*) (Rahwan et al., 2019). Bonnefon and Rahwan (2020) argue that humans instinctively apply mental

frameworks used to assess human behavior, such as detecting signs of Fast (intuitive) and Slow (analytical) thinking, to their interactions with AI. This phenomenon underscores the universality of DPT principles in interpreting agency, whether human or artificial.

When machines exhibit characteristics of *Thinking Fast or Thinking Slow*, humans instinctively draw inferences about the tasks these machines can handle and the level of trust they merit. Research indicates that people trust AI-driven decisions more when they align with the expected cognitive style for a given task (Glikson & Woolley, 2020). This behavior mirrors the way we assess human capabilities and reliability. Understanding these signals, both in humans and machines, offers valuable insights into the learning process itself, particularly in contexts where artificial systems are integrated into educational tools or adaptive language platforms. However, as Bonnefon and Rahwan caution, this raises ethical concerns regarding the potential for manipulation by machines designed to mimic human cognitive styles.

By incorporating these multidimensional perspectives, spanning human cognitive processes, language learning, and human-machine interactions, this chapter underscores the relevance of DPT beyond traditional contexts. It not only highlights the intersection of cognitive science and language learning but also proposes a novel framework for optimizing strategies in an increasingly interconnected and technologically advanced world.

Daniel Kahneman's research into cognitive processes provides significant insights into how learners acquire language and manage cognitive demands. Kahneman's distinction between System 1 and System 2 processing is particularly relevant for understanding the dual nature of language learning. System 1 represents the automatic, intuitive processes that operate effortlessly and rapidly. This system is essential for implicit learning, where learners absorb language patterns and vocabulary through exposure and repetition without conscious effort. For instance, when learners are immersed in a language-rich environment, they naturally pick up on frequent phrases and vocabulary through System 1 processing. This automaticity fosters fluency and helps learners develop a sense of the language's rhythm and structure.

In contrast, System 2 involves more deliberate, analytical thought processes. This system is engaged when learners tackle complex grammatical rules, analyze language structure, or consciously attempt to use new vocabulary in specific contexts. System 2 processing requires effort and concentration, making it crucial for explicit learning tasks. For example, when learners practice writing sentences or solving grammar exercises, they are engaging System 2 to apply rules and structures they have been taught. Effective language instruction should therefore include strategies that engage both systems. Immersive experiences and extensive practice can enhance System 1 processing, while explicit teaching methods and problem-solving activities can develop System 2 skills. This dual

approach helps learners achieve a balance between intuitive language use and analytical understanding.

Kahneman's work on cognitive load further informs our understanding of language acquisition. Cognitive load theory posits that individuals have a limited capacity for processing information at any given time. In language learning, managing cognitive load is crucial to ensuring that learners do not become overwhelmed by too much information. For instance, introducing new vocabulary and grammatical concepts gradually, using clear explanations, and providing contextual support can help manage cognitive load. This approach aligns with Krashen's theory of comprehensible input, which suggests that learners benefit from exposure to language that is just beyond their current level of proficiency. By presenting information that is challenging yet accessible, educators can optimize cognitive resources and facilitate effective learning.

Additionally, Kahneman's concept of ego depletion offers valuable insights into the cognitive and motivational aspects of language learning. Ego depletion refers to the idea that self-control and cognitive resources are finite and can be depleted with prolonged effort. In the context of language acquisition, this means that learners' ability to engage in effortful cognitive tasks may diminish over time. For instance, after extended periods of study or practice, learners might experience reduced motivation and cognitive fatigue, which can affect their performance and retention. To mitigate the effects of ego depletion, educators can incorporate strategies such as varying learning activities, providing regular breaks, and offering positive reinforcement. Encouraging learners to engage in enjoyable and meaningful practice can help replenish their cognitive and motivational resources, supporting sustained language learning.

By integrating Kahneman's insights into language acquisition, we can develop a more nuanced understanding of how cognitive processes influence learning. The balance between System 1 and System 2 processing, effective management of cognitive load, and strategies to address ego depletion all contribute to a more comprehensive approach to language instruction. This understanding not only enriches theoretical perspectives on language learning but also offers practical applications for designing effective teaching strategies and supporting learner outcomes.

Understanding the cognitive processes involved in language learning, particularly the reliance on System 1 and System 2, can significantly inform instructional strategies. Being aware of whether learners are primarily engaging System 1 or System 2 enables instructors to tailor their courses to manage cognitive load effectively. System 1 processing is automatic and intuitive, often engaged during activities such as immersion and repetitive practice. Conversely, System 2 processing requires more conscious, effortful engagement, which is crucial for tasks that involve deliberate analysis and application of language rules.

When learners predominantly rely on System 1, they benefit from exposure to language in context and repetitive practice, which helps build fluency and automaticity. However, if learners are tasked with processing too much new information simultaneously, or if the material is overly complex, their cognitive load can become overwhelming. This can impede their ability to absorb and internalize new language effectively. To manage cognitive load and support System 1 processing, instructors can incorporate strategies such as gradual introduction of new concepts, repetitive exposure, and contextual practice. This approach aligns with Krashen's principle of comprehensible input, which emphasizes that learners should be exposed to language that is slightly above their current proficiency level, ensuring it is challenging yet manageable.

On the other hand, when learners engage System 2, they are involved in more effortful, analytical tasks such as studying grammar rules or performing complex language exercises. These tasks can lead to high cognitive load if not carefully designed. To avoid overwhelming learners, instructors can use text enhancements to make information more accessible and manageable. For example, incorporating visual aids, annotations, and structured practice activities can help break down complex information into more digestible chunks. This approach ensures that learners can focus their cognitive resources on understanding and applying new concepts without becoming overwhelmed.

Integrating text enhancements into language instruction can thus support both System 1 and System 2 processing by calibrating the cognitive load according to learners' needs. Enhanced texts can make new vocabulary and grammatical structures more salient, facilitating easier processing and retention. By providing clear, contextualized examples and interactive elements, instructors can support learners in transitioning from initial exposure to deeper understanding and application. This tailored approach not only aligns with Krashen's principles but also ensures that learners are neither under-challenged nor overburdened by the complexity of the material.

In sum, awareness of the cognitive processes learners engage in allows instructors to design courses that balance the demands placed on System 1 and System 2 processing. By managing cognitive load through careful integration of text enhancements and adhering to principles of comprehensible input, educators can create a supportive learning environment that fosters both intuitive and analytical language skills. This balanced approach enhances the overall effectiveness of language instruction and supports learners in achieving proficiency with greater ease.

Incorporating gamification strategies into language learning can significantly enhance the effectiveness of instructional approaches by creating a more engaging and supportive learning environment. Gamification reinforces elements commonly found in games, such as points, badges, and leaderboards, to motivate learners and make the learning process more interactive and enjoyable.

This approach not only fosters a positive learning atmosphere but also contributes to cognitive ease and reduces the affective filter.

Gamification strategies can recreate a familiar and friendly environment by integrating game-like elements that resonate with learners' everyday experiences. For example, incorporating challenges, rewards, and interactive scenarios can make language learning more relatable and less intimidating. By framing learning activities as games, instructors can tap into learners' intrinsic motivation and natural enjoyment of play, which facilitates a more relaxed and open mindset towards language acquisition.

Creating an environment that emphasizes cognitive ease is crucial for effective learning. When learners engage in gamified activities, they are often more relaxed and focused, experiencing a state of flow where cognitive demands are balanced with their skill level. This alignment reduces the cognitive load associated with language learning tasks, allowing learners to process information more effortlessly and efficiently. By presenting new vocabulary and grammatical structures in a playful and interactive context, learners can approach these elements with less anxiety and greater confidence.

Moreover, gamification can play a key role in lowering the affective filter, a concept introduced by Krashen to describe the emotional barriers that can hinder language acquisition. When learners perceive the learning environment as enjoyable and supportive, they are more likely to engage with the material and participate actively. Game-like elements such as immediate feedback, progress tracking, and personalized challenges can help create a sense of achievement and encourage persistence. This positive emotional state reduces anxiety and fear of making mistakes, which in turn enhances learners' willingness to take risks and practice their language skills.

The integration of gamification strategies aligns with the principles of comprehensible input and cognitive load management by offering a context that is both motivating and cognitively manageable. By designing language learning activities that incorporate game mechanics, instructors can create an environment that facilitates both System 1 and System 2 processing. For instance, simple, repetitive tasks can cater to System 1's automatic processing, while more complex challenges can engage System 2's analytical processing in a supportive and engaging manner.

In summary, gamification strategies offer a powerful means of enhancing language learning by creating a familiar, friendly environment that supports cognitive ease and lowers the affective filter. By integrating game elements into instructional design, educators can foster a positive learning experience that motivates learners, reduces cognitive load, and promotes active engagement with the material. This approach not only aligns with theoretical principles of language acquisition but also contributes to a more effective and enjoyable learning process.

Creating the right cognitive conditions is essential for optimizing the capacity to store and retain new vocabulary in language learning. When learners are provided with an environment that minimizes cognitive overload and maximizes engagement, their ability to internalize and recall new words improves significantly. By strategically incorporating enhancements, such as personalized feedback, contextualized practice, and interactive activities, educators can create conditions that align with learners' cognitive needs. These enhancements facilitate better encoding and retrieval of vocabulary by ensuring that the learning experience is both manageable and meaningful. When learners are not overwhelmed by excessive complexity and are actively engaged through well-designed text enhancements, they are more likely to develop a robust mental lexicon. Thus, achieving the right balance in cognitive load and engagement through targeted enhancements can significantly enhance the effectiveness of vocabulary acquisition, setting a solid foundation for further exploration of the mental lexicon.

1.9.2 Evidence In Mental Lexicon Studies

A central question in cognitive science is how semantic information is mentally represented. Two dominant theories address this issue: language-based distributional semantic models, which propose that word meaning is based on patterns of word co-occurrence in language (Harris, 1954), and semantic network models, which suggest that words are represented as a network where closely related meanings are more directly linked (Collins & Quillian, 1969).

Kroll and Stewart's Revised Hierarchical Model (RHM) provides an insightful framework for understanding the lexical organization and access in bilinguals. The model posits that bilingual individuals maintain two separate but interconnected lexical systems for their first language (L1) and second language (L2). In the early stages of L2 acquisition, learners heavily rely on their L1 lexicon, often translating from L1 to L2, as the connections between L2 words and their meanings are initially weak. As proficiency in L2 increases, these direct links between L2 words and their conceptual meanings strengthen, reducing the reliance on L1. Consequently, advanced learners can access meanings directly from L2 without the mediation of L1 (Kroll & Stewart, 1994). The model suggests strong lexical connections from L1 to L2 initially, which become more balanced as L2 proficiency grows. Furthermore, conceptual links, direct connections between words and their meanings, are robust for L1 but initially weak for L2, gradually strengthening with advanced proficiency (Tokowicz & Kroll, 2007). This progression highlights the shift from dependency on L1 to more autonomous L2 use. The RHM has significant implications for second language acquisition, emphasizing the importance of developing direct conceptual links in L2 and suggesting that effective learning strategies and teaching approaches should focus on immersive experiences and contextual L2 usage.

to build these connections (Kroll & Tokowicz, 2005). While the RHM has been influential, it is also subject to critiques and extensions, particularly regarding the variability in the strength of lexical connections based on individual differences and contextual factors such as the similarity between L1 and L2, age of acquisition, and frequency of language use (Talamas et al., 1999).

While the concept of the mental lexicon as a complex network of interconnected words is intellectually appealing, Meara cautions that translating this abstract model into tangible, empirical research is far from straightforward. He suggests that despite the theoretical promise, the current state of research has yet to fully capture the intricate nature and dynamics of the mental lexicon, indicating a significant gap between conceptualization and practical realization (Meara, 1982).

Meara proposes a paradigm shift in vocabulary research, advocating for a broader perspective beyond the granular analysis of individual words. He argues that a comprehensive understanding of lexicons, as holistic systems, is more fruitful than an in-depth examination of their constituent parts. By drawing an analogy to forestry, Meara emphasizes the importance of seeing the forest (lexicon) rather than becoming preoccupied with the trees (individual words). He criticizes existing models that risk overlooking the systemic properties of lexicons by excessively focusing on word-level details. (Meara, 1996)

We further integrate into our research the understanding that the organization of the mental lexicon varies between native and non-native speakers. Native speakers typically organize their mental lexicon paradigmatically, which means they group words based on shared grammatical features and lexical categories. This paradigm-based organization reflects a deeper integration of words within lexical categories, facilitating more abstract and rule-based connections.

As individuals gain a deeper understanding of vocabulary, semantic links, relationships based on meaning, become more robust and tend to surpass phonetic links, relationships based on sound (Wolter, 2001). This shift reflects an increase in lexical competence, where words become integrated into the core vocabulary and thus exhibit more complex interrelations.

The transition from syntagmatic to paradigmatic associations is facilitated when words are firmly established within the core vocabulary. In this model, paradigmatic responses, associations based on shared lexical categories, are prominent among words located centrally in the mental lexicon. Syntagmatic connections, which are based on word co-occurrence in specific contexts, are found among words situated slightly further out. Phonological associations, relating to sound similarities, are represented on the periphery of the lexical network (Wolter, 2001).

To investigate this, researchers Wang et al. (2022) examined core vocabulary, the set of words most central in the mental lexicon, through experiments designed to test which measure of core vocabulary aligns most closely with human behavior. In their experiment, participants engaged in a word-

guessing game where they identified target words based on sets of semantically related hint words. These hint and target words were generated from different core vocabulary lists corresponding to the two theories (De Deyne et al., 2021). The results of (Wang et al., 2022) indicated that the type of hint words did not significantly impact performance. However, target words derived from word associations led to better identification compared to those derived from natural language distributional statistics (Vankrunkelsven et al., 2018). Follow-up analyses ruled out several alternative explanations. These findings suggest that the semantic information reflected in word associations may be more effective for the efficient identification of lexical meanings (Wierzbicka, 1996).

The real challenge in vocabulary acquisition lies in observing and predicting the growth of a learner's lexical inventory. Understanding the mechanisms and connections between existing concepts and newly acquired ones can provide valuable insights for course design and didactic interventions tailored to an individual's current state of acquisition.

Among the various findings and approaches, several common parameters can be identified for consideration in vocabulary acquisition. These include the age of acquisition, usage frequency, semantic transparency of a given word, and the native language of the learner. Lexical inventories can share common ground in word morphological similarity, especially when derived from the same linguistic family, such as Italian-Latin and English-Latin stems. This phenomenon can decrease the semantic transparency of a word, thus increasing the difficulty in detecting its meaning.

As demonstrated in the model for semantic growth proposed by Steyvers & Tenenbaum (2005), each new word or concept is connected to an existing network by differentiating the connectivity pattern of an existing node. This model generates appropriate small-world statistics and power-law connectivity distributions, and it also proposes a potential mechanistic basis for the effects of learning history variables (such as age of acquisition and usage frequency) on behavioral performance in semantic processing tasks. Steyvers & Tenenbaum (2005) findings indicate that semantic networks built from Latent Semantic Analysis (LSA) representations generally lack the large-scale statistical regularities characteristic of real-world semantic networks.

In Rodd's (2025) *Settling Into Semantic Space: An Ambiguity-Focused Account of Word-Meaning Access*, she emphasizes that word-meaning access is a highly interactive process. A variety of linguistic and nonlinguistic cues available during comprehension can influence the lexical-semantic settling process. Rodd argues that fluent disambiguation of ambiguous words within sentence contexts is largely dependent on the system's sensitivity to all relevant cues, which reliably predict the intended meaning. Furthermore, learning mechanisms play a crucial role in shaping and maintaining high-quality lexical-semantic representations throughout the lifespan, enabling individuals to adapt to changing linguistic environments and utilize recent experiences to enhance the

availability of meanings most likely to be encountered in the near future. These mechanisms ensure sensitivity to the cues that support fluent disambiguation.

Lexical acquisition has consistently been a primary focus in the design of language courses and digital applications, as detailed in the following sections. However, it is important to highlight that the development of tools for lexical acquisition often fails to adequately address the need for tracking learners' lexical inventories. This oversight can hinder the effectiveness of these tools by not fully considering the current state of each learner's vocabulary knowledge.

On the other hand James et al. (2023) investigate the encoding and consolidation processes in word learning, emphasizing the differential effects of repeated exposure on word-form and meaning knowledge. Consistent with previous research, significant gains in word-form knowledge were observed with repeated testing over a week, supporting the hypothesis that new word knowledge is strengthened offline through hippocampal reactivation during sleep (Davis & Gaskell, 2009; McClelland et al., 1995). However, gains in meaning knowledge were limited, possibly due to fewer re-exposure opportunities. This finding aligns with the complementary learning systems perspective, which suggests that the integration of new word knowledge with existing representations varies across different aspects of word knowledge (Goossens, 2014; L. M. James Eleanor, Gaskell, Gareth M. ,. Henderson, 2020). Prior studies, such as Henderson et al. (2013), L. M. James Eleanor, Gaskell, Gareth M. ,. Henderson, (2019), and (Tamminen, 2010), similarly report that word-form recall improves with repeated retrieval practice, whereas meaning recall remains stable despite repeated encoding opportunities. These results underscore the importance of balanced re-exposure to both word-forms and meanings to optimize vocabulary acquisition.

The findings on encoding and consolidation processes in word learning have significant implications for second language (L2) acquisition. These insights emphasize the differential effects of repeated exposure on word-form and meaning knowledge and highlight the critical role of balanced re-exposure in vocabulary acquisition.

Moreover, repeated exposure to word forms significantly enhances word-form knowledge, as demonstrated by improvements observed through repeated testing over the course of a week in the study discussed. This finding aligns with previous research, which suggests that offline hippocampal reactivation during sleep strengthens word-form knowledge (Davis & Gaskell, 2009; McClelland, 1995). For second language (L2) learners, consistent practice with new word forms leads to better retention and recall. Strategies such as spaced repetition and regular testing are particularly effective in reinforcing the phonological aspects of new vocabulary (Schmitt, 2000). However, the limited gains in meaning knowledge despite repeated exposure underscore the need for more effective teaching methods in L2 semantics. The study indicates that, in the absence of re-exposure, semantic

knowledge does not improve as robustly as word-form knowledge. Consequently, L2 instructional approaches should include repeated and varied exposures to new word meanings, integrating contextualized practice and meaningful communication activities (Nation, 2001). Furthermore, the enhancement of word-form recall following sleep highlights the critical role sleep plays in consolidating L2 vocabulary. Educators should stress the importance of adequate sleep for memory consolidation after intensive vocabulary learning sessions (Tamminen, 2010; James, 2019). Promoting healthy sleep habits can serve as an indirect but effective strategy for supporting L2 vocabulary acquisition.

The role of sleep in vocabulary consolidation is another area ripe for exploration. While current research underscores its importance in solidifying word-forms, the precise mechanisms by which sleep influences semantic memory in second language acquisition remain less clear. Further investigations into how different sleep patterns, such as napping versus nocturnal sleep, impact various aspects of vocabulary learning could lead to more targeted recommendations for learners, optimizing the timing and scheduling of their study sessions.

As technological advancements continue to reshape educational landscapes, adaptive learning platforms hold significant promise in enhancing language acquisition. These platforms can monitor individual progress in both word-form and semantic knowledge, offering tailored learning paths that adjust in real-time to the learner's needs. For example, learners who exhibit strong word-form recall but struggle with meanings could be directed towards exercises that reinforce semantic understanding. Integrating natural language processing tools into these platforms could provide dynamic feedback, allowing learners to focus on areas that need improvement.

Cultural relevance in vocabulary instruction is another critical consideration. Learners often encounter words and phrases deeply rooted in the target language's culture, and understanding these contexts can greatly enhance semantic comprehension. By incorporating authentic materials, such as articles, videos, and social media content, learners can engage with vocabulary in a manner that bridges the gap between word-form recognition and meaningful usage.

Long-term retention of vocabulary is crucial for true mastery, and practical application plays a key role in this process. Providing opportunities for learners to use new vocabulary in real-life communication, such as through role-playing, debates, or collaborative projects, can help solidify both word-form and meaning knowledge. Longitudinal studies tracking learners over extended periods could offer valuable insights into which instructional strategies are most effective for sustained vocabulary acquisition.

In the early stages of language learning, non-native speakers often acquire vocabulary in a syntagmatic manner, focusing on word sequences and phrases as they appear in context. This

approach is essential for understanding the structural patterns of the target language and for building basic communicative competence. However, as learners progress, their exposure to connections and semantic maps becomes increasingly beneficial. Engaging with these connections helps learners to link new vocabulary with words they already know whether in the target language or through cognates and synonyms in their native language (L1). This process not only reinforces their existing knowledge but also aids in the integration of new terms into their mental lexicon. For instance, when a learner encounters a new word that is semantically related to a familiar term in their L1, the association can facilitate quicker recall and more robust understanding. By drawing on their prior linguistic knowledge, learners can create a more interconnected and nuanced vocabulary system, which ultimately supports their overall language acquisition process. This semantic mapping approach is particularly effective in moving learners from rote memorization to a deeper, more flexible understanding of the target language, allowing them to navigate and use new vocabulary with greater confidence and accuracy.

These considerations emphasize the importance of a comprehensive approach to vocabulary instruction, one that balances the need for repetition with the rich, contextualized use of language. As research continues to evolve, the combination of effective teaching strategies, adaptive technologies, and an understanding of cognitive processes like sleep will undoubtedly contribute to more personalized and effective methodologies for second language learners.

1.10 The CEFR Framework: Linking Proficiency to Pedagogy

The Common European Framework of Reference for Languages (CEFR) has become one of the most influential frameworks for describing, assessing, and guiding language learning across educational systems worldwide. First introduced by the Council of Europe (2001) and later expanded in the Companion Volume (2020), the CEFR provides a shared reference point for language proficiency, offering descriptors that transcend individual curricula and allow for comparability across contexts. Its impact on second language acquisition (SLA) research and practice lies in its capacity to operationalize the otherwise abstract notion of communicative competence into observable, assessable ‘can-do’ statements.

At its core, the CEFR divides language proficiency into six levels: A1 (Beginner), A2 (Elementary), B1 (Intermediate), B2 (Upper Intermediate), C1 (Advanced), and C2 (Mastery). These levels are cumulative and progressive, describing the learner’s evolving ability to perform tasks in listening, reading, spoken interaction, spoken production, and writing. The framework emphasizes that language competence is not monolithic but multidimensional, integrating linguistic, pragmatic, and

sociolinguistic components. This multidimensionality aligns closely with SLA theories such as Canale and Swain's (1980) model of communicative competence, which distinguishes grammatical competence, sociolinguistic competence, discourse competence, and strategic competence. By translating such constructs into practical descriptors, the CEFR provides a scaffold for curriculum design, teaching materials, assessment practices, and learner self-reflection.

The CEFR explicitly prioritizes communicative competence, shifting the focus from abstract grammatical knowledge to the ability to use language effectively in context. Each level specifies not only linguistic resources but also pragmatic capabilities and sociolinguistic appropriateness. For example, at the A1 level, learners are expected to understand and use familiar expressions and basic phrases, while at the B2 level, they should manage extended interactions on abstract or technical topics, and at C1–C2, they should demonstrate fluency, precision, and the ability to adapt language flexibly to different registers and purposes.

This alignment of linguistic and pragmatic descriptors reflects key insights from SLA research. Pragmatics scholars such as Searle (1969) and Taguchi (2019) have shown that communicative success depends on more than syntactic accuracy; learners must also master speech acts, politeness strategies, and discourse management. Similarly, phonetics and phonology researchers such as Ladefoged and Johnson (2014) highlight that intelligibility relies not only on phonetic accuracy but also on prosodic control. The CEFR captures this broader view of competence by embedding pragmatic and phonological performance criteria into its descriptors, ensuring that assessments and curricula reflect authentic communication.

Each of the CEFR's five key abilities, listening, reading, spoken interaction, spoken production, and writing, is elaborated through qualitative descriptors. For listening, progression is reflected in the ability to comprehend increasingly complex speech, from slow, clear sentences at A1 to idiomatic, colloquial, and fast-paced discourse at C2. Reading follows a similar path, beginning with short, simple texts and culminating in the ability to critically evaluate complex literature or technical material. Spoken interaction, perhaps the most distinctive of the CEFR categories, underscores the dialogic nature of communication, requiring learners to manage turn-taking, negotiate meaning, and adjust register. Spoken production highlights monologic skills such as presentations, storytelling, or formal argumentation, while writing ranges from simple notes to sophisticated academic or professional compositions.

What is particularly valuable is how these descriptors interlock across modalities. For example, vocabulary growth (Nation, 2001) underpins not only reading comprehension but also spoken fluency and written production. Syntax and morphology (Ellis, 2006; Aronoff & Rees-Miller, 2001) enable accurate sentence formation in both speech and writing, while pragmatic awareness (Searle, 1969;

Blum-Kulka et al., 1989) ensures that spoken interaction is contextually appropriate. In this way, the CEFR implicitly affirms the interconnectedness of linguistic branches, laying the groundwork for the holistic curricular design discussed in the following chapter.

The CEFR's influence on pedagogy is profound. By defining proficiency in terms of 'can-do' statements, it encourages educators to design tasks that mirror real-world communication. For instance, rather than drilling grammar in isolation, a CEFR-aligned syllabus at the B1 level might include role-plays where learners narrate experiences or exchange opinions. At C1, tasks might include summarizing complex information or mediating a discussion. This task-based orientation resonates with Ellis's (2003) advocacy for task-based language teaching (TBLT), which views tasks as the central unit of planning, instruction, and assessment.

Moreover, the CEFR supports learner autonomy by providing benchmarks that learners can use for self-assessment. By situating themselves within the A1–C2 continuum, learners can monitor progress, set goals, and recognize gaps. This aligns with Krashen's (1985) emphasis on comprehensible input at the right level ($i+1$), as CEFR levels can guide material selection that is challenging yet accessible. Similarly, Swain's (1995) Output Hypothesis finds a natural home in CEFR descriptors for spoken production and interaction, which require learners to go beyond comprehension and actively engage in communicative tasks.

The Companion Volume (Council of Europe, 2020) expands the CEFR's scope by adding descriptors for mediation, online interaction, and plurilingual/pluricultural competence. Mediation, for example, involves reformulating or simplifying messages for others, summarizing texts, or facilitating dialogue across cultural and linguistic boundaries. This addition reflects the growing recognition in applied linguistics that language learning is not only about individual competence but also about collaborative meaning-making in multilingual environments. Online interaction descriptors respond to the digital transformation of communication, emphasizing skills such as participating in chats, moderating forums, or managing video calls. These updates align with contemporary CALL and ICALL research (Chapelle, 2001; Heift & Schulze, 2007), which explores how technology can support authentic interaction and mediation in virtual spaces.

The CEFR has also reshaped language assessment, particularly in high-stakes contexts. International exams such as IELTS, Cambridge English Qualifications, and many national proficiency tests are now benchmarked to CEFR levels. This standardization facilitates mobility, mutual recognition of qualifications, and comparability of outcomes across institutions and countries. However, researchers caution that aligning tests to CEFR levels is not straightforward: descriptors are intentionally broad and context-sensitive, which can make precise calibration challenging (Fulcher, 2004). Still, the

CEFR remains the most widely accepted framework for linking instruction, assessment, and learner progression.

Despite its widespread adoption, the CEFR has attracted critique. Some scholars argue that the framework's broad descriptors can be vague, leaving significant room for interpretation (Alderson, 2007). Others question whether the six-level scale adequately captures the complexity of language proficiency, particularly in specialized or academic registers. Still others note that while the CEFR claims to be neutral, it reflects European communicative traditions and may not fully account for global diversity in language use. Nevertheless, these critiques often lead to refinements and adaptations rather than wholesale rejection. The CEFR's flexibility allows it to be contextualized, supplemented with local syllabi, and adapted for specific learner populations.

For SLA research, the CEFR offers a bridge between theory and practice. By providing operationalized descriptors of proficiency, it allows researchers to map theoretical constructs, such as input processing, noticing, or pragmatic development, onto observable performance outcomes. For curriculum designers, it provides a roadmap for sequencing instruction and aligning assessment with real-world communicative abilities. Importantly, it also affirms the necessity of integrating all linguistic branches, phonetics, phonology, morphology, syntax, semantics, pragmatics, and grammar, into a coherent whole, since proficiency descriptors invariably touch on multiple areas at once.

Thus, the CEFR serves as a conceptual and practical anchor for language education. It encapsulates the communicative orientation that SLA research has long advocated, while also providing tools for pedagogy, assessment, and learner autonomy. Positioned between theoretical models of acquisition (Krashen, Swain, Long, Chomsky) and the practical integration of linguistic fields, the CEFR lays the groundwork for a curriculum that is both empirically informed and pedagogically actionable.

Thus, the CEFR serves as a conceptual and practical anchor for language education, encapsulating the communicative orientation that SLA research has long advocated, while also providing tools for pedagogy, assessment, and learner autonomy. Positioned between theoretical models of acquisition (Krashen, Swain, Long, Chomsky) and the practical integration of linguistic fields, the CEFR lays the groundwork for a curriculum that is both empirically informed and pedagogically actionable. Building on this framework, the next chapter turns to the major linguistic branches, phonetics and phonology, morphology, syntax, semantics, pragmatics, and grammar, and examines how each contributes to a holistic approach to SLA curriculum design.

1.10.1 Curriculum Design

A well-structured language curriculum or didactic intervention should adopt a holistic approach, where various linguistic fields are integrated to reflect the complexity and interconnectedness of

language use in real-world communication. Rather than isolating linguistic components such as phonetics, morphology, syntax, semantics, and pragmatics, an effective syllabus interweaves these areas to create a more meaningful learning experience. CEFR provides the overarching framework for describing and assessing language proficiency, but its descriptors are only meaningful when grounded in the actual linguistic fields that make up communicative competence. A well-structured language curriculum or didactic intervention should therefore adopt a holistic approach, where phonetics, morphology, syntax, semantics, and pragmatics are not treated in isolation but are interwoven to reflect the complexity and interconnectedness of real-world language use. By mapping these linguistic domains onto CEFR proficiency levels, educators can ensure that instruction not only aligns with internationally recognized standards but also addresses the specific structural and functional components of language that learners must master to progress effectively.

This integrated approach ensures that learners can understand how different aspects of language function together in practical contexts, thereby strengthening both their theoretical understanding and applied communicative skills. For instance, thematic units can be designed to address multiple linguistic dimensions simultaneously, such as syntax, semantics, and pragmatics, within meaningful communicative tasks. This holistic vision builds upon the complementary insights of Krashen's Input Hypothesis (1982), Swain's Output Hypothesis (1985, 1995), and Long's Interaction Hypothesis (1983, 1996). Each of these frameworks emphasizes that language acquisition is not driven by a single factor but by the dynamic interplay of input, output, and interaction, where comprehension, production, and negotiation collectively facilitate learning. Having considered these foundational perspectives, it is therefore essential to examine how the core linguistic branches, phonology, morphology, syntax, semantics, and pragmatics, contribute to the overall process of second language acquisition.

At the grammatical level, Chomsky's (1981) theory of Universal Grammar provides the foundation for understanding how learners internalize complex rules and constraints. His principles-and-parameters framework emphasizes that all human languages share a set of innate structural principles, while cross-linguistic variation results from parameter settings. In SLA research, this has led to debates on whether adult learners retain full access, partial access, or no access to Universal Grammar, with implications for grammar instruction. White (2003) and Hawkins (2001) show that second language learners often respect deep syntactic constraints, such as subadjacency or binding principles, but struggle with surface-level morphology or agreement. This suggests that while learners' internal grammar may be generative and rule-governed, performance limitations and transfer from the first language can create persistent challenges. For curriculum design, this means that grammar cannot be

taught as a detached set of prescriptive rules; rather, it must be embedded within communicative practice that highlights both structural accuracy and meaningful use.

Morphology, the study of word formation and inflection, is equally essential. Aronoff and Rees-Miller (2001) stress that morphemes, as the smallest units of meaning, serve as the building blocks of vocabulary and syntax. In SLA, morphological development often lags behind syntactic growth, as learners produce complex sentences while variably marking tense, agreement, or case. McQuillan and Krashen (2007) emphasize the importance of repeated, meaningful exposure to morphological patterns through comprehensible input, which can help learners internalize inflectional systems more naturally. However, explicit attention is often necessary when forms are not salient in the input, particularly for low-frequency or irregular morphology. This resonates with Ellis's (2003) argument for task-based grammar instruction: rather than isolating morphology, instruction should encourage learners to notice and practice morphological patterns in context, such as through guided writing, reading passages with recurring forms, or input enhancement strategies that draw attention to inflections without disrupting meaning.

Syntax, or sentence structure, governs how words are combined to form meaningful utterances. As Ellis (1994, 2003) explains, learners progress through predictable stages in acquiring syntactic structures such as negation, question formation, or subordination. Yet instruction can accelerate progress by making target structures salient and by providing feedback that encourages restructuring. Chomsky's (1981) generative framework views syntax as constrained by innate principles, while usage-based approaches emphasize the frequency and salience of patterns in the input. Both perspectives are instructive for pedagogy: learners need abundant exposure to authentic syntactic constructions, but they also benefit from tasks that create communicative pressure to produce increasingly complex sentences. Activities such as collaborative storytelling, debates, or problem-solving tasks naturally elicit subordinate clauses, cohesive devices, and varied word orders, giving learners opportunities to test and refine their syntactic hypotheses.

Semantics, the study of meaning, provides the interpretive core of communication. Cruse (2000) emphasizes the importance of lexical semantics, the meanings of words and their relationships, including polysemy, synonymy, and antonymy, as well as compositional semantics, which explains how word meanings combine in larger units. For L2 learners, building semantic networks is crucial for vocabulary expansion and comprehension. Nation (2001) highlights that vocabulary knowledge must be measured in terms of both breadth (the number of words known) and depth (the quality of understanding, including collocations, register, and connotation). Instruction that integrates semantic learning with contextual use, such as reading authentic texts, discussing meanings in discourse, and

practicing collocations, supports both receptive and productive competence. Importantly, semantics also interacts with pragmatics, as meaning is always filtered through context.

Pragmatics addresses how language use is shaped by social norms, intentions, and context. John Searle's (1969) speech act theory underscores that utterances are actions with illocutionary force, requests, promises, apologies, directives, that depend on felicity conditions. For second language learners, pragmatic competence often lags behind grammatical competence, leading to situations where formally correct sentences misfire socially. Blum-Kulka, House, and Kasper (1989) demonstrated that requests and apologies vary significantly across cultures, making pragmatics a key site of cross-linguistic transfer and potential breakdowns in communication. Taguchi (2019) argues that pragmatic competence requires both pragmalinguistic knowledge (the linguistic forms that realize functions) and sociopragmatic awareness (understanding norms of appropriateness in a given context). Swain's notion of comprehensible output complements this perspective, as being 'pushed' to produce language compels learners not only to retrieve appropriate vocabulary and grammatical forms but also to adjust their language to fit contextual and interpersonal demands, thereby sharpening pragmatic awareness and communicative sensitivity (Swain, 1985, 1995). From a pedagogical standpoint, pragmatic competence can be developed through scenario-based tasks, role-plays, and discourse completion activities that simulate authentic communicative situations and prompt learners to reflect on the social meaning and appropriateness of their linguistic choices (Taguchi, 2011; Ishihara & Cohen, 2010). Such activities operationalize the interaction between form, meaning, and context, aligning with the broader principles of task-based and interaction-oriented approaches to second language acquisition.

The branches of phonetics and phonology focus on the perception and production of speech sounds. Ladefoged and Johnson (2014) outline how articulatory phonetics describes the physical production of sounds, while phonology accounts for how these sounds are organized cognitively into systems of contrasts. SLA research shows that learners' first language phonological categories shape their perception of L2 sounds, often leading to persistent difficulty distinguishing contrasts such as /r/-/l/ or /ɪ/-/i:/ (Flege, 1995). Derwing and Munro (2015) argue that intelligibility, rather than native-like accuracy, should be the central goal of pronunciation instruction, as suprasegmental features like stress, rhythm, and intonation often carry greater weight in communication than individual segmental accuracy. Pedagogical interventions such as high-variability phonetic training, visualizations of formants, and tasks that integrate perception and production have proven effective in improving learners' phonological competence. Celce-Murcia et al. (2010) further emphasize that pronunciation should be embedded in communicative tasks rather than treated as isolated drills, aligning with the broader holistic approach to curriculum design.

Building on the discussion of communicative and pragmatic competence, grammatical competence, encompassing all linguistic domains, can be viewed as the structural backbone of communicative ability. Canale and Swain (1980) situated it within the broader construct of communicative competence, alongside sociolinguistic, discourse, and strategic competencies. In SLA, grammatical competence develops not only through explicit instruction but also through implicit processes of noticing, automatization, and restructuring. As VanPatten (1996) argues, learners naturally prioritize meaning over form, implying that pedagogical interventions must highlight grammar in ways that do not overwhelm cognitive processing. Techniques such as textual enhancement, input flooding, and task-based interaction thus provide balanced avenues for integrating comprehension with opportunities for form-focused learning.

Integrating these linguistic branches within a single curriculum requires a dynamic and layered approach. For example, a thematic unit on *Describing Experience* might engage phonetics (Ladefoged & Johnson, 2014) in the pronunciation of past tense forms, morphology (Aronoff & Rees-Miller, 2007) in regular and irregular verb marking, syntax (Chomsky, 1965) in narrative sequencing with subordinate clauses, semantics (Cruse, 1986) in lexical choices for emotion and evaluation, pragmatics (Searle, 1969) in the appropriate narration of events across formal and informal contexts, and grammar (Canale & Swain, 1980) in the control of tense–aspect distinctions.

Technology-enhanced environments, such as Intelligent Computer-Assisted Language Learning (ICALL) systems, further expand these possibilities by adapting input and feedback to learners' evolving profiles across all linguistic domains (Nation, 2001; Ellis, 2008). In this sense, the integration of linguistic theory and digital innovation reflects the broader trajectory of applied linguistics, one that seeks to bridge theory and practice in the service of effective language education. In sum, the study of SLA demonstrates that no single linguistic branch can fully explain or sustain language learning in isolation. Morphology and syntax provide the scaffolding of sentence structure; semantics ensures meaning; pragmatics governs contextual appropriateness; phonetics and phonology shape perception and production; and grammar binds these elements into a coherent system. Scholars from Chomsky (1965) to Nation (2001), from Aronoff and Rees-Miller (2007) to Searle (1969), Cruse (1986), and Ladefoged and Johnson (2014) each illuminate a distinct facet of this intricate whole. Collectively, their contributions build upon the foundational insights of Krashen (1982), Swain (1985), Long (1996), and Ellis (2008), underscoring the need for a holistic, integrated approach to curriculum design, one that recognizes the interdependence of linguistic branches and aligns them with learners' real-world communicative goals.

1.10.2 CEFR–Linguistic Branch Alignment

The CEFR levels (A1–C2) can be fruitfully mapped onto the major linguistic branches of SLA, offering a developmental perspective on how learners' competence unfolds. At the A1 beginner stage, the focus is primarily phonetic and phonological, as learners must establish control over core segmental contrasts and word stress to achieve intelligibility. As Ladefoged and Johnson (2014) emphasize, segmental contrasts and stress patterns form the perceptual foundation of communication, and failure to master them early can compromise intelligibility in later stages. Morphology at this stage is minimal, with learners managing only the most frequent inflections and simple verb forms (Aronoff & Rees-Miller, 2001). Syntax is limited to canonical SVO structures and formulaic questions, while vocabulary is restricted to survival needs, around 500–1,000 word families (Nation, 2001). Pragmatically, learners can perform only the most basic speech acts, such as greetings or requests (Searle, 1969). At this stage, the role of comprehensible input (Krashen, 1985) is paramount: learners need high-frequency, simplified input that is slightly above their level ($i+1$) to acquire core forms and patterns.

At the A2 level, competence expands. Learners consolidate phonological contrasts and begin to use sentence stress and intonation to distinguish meaning (Ladefoged & Johnson, 2014). Morphology extends to productive control of past tense and comparatives, with irregular forms beginning to appear, while syntax develops coordination and basic subordination, consistent with Ellis's (1994) observations on staged syntactic development. Vocabulary broadens to 1,000–2,000 word families (Nation, 2001), covering domains such as shopping, work, and health. Pragmatics extends to polite requests and refusals, marking the emergence of sociopragmatic awareness (Blum-Kulka, House, & Kasper, 1989). Here too, input remains critical, but learners begin to require opportunities for guided production, echoing Swain's (1985) claim that output helps consolidate forms by forcing attention to gaps.

At the B1 intermediate stage, learners' phonology acquires greater naturalness, with rhythm and connected-speech phenomena emerging (Derwing & Munro, 2015). Morphological competence extends to aspectual contrasts and modal verbs, while syntax develops through subordination, relative clauses, and reported speech. Ellis (2006) highlights that these developments benefit from input that makes form salient while retaining communicative meaning. Vocabulary expands to 2,000–3,500 word families (Nation, 2001), supporting discussions of semi-abstract topics and functional reasoning. Pragmatically, learners handle disagreements politely and repair misunderstandings, reflecting Searle's (1969) view of language as action within social norms. At this stage, Swain's Output Hypothesis becomes more visible: learners must be 'pushed' to produce language in extended discourse, which promotes noticing of gaps and hypothesis-testing about form.

At the B2 stage, learners achieve flexibility and begin to approximate professional competence. Phonology incorporates reduced forms and prosodic control across paragraphs. Morphology now includes passives, conditionals, and perfect aspect, with derivational processes supporting academic vocabulary growth. Syntax is varied and flexible, with learners producing complex subordination and nominalizations. Vocabulary reaches 3,500–5,000 families (Nation, 2001), sufficient for independent engagement with a wide range of texts. Pragmatics involves register-shifting, hedging, and efficient negotiation of meaning (Taguchi, 2019). Here, the balance of input and output is critical: learners require exposure to authentic discourse to expand structures and lexis, but also structured tasks that demand precise production, aligning with Swain's emphasis on metalinguistic reflection through collaboration.

At the C1 advanced level, phonology supports discourse-level structuring, with prosody signaling contrast, concession, and stance (Ladefoged & Johnson, 2014). Morphology and syntax support dense packaging of information, including nominalizations and complex embeddings, reflecting what Ellis (2003) describes as the interface between complexity and communicative demand. Vocabulary extends to 6,000–8,000 families (Nation, 2001), enabling idiomaticity, figurative language, and nuanced stance (Cruse, 2000). Pragmatically, learners deploy advanced politeness strategies, irony, and cultural adaptation (Searle, 1969; Blum-Kulka et al., 1989). At this stage, interaction (Long, 1996) and output play key roles, as advanced learners refine accuracy and appropriateness through negotiation of meaning, mediation, and feedback.

Finally, C2 mastery represents near-native command. Phonologically, learners use prosody and accent variation deliberately for stylistic purposes. Morphology and syntax are fully flexible, enabling elegant restructuring and stylistic precision. Vocabulary surpasses 8,000–10,000 families (Nation, 2001), extending to discipline-specific terminology and subtle connotations. Pragmatic competence encompasses humor, intertextuality, and effortless register-shifting across multilingual contexts (Taguchi, 2019). At this stage, Krashen's comprehensible input remains relevant in sustaining exposure to nuanced registers, but Swain's comprehensible output and Long's interaction are decisive for maintaining accuracy and rhetorical sophistication.

Viewed in this way, the CEFR not only describes language proficiency but also provides a roadmap for sequencing linguistic development in line with SLA theory. Each level demonstrates the interplay of phonology, morphology, syntax, semantics, and pragmatics, while also reflecting how learners depend on input for comprehension, output for consolidation, and interaction for refinement. This integration reinforces Ellis's (2006) call for explicit form-focused instruction alongside communicative practice, Nation's (2001) emphasis on systematic vocabulary growth, and Searle's (1969) recognition of pragmatics as inseparable from communication. The CEFR thus affirms the

value of a holistic, theory-driven syllabus that interweaves linguistic domains while guiding learners along the trajectory toward communicative competence.

1.11 Conclusion

Taken together, the theories and frameworks reviewed, from Krashen's Input Hypothesis (Krashen, 1982) and Swain's Output Hypothesis (Swain, 1985, 1995) to Long's Interactionist model (Long, 1996), Ellis's work on form-focused instruction (Ellis, 2008), Chomsky's insights on Universal Grammar (Chomsky, 1965, 1986, 1995), Nation's emphasis on vocabulary learning (Nation, 2001, 2013), Ladefoged and Johnson's phonetic foundations (Ladefoged & Johnson, 2014), and Searle's exploration of pragmatics and speech acts (Searle, 1969), offer a multidimensional understanding of second language acquisition.

They reveal language learning as a process shaped by both implicit and explicit mechanisms, by input and output, and by the interplay of cognitive, linguistic, and social factors. The Common European Framework of Reference (CEFR) (Council of Europe, 2001; North, 2014) provides a unifying framework for aligning these perspectives into curricula that integrate phonology, morphology, syntax, semantics, and pragmatics into coherent developmental pathways.

Yet while these theoretical and pedagogical insights remain indispensable, the growing role of technology in education compels us to consider how they can be translated into practice within digital learning environments. The next chapter therefore turns to innovations in Computer-Assisted Language Learning (CALL) and its evolution into Intelligent CALL (ICALL), examining how advances in technology can operationalize SLA theory, scaffold learning, and create adaptive systems that extend communicative competence into new, interactive domains

2 Innovations in Second Language Learning: Computer-Assisted and Intelligent CALL

2.1 Introduction

The integration of technology into language education has fundamentally reshaped the ways in which learners engage with linguistic input, practice communicative skills, and receive feedback. What began in the mid-twentieth century as drill-and-practice software under the umbrella of Computer-Assisted Language Learning (CALL) has evolved into increasingly sophisticated, interactive platforms that seek to replicate and enhance authentic learning experiences. Early CALL applications mirrored behaviorist approaches, emphasizing repetitive grammar and vocabulary drills. These tools, while limited, demonstrated the potential of computers to provide individualized practice and instant feedback. With the rise of communicative language teaching, CALL entered a new phase, focusing on meaning, interaction, and the integration of authentic materials. This evolution aligned more closely with the theoretical frameworks outlined in the previous chapter, including Krashen's emphasis on comprehensible input, Swain's output hypothesis (Swain, 1985), Long's interactionist model (Long, 1996), and Ellis's advocacy for task-based learning (R. Ellis, 2003).

2.2 Early CALL and Its Drill-Based Roots

The origins of Computer-Assisted Language Learning (CALL) can be traced back to the 1960s and 1970s, when rapid developments in computer technology encouraged educators and linguists to experiment with digital tools for language instruction. During this early phase, CALL was strongly influenced by the behaviorist paradigm of learning, which conceptualized acquisition as the result of habit formation through repetition and reinforcement (Skinner, 1957; Warschauer & Healey, 1998). Language learning was understood primarily as the mastery of correct linguistic forms, grammar rules, vocabulary items, and pronunciation patterns through drill-and-practice routines. Computers were particularly well-suited to this instructional model: they could deliver endless repetitions of the same task, provide immediate feedback, and record learner performance data, capabilities that far exceeded what traditional classrooms could offer at the time (Levy, 1997; Beatty, 2010).

Prominent systems such as PLATO (Programmed Logic for Automated Teaching Operations), developed at the University of Illinois in the early 1960s, exemplify the ambitions and technological

optimism of early CALL (Ahmad, Corbett, Rogers, & Sussex, 1985; Levy, 1997). These platforms enabled learners to complete fill-in-the-blank exercises, multiple-choice drills, and grammar-translation tasks within an interactive computer-based environment. Although rudimentary by modern standards, PLATO represented a groundbreaking innovation in language pedagogy, introducing features such as individualized pacing, automated feedback, and data-driven tracking of learner progress, capabilities that were nearly impossible to achieve in traditional classrooms at the time (Beatty, 2010). Other early projects followed similar design principles, emphasizing vocabulary memorization, verb conjugation drills, and pronunciation exercises using the limited audio capabilities then available in educational computing.

The advantages of this early CALL model were evident. For one, the computer's patience and consistency were valued assets. Unlike human teachers constrained by time and attention, the computer could provide unlimited practice opportunities without fatigue or bias. Learners could proceed at their own pace, repeating exercises as often as necessary until mastery was achieved. The immediate feedback offered by these systems reinforced correct responses and highlighted errors, creating what proponents argued was a more efficient learning cycle.

Nevertheless, early CALL quickly revealed its limitations. Critics argued that drill-and-practice software failed to capture the communicative essence of language. While learners could memorize conjugations or word lists, they often struggled to transfer this knowledge into meaningful communicative contexts. The emphasis on form over meaning risked producing learners who were grammatically accurate in controlled tasks but unable to engage in authentic interaction. Furthermore, the rigid, text-based nature of early programs excluded the rich input, visual, auditory, and contextual, that SLA research has since shown to be critical for acquisition (Krashen, 1985; Long, 1996). By the late 1970s, dissatisfaction with behaviorist approaches in education more broadly, coupled with the rise of communicative language teaching (CLT), began to shift the trajectory of CALL toward more interaction-oriented models.

Still, the legacy of early CALL should not be dismissed. Many of its core ideas, individualized pacing, automatic feedback, and learner tracking, remain central to contemporary digital learning environments. Moreover, the drill-and-practice model aligns with certain aspects of SLA research, particularly where explicit learning of forms plays a role. Ellis (2006), for example, notes that explicit instruction can accelerate acquisition of features that are otherwise not salient in input. Similarly, Nation (2001) acknowledges the role of deliberate vocabulary learning, often achieved through repetition and recall, in building the lexical foundations necessary for comprehension and production. Thus, while limited in scope, early CALL foreshadowed enduring debates about the balance between form-focused practice and meaning-focused communication in language pedagogy.

In sum, the drill-and-practice era of CALL represents both a starting point and a cautionary tale. It demonstrated the potential of computers to extend language instruction beyond the classroom and highlighted the value of automation in practice and feedback. At the same time, its shortcomings made clear that technology must align with broader pedagogical goals if it is to truly enhance language acquisition. This realization set the stage for the communicative turn in CALL, where the emphasis shifted from repetitive form practice to fostering authentic interaction and communicative competence, a transition that will be examined in the next section.

2.3 The Move Toward Communicative CALL

By the late 1970s and early 1980s, the limitations of drill-and-practice CALL had become increasingly apparent. While early systems such as PLATO offered valuable opportunities for repetition and feedback, they failed to capture the communicative nature of language, leaving learners with decontextualized skills that often did not transfer to real-world interaction. Simultaneously, language pedagogy was undergoing a paradigm shift with the rise of Communicative Language Teaching (CLT), which emphasized meaning-making, authentic interaction, and the development of communicative competence (Canale & Swain, 1980; Littlewood, 1981). CALL evolved in parallel, moving away from behaviorist drills toward methods that reflected the goals of CLT, an era commonly referred to as Communicative CALL (Underwood, 1984; Warschauer, 1996).

Communicative CALL placed far greater emphasis on learners' ability to use language for meaningful purposes rather than manipulate forms in isolation. The computer became less a vehicle for mechanical practice and more a platform for exploration, interaction, and problem-solving. Early examples included text reconstruction programs, adventure-style games requiring linguistic input, and software that prompted learners to generate original sentences or narratives rather than select from pre-coded responses. The introduction of multimedia, particularly audio and video, expanded opportunities for authentic input, aligning closely with Krashen's Input Hypothesis (1982, 1985), which stressed the role of comprehensible input in naturalistic contexts.

Yet the shift was not limited to input. Communicative CALL also fostered opportunities for learner output, echoing Swain's Output Hypothesis (1985, 1995). By engaging learners in written or spoken production tasks, these systems encouraged them to notice gaps in their interlanguage and refine their linguistic competence through use. Collaborative CALL tasks, where learners worked together on problem-solving activities or joint text production, mirrored classroom interaction and embodied the negotiation of meaning processes described in Long's Interaction Hypothesis (1996).

One of the key contributions of this stage was the use of computers to simulate authentic contexts for communication. Programs incorporated dialogues, simulations, and role-playing exercises that required learners to make decisions and respond to prompts in ways that mimicked real-life interaction. This development aligned with Searle's (1969) theory of speech acts and subsequent pragmatics research (Blum-Kulka et al., 1989), which underscored that successful language use depends not only on grammatical accuracy but also on the ability to perform socially appropriate actions such as requesting, apologizing, or refusing. CALL activities in this era thus began to integrate pragmatic awareness into digital learning environments.

Despite these advances, Communicative CALL was not without challenges. Computers of the 1980s had limited processing power, restricting the complexity of interaction and feedback that software could provide. Much of the communication remained text-based, and while learners were encouraged to be creative, the computer's capacity to interpret and respond meaningfully was minimal compared to a human interlocutor. Still, the pedagogical orientation of CALL shifted decisively: technology was no longer conceived merely as a tutor delivering drills but as a tool to engage learners in exploration, communication, and interaction. This phase also marked the growing influence of task-based language teaching (TBLT), championed by scholars such as Ellis (2003, 2006). Task-oriented CALL activities paralleled classroom TBLT practices, with learners completing tasks that required the integration of multiple linguistic skills. For example, reconstructing a scrambled text or engaging in a problem-solving game required attention to syntax, lexis, and discourse coherence. Such tasks reflected the interdependence of linguistic domains highlighted in both SLA theory and CEFR descriptors. In short, the move toward Communicative CALL reflected broader changes in language pedagogy and SLA research. Communicative CALL acknowledged the limitations of repetitive drills and adopted a more holistic view of language learning, one that balanced input, output, and interaction while situating language use in authentic, meaningful contexts (Warschauer & Healey, 1998; Chapelle, 2001). At the same time, it highlighted the emerging potential of computers as platforms for learner autonomy and creativity, rather than as tools for mechanical instruction. This reorientation toward meaning and learner agency set the stage for the next major development in CALL, the rise of Integrative CALL in the 1990s and 2000s, when the growth of the internet, multimedia, and networked communication would further expand opportunities for collaboration, authentic input, and individualized learning (Warschauer, 1996, 2002; Egbert, 2005).

2.4 Integrative CALL in the Age of Digital Learning

The late 1990s and early 2000s ushered in a new phase of Computer-Assisted Language Learning, often referred to as Integrative CALL. This stage was shaped by two converging forces: the rapid

development of the internet and digital multimedia and the pedagogical momentum of integrative, task-based, and communicative approaches in second language teaching. Unlike earlier stages, which tended to isolate specific skills or emphasize either input or output, Integrative CALL sought to bring together the full range of linguistic competencies, listening, speaking, reading, writing, and pragmatic awareness, within a single, interconnected digital environment.

With the widespread availability of the internet, learners gained unprecedented access to authentic language resources. Online newspapers, videos, podcasts, and discussion forums provided input that aligned closely with Krashen's (1985) comprehensible input hypothesis, exposing learners to rich, naturalistic language in diverse contexts. Yet, exposure was no longer limited to receptive modes. Web-based communication tools such as email, chat, and later video conferencing allowed learners to interact with peers, teachers, and native speakers in real time, facilitating the negotiation of meaning central to Long's (1996) interaction hypothesis. In this way, integrative CALL environments blurred the boundary between classroom practice and real-world communication.

Another hallmark of this period was the incorporation of multimodality. Digital platforms could combine text, sound, and video, allowing learners to process input and produce output through multiple channels. For instance, learners might watch a subtitled video, discuss it in an online forum, and then write a reflective essay, all within the same platform. Such integrated tasks align with Ellis's (2003, 2006) task-based learning framework, where learners are required to mobilize different linguistic resources simultaneously to achieve communicative goals.

The growth of online corpora and concordancing tools further expanded the scope of CALL. Learners could now explore authentic language use at scale, identifying patterns in collocation, register, and discourse organization. This development directly supported Nation's (2001) emphasis on vocabulary acquisition, particularly his call for attention to both breadth and depth of lexical knowledge. Digital corpora allowed learners to see how words and structures function in real contexts, bridging the gap between isolated word lists and meaningful use.

Pragmatics, too, gained new relevance in integrative CALL. As Searle (1969) and subsequent pragmatics researchers such as Blum-Kulka, House, and Kasper (1989) emphasized, communicative competence requires not only grammatical accuracy but also appropriateness in social interaction. Online communication tools created authentic but lower-stakes contexts for learners to practice requests, apologies, and refusals in intercultural exchanges, offering opportunities to develop both pragmalinguistic and sociopragmatic competence (Taguchi, 2019).

Moreover, the rise of learning management systems (LMS) and virtual learning environments allowed institutions to embed CALL into the structure of formal education. Platforms such as Moodle and Blackboard provided centralized access to materials, assessments, and communication tools, enabling

a blended learning model in which digital and face-to-face instruction complemented each other. This integrative approach resonates with the CEFR framework (Council of Europe, 2001, 2020), which emphasizes the balanced development of receptive, productive, and interactive skills across proficiency levels. By aligning digital tasks to CEFR descriptors, instructors could design learning pathways that were both standardized and adaptable to individual needs.

Despite these advances, Integrative CALL also faced significant challenges. The sheer abundance of digital resources risked cognitive overload, as learners were often required to navigate vast, uncurationed materials without sufficient scaffolding or guidance (Hampel & Stickler, 2005). Instructors, too, struggled with the growing demand for technological literacy needed to implement CALL effectively (Egbert, 2005; Chapelle, 2001). Furthermore, inequities in access to hardware, software, and internet connectivity created disparities among learners, limiting the democratizing potential of digital language learning (Warschauer, 2002).

Nevertheless, the Integrative stage of CALL represented a major step forward in conceptualizing technology as a collaborative partner in language learning rather than a mere supplement. Integrative CALL expanded the pedagogical landscape by embedding digital resources, authentic communication, and task-based methodologies into cohesive online learning environments (Egbert, Chao, & Hanson-Smith, 1999). It marked a turning point in which technology became central to curriculum design and instructional practice. This integrative orientation laid the foundation for the next and most transformative stage: Intelligent CALL (ICALL), where advances in artificial intelligence (AI) and natural language processing (NLP) promise to personalize instruction dynamically according to the evolving needs of individual learners (Heift & Schulze, 2015; Schulze, 2008).

2.5 Intelligent CALL: From Static Programs to Adaptive Systems

While Integrative CALL marked a turning point in embedding technology within communicative and task-based pedagogy, the twenty-first century has ushered in an even more transformative phase: Intelligent Computer-Assisted Language Learning (ICALL). Unlike earlier CALL systems, which were largely static and limited to preprogrammed responses, ICALL draws on advances in artificial intelligence (AI), natural language processing (NLP), speech recognition, and adaptive learning algorithms to create dynamic, personalized, and interactive environments (Heift & Schulze, 2015; Schulze, 2008). These intelligent systems not only deliver input and record learner responses but also analyze learner output, diagnose errors, and adapt instructional sequences in real time, thereby operationalizing key theoretical insights from second language acquisition (SLA) research (Chapelle,

2009; Meurers, 2012). In doing so, ICALL represents the convergence of linguistic theory, cognitive psychology, and computational innovation, transforming computers from passive tools into active pedagogical partners in the language learning process.

At the core of ICALL is the capacity to provide individualized feedback. Traditional CALL could indicate right or wrong answers but rarely explain why an error occurred or how to correct it. ICALL systems, by contrast, can draw on linguistic databases and parsing technologies to identify error patterns, categorize them (phonological, morphological, syntactic, lexical, pragmatic), and provide feedback tailored to the learner's interlanguage stage. For example, a pronunciation-training tool informed by Ladefoged and Johnson's (2014) phonetic framework can visually display formant structures or voice-onset times, helping learners notice contrasts such as /r/ vs. /l/ that are filtered through their L1 phonological categories. Similarly, ICALL grammar parsers can highlight issues in agreement or word order, offering explanations aligned with the learner's proficiency level.

Another distinctive strength of ICALL is its ability to adapt instruction dynamically. Using learner analytics, ICALL platforms can monitor progress over time, identify strengths and weaknesses, and adjust the sequence and complexity of tasks accordingly. This approach resonates with Krashen's Input Hypothesis (1985), since adaptive systems can ensure that learners are consistently exposed to input at the 'i+1' level, slightly beyond their current competence. At the same time, ICALL supports Swain's Output Hypothesis (1995) by creating tasks that push learners to produce comprehensible output under communicative pressure, thus fostering noticing, hypothesis testing, and metalinguistic reflection.

ICALL also excels in supporting interaction, both human-computer and peer-to-peer. Advances in NLP and conversational AI allow for dialogue systems that simulate authentic interaction, providing learners with opportunities to practice pragmatic skills in controlled yet realistic contexts. This aligns with Long's Interaction Hypothesis (1996), which highlights the role of negotiation of meaning in acquisition. For example, an ICALL environment might simulate a business meeting or service encounter where the learner must make requests, clarify misunderstandings, or adjust register, tasks that require not just grammatical competence but also the pragmatic sensitivity described by Searle (1969) and further elaborated by Blum-Kulka, House, and Kasper (1989).

Vocabulary acquisition, a central component of SLA, is another area where ICALL demonstrates specific promise. By drawing on frequency data and lexical databases, ICALL systems can scaffold learners' vocabulary growth in line with Nation's (2001) model of breadth and depth. Adaptive algorithms can introduce new words at optimal intervals (spaced repetition), track collocational knowledge, and present contextualized examples that deepen semantic understanding. Unlike

traditional word lists, these systems can personalize vocabulary instruction, ensuring that learners encounter high-utility items relevant to their level, goals, and disciplinary interests.

Beyond linguistic domains, ICALL provides a platform for integrating assessment with instruction. Systems can align tasks with CEFR proficiency descriptors (Council of Europe, 2001, 2020), automatically mapping learner performance to standardized levels. This enables continuous formative assessment without interrupting the learning process, giving learners and teachers ongoing insight into progress while maintaining alignment with international benchmarks.

Despite these advances, ICALL is not without limitations. Speech recognition, though improving, still struggles with accented speech and spontaneous production, which can lead to frustration or inaccurate feedback. NLP systems, while increasingly sophisticated, cannot yet fully capture the nuances of pragmatics, irony, or culturally embedded language use. Moreover, questions remain about the pedagogical implications of relying heavily on automated systems: Will learners become overly dependent on machine feedback? How can teachers integrate ICALL tools effectively without reducing instruction to error correction? These challenges underscore the need for ICALL to be understood not as a replacement for human teaching but as a complement that amplifies its reach and precision.

In sum, Intelligent Computer-Assisted Language Learning (ICALL) represents a major advance in the evolution of language learning technologies. By integrating input-rich exposure, output-oriented tasks, and interactional simulations, ICALL aligns closely with the core principles of second language acquisition (SLA) theory while offering unprecedented personalization, adaptivity, and scalability (Chapelle, 2009; Heift & Schulze, 2015; Meurers, 2012). At its best, ICALL operationalizes the insights of Krashen's Input Hypothesis (1982, 1985), Swain's Output Hypothesis (1985, 1995), Long's Interaction Hypothesis (1996), Ellis's work on form-focused instruction (2008), Nation's research on vocabulary learning (2001, 2013), Ladefoged and Johnson's phonetic foundations (2014), and Searle's theory of speech acts (1969) within a unified technological ecosystem. It exemplifies the transformation of computers from static tutors to intelligent pedagogical partners, laying the groundwork for adaptive, responsive, and learner-centered models of language education.

2.6 Previous Works in the Field of CALL

Given the large number of tools available, it is essential to focus on those that have made significant academic contributions, particularly those supported by rigorous documentation and empirical research. While many commercial platforms exist, this section emphasizes tools whose development

and evaluation have been grounded in scholarship, as these provide clearer insights into the evolution of language learning technologies.

Several early systems laid the groundwork for contemporary CALL. (Kılıçkaya, 2007) pioneered hyperlinked dictionaries, enabling learners to access definitions and multimedia resources instantly, thereby fostering vocabulary learning in context (Godwin-Jones, 2008). Hot Potatoes revolutionized exercise creation, empowering teachers to build customized interactive materials that addressed grammar, vocabulary, and reading comprehension, democratizing CALL and enabling systematic research on exercise effectiveness (Bauer, 2005). *AntLap*, with its emphasis on vocabulary recycling, contributed to the understanding of repeated exposure in word acquisition and informed later designs for spaced repetition systems (Anthony, 2014). *Lextutor* provided learners with online concordances and access to corpus-based resources, encouraging the exploration of authentic language use and influencing both SLA research and pedagogical practices (Cobb, 2007). Similarly, the French Learner's Workbench showcased the promise of multimedia integration by combining text, audio, and visuals for a richer learning experience (Hayet, 1994). More recent examples include *SmartReader*, an adaptive e-reader that personalizes text complexity and multimedia scaffolding to support vocabulary acquisition (Zhao, Guo, & Dynia, 2017), and *WERTi*, which transformed authentic web pages into interactive grammar and vocabulary practice environments (Meurers et al., 2010; 2011). Finally, FLAIR demonstrated how linguistically enriched input can be retrieved and tailored to a learner's level, advancing adaptive reading support in CALL (Chinkina & Meurers, 2016).

Taken together, these systems illustrate the gradual evolution of CALL from relatively static, form-focused applications to more interactive, learner-centered platforms. While tools such as *WordChamp*, (Half-Baked Software & Centre, 2022), and *Lextutor* highlighted the value of interactivity, vocabulary exposure, and corpus-based exploration, they were often limited in their capacity to respond dynamically to individual learner profiles. The next generation of CALL systems, by contrast, integrates insights from SLA theory with advances in natural language processing and machine learning to create environments that adapt in real time to learner needs. This shift marks the transition from CALL to ICALL, where the focus is not only on providing access to resources but also on personalizing instruction, scaffolding learning processes, and generating data-driven feedback.

Against this backdrop, two recent systems, *SMILLE* and *Revita*, represent significant steps forward. Both illustrate how the integration of intelligent technologies with pedagogical principles can overcome the limitations of earlier CALL tools and align more closely with frameworks such as the

CEFR. In the following subsections, their functionalities, methodologies, and contributions will be examined in detail.

2.6.1 SMILLE

The *SMILLE* (Zilio et al., 2017) system is an advanced educational tool designed to support second language acquisition through personalized, adaptive techniques. It accommodates learners from CEFR levels A1 to C1, providing grammatical information, vocabulary support, and interactive feedback tailored to individual profiles. A central feature of *SMILLE* is its personalized learning database, which tracks learner progress and dynamically adjusts content difficulty in response to performance. This ensures that learners encounter material that is both level-appropriate and aligned with their developing needs.

The system integrates existing NLP resources to enrich raw online texts, enabling users to extract passages from web pages and annotate them with grammatical structures, word glosses (via WordNet: Miller et al., 1990), and dictionary definitions. By doing so, *SMILLE* encourages learners to engage in independent study while providing explicit guidance on linguistic features in context.

Evaluation of *SMILLE* has shown promising results, with over 85% precision in detecting grammatical structures, though recall, particularly for short answers, remains an area requiring improvement. Planned updates include refining error analysis, expanding the language base, and implementing databases that track learner interactions (e.g., *words looked up*, *rules consulted*). These enhancements aim to create a more scaffolded and autonomous learning environment, where learners receive both input support and real-time analysis of their output.

In short, *SMILLE* illustrates the potential of intelligent CALL to combine adaptive technology with explicit SLA principles. Its design emphasizes grammatical awareness, personalized scaffolding, and data-driven insights, though further refinement is needed to ensure accuracy and comprehensive coverage across languages.

2.6.2 Revita

These tools exemplify the synergy between Second Language Acquisition (SLA) and Computer-Assisted Language Learning (CALL), demonstrating how technological advancements can enhance language learning processes. By effectively integrating cutting-edge technology with established pedagogical principles, these tools contribute significantly to various facets of language acquisition, from vocabulary development to the creation of interactive exercises. This integration underscores

the importance of interdisciplinary collaboration in advancing the field of educational technologies, particularly in language learning (Katinskaia, 2018).

The field of Intelligent Computer-Assisted Language Learning (ICALL) and Intelligent Tutoring Systems (ITS) represents a dynamic and evolving area of research. A thorough review of the literature and current tools reveals both notable advancements and existing limitations within these systems. While ICALL systems have made strides in personalized language learning, they often lack the intelligent capabilities needed to fully adapt to learners' evolving needs. Similarly, ITS systems, despite their promise, often remain in experimental phases and fail to achieve wide-scale practical application. This review seeks to address these challenges, highlighting how newer tools, such as *Revita*, are overcoming these obstacles.

Revita, as described in Katinskaia's (2018), marks a significant step forward in language learning technology by successfully merging Intelligent Tutoring Systems (ITS) and Computer-Assisted Language Learning (CALL). By blending these two approaches, *Revita* overcomes many of the limitations seen in existing systems, providing a more dynamic and effective solution to language education.

Central to *Revita*'s innovation is its use of Adaptive Learning Technology, which tailors instructional content to the individual learner based on their performance. This personalized approach marks a clear departure from traditional rule-based ICALL systems, which often fail to adjust content based on specific learner needs. Additionally, *Revita*'s accessibility sets it apart from other ITS systems, as it is freely available online, allowing a broad range of learners and educators to engage with the learning platform. Unlike many ITS systems, which are often confined to experimental environments, *Revita* bridges the gap between theoretical advancements and real-world application, making it a valuable resource for practical language learning.

A key feature of *Revita* is its focus on vocabulary acquisition, incorporating interactive exercises designed to help learners build and reinforce their vocabulary. This emphasis aligns with current trends in language learning, where vocabulary development is recognized as central to overall language proficiency. Moreover, *Revita* excels in its educational data collection capabilities, gathering valuable insights into learner behavior and teaching effectiveness. This data, combined with ongoing feedback from experts in Second Language Acquisition (SLA) and Foreign Language Acquisition (FLA), enables continuous refinement of the system, enhancing its effectiveness over time.

Katinskaia's underscores how *Revita* exemplifies the integration of ITS and CALL, demonstrating its significant impact in addressing the limitations of existing systems while focusing on practical, accessible language learning solutions. Katinskaia's (Katinskaia, 2018) highlights how *Revita* exemplifies the successful integration of ITS and CALL, showcasing the potential of such

interdisciplinary approaches in advancing language learning technologies. By addressing the limitations of existing systems and focusing on practical application, *Revita* represents a significant step forward in creating more intelligent, accessible, and effective language-learning tools.

The learning platform's emphasis on adaptive learning, user accessibility, and data-driven improvements underscores its potential to make a meaningful impact on language education. *Revita* not only enhances the learning experience for individual users but also contributes to the broader field of educational technology by demonstrating the benefits of combining ITS and CALL approaches.

Revita systems have been largely integrated with a number of modules and implementation which go in the direction of more modern technologies in vogue and, also in the direction of machine learning. The study on assessing grammatical correctness in language learning (Katinskaia & Yangarber, 2021), particularly through the lens of evaluating alternative-correct answers in cloze exercises, highlights significant challenges and advancements in error detection. It underscores the limitations of traditional pre-trained models, like BERT, and explores the potential of fine-tuning these models on synthetic data to improve performance. This approach aims to enhance the accuracy of grammatical assessments and provide more reliable feedback for learners.

In parallel with these advancements, *Revita* has been at the forefront of integrating modern technologies and machine learning into language education. The learning platform has incorporated a range of modules and implementations that reflect current trends in educational technology. *Revita's* emphasis on adaptive learning and real-time data utilization aligns with the latest developments in the field, showcasing a commitment to leveraging cutting-edge technologies to improve language learning outcomes.

Revita's integration of ITS and CALL with modern machine learning techniques represents a natural progression from traditional systems to more sophisticated, data-driven approaches. While the study focuses on refining grammatical correctness through specialized model training, *Revita* applies similar principles on a broader scale, utilizing machine learning to personalize learning experiences and enhance feedback mechanisms.

The alignment between the study's exploration of advanced error detection methods and *Revita's* implementation of modern technologies highlights a shared goal: advancing language learning through innovative solutions. Both endeavors contribute to the development of more effective, adaptive, and intelligent language learning tools, bridging the gap between theoretical advancements and practical applications in the field of language education.

Revita addresses two central challenges in language learning technology: adaptability and personalization, and automated grammatical assessment with feedback. The learning platform's adaptive architecture ensures that learners receive content calibrated to their performance history and

learning profiles, creating a dynamic environment that responds to individual progress (Katinskaia, Yangarber, & Lazareva, 2018). Personalization extends this adaptability by continuously adjusting task difficulty, exercise type, and linguistic focus according to each learner's evolving proficiency. Equally crucial is *Revita*'s capacity for grammatical correctness assessment and reliable feedback delivery, which are foundational to effective second language development. Through natural language processing (NLP) and machine learning techniques, the system can identify and classify learner errors, provide targeted feedback, and guide users toward self-correction and linguistic refinement (Yangarber et al., 2022). Such intelligent feedback mechanisms not only support accuracy and fluency gains but also foster deeper engagement with the learning process, enabling learners to consolidate strengths and address weaknesses more effectively (Heift & Schulze, 2015; Meurers, 2012).

One of the significant challenges highlighted by the *Revita* project is word-sense disambiguation (Quecedo et al., 2020). Disambiguating word meanings, particularly in languages with complex morphological structures, is a critical issue for many language learning systems. The *Revita* project has shown promising results in disambiguating morphological analyses for languages like Finnish, Russian, and Spanish, demonstrating the effectiveness of its approach in handling different linguistic structures. The process hinges on having a high-quality morphological analyzer, a large text corpus, and ideally annotated data for evaluation. While the performance of disambiguation showed only slight differences between Part-of-Speech (POS) and lemma cases, the findings suggest that further improvements are possible. A promising direction is the combination of POS and lemma predictions into an ensemble model, which could potentially boost accuracy from 85% to 94%.

Additionally, the exploration of multi-task learning in the *Revita* project has revealed both challenges and opportunities for enhancement. Integrating a model for predicting morphological tags could significantly improve disambiguation accuracy. This approach holds great potential for refining language learning tools, particularly in languages with complex grammatical systems. Through these advancements, the *Revita* system not only contributes to improving language learning technology but also sets the stage for future innovations in error detection, grammatical assessment, and multilingual word sense disambiguation.

A key strength of this project lies in its rigorous approach to disambiguating morphological analyses across a range of languages, particularly those characterized by complex morphological structures. The project demonstrates considerable potential for achieving high accuracy in morphological analysis, with the possibility of extending its applicability to additional, morphologically rich languages. The development of ensemble models and the continued exploration of multi-task learning underscore the project's commitment to advancing morphological analysis techniques, with the aim of attaining near-optimal performance. Such advancements are critical for enhancing the effectiveness

of language learning technologies, particularly in languages with intricate grammatical and morphological systems.

The study presented in Katinskaia & Yangarber (2021) focuses on the evaluation of grammatical correctness in alternative-correct answers provided by advanced learners, specifically in the context of the Russian version of *Revita*. A central challenge within this domain is the issue of alternative-correct answers, where responses, while grammatically valid, deviate from expected formulations. This variability introduces significant complexity in the accurate assessment of grammatical correctness. Initial tests using pre-trained BERT models revealed limitations, particularly in the evaluation of Russian grammatical structures as compared to English. However, fine-tuning BERT on synthetic data resulted in notable improvements, indicating that tailored approaches may substantially enhance model performance. The study outlines key future directions, including the refinement of synthetic data generation processes and the enhancement of model confidence estimation. Both of these factors are crucial for ensuring the reliability and precision of feedback provided to learners. Moreover, the study proposes the specification of answer positions, a strategy aimed at improving the contextual relevance of feedback.

A broader understanding of current developments in language learning can be achieved by considering contextualized approaches. Traditional error detection models, which rely predominantly on rule-based systems, often encounter challenges when addressing the complexities of advanced grammatical correctness. In contrast, the shift towards data-driven methods, such as the use of BERT and synthetic data, represents a more sophisticated and flexible approach to error detection. Additionally, the integration of adaptive learning systems, which dynamically adjust content based on individual learner performance, presents a complementary approach to the evaluation of grammatical correctness. These systems provide more context-sensitive feedback, adjusting in real-time to the learner's evolving proficiency.

The incorporation of real-time data into contemporary language learning platforms further refines error detection. Unlike pre-collected and synthetic data, real-time data allows for the analysis of ongoing learner interactions, which facilitates a more immediate and nuanced form of feedback. Furthermore, the study emphasizes the need for comprehensive error analysis in language learning, which should extend beyond grammatical errors to include lexical and pragmatic errors. By adopting a broader framework for error detection, it is possible to significantly enhance the accuracy and effectiveness of error correction systems.

The study places specific emphasis on model confidence and feedback reliability, key factors in ensuring the consistency and dependability of automated feedback mechanisms. Concurrently, other approaches in the field are also investigating methods to assess and improve the reliability of

feedback, thereby ensuring that learners receive continuous, accurate, and trustworthy guidance throughout their language acquisition process.

Another strong point of the system proposed is its comprehensive approach to incorporating activities and gamification features within the *Revita* Intelligent Tutoring System (Hou et al., 2022). The system's design aims to enhance learning effectiveness by actively engaging users through a diverse range of interactive elements. Our research involves tracking 600 learners across multiple universities to evaluate how these features impact learner engagement and overall effectiveness. *Revita's* approach is rooted in the belief that a varied selection of activities and gamification can sustain learner interest and improve educational outcomes, compared to a narrower focus on exercise types. Achieving robust quantitative evidence of these benefits requires longitudinal studies involving a large and diverse user base, supported by systems that are sufficiently user-friendly to sustain long-term engagement (Chapelle, 2001; Egbert, 2005). Although *Revita's* gamification framework remains in its early stages and is not yet grounded in a specific theoretical model, the growing body of user interaction data offers valuable opportunities for empirical validation and for informing the development of new pedagogical and motivational frameworks (Deterding, Dixon, Khaled, & Nacke, 2011). Future work aims to expand *Revita's* gamified features, such as adaptive challenges, progress visualization, and reward systems, and to conduct detailed analyses of learner engagement metrics, with the goal of better understanding how motivation, persistence, and learning outcomes interact within AI-driven language learning environments (Reinhardt & Sykes, 2014; Reiners & Wood, 2015).

SMILLE and *Revita's* approaches illustrate the significant impact of technological advancements, such as Natural Language Processing and machine learning, on enhancing Computer-Assisted Language Learning (CALL) and second language acquisition (SLA) tools. These systems provide personalized, context-aware support that adapts to individual learning needs, improving the effectiveness of language learning. However, aligning these technologies with the CEFR framework poses challenges. The complexity of covering all CEFR proficiency levels, from basic communication at A1 to nuanced language use at C2, requires advanced systems to handle subtle linguistic nuances and complex language tasks. Additionally, providing effective scaffolding and ensuring comprehensive support across all language skills, listening, reading, speaking, and writing, are crucial for achieving the depth and breadth of proficiency outlined by CEFR. Addressing these challenges is essential for creating tools that fully support learners at every stage of their language development.

One of the main complexities in developing effective CALL tools is anticipating all potential flaws and implicit requests that learners might have during their interaction with the system. Unlike traditional classroom settings where immediate feedback and personalized support are possible, asynchronous CALL tools face limitations due to the distance between the learner and any human

assistance. Consequently, the tool must be meticulously designed to address and accommodate a wide range of learner needs and queries without real-time intervention. This requires the system to preemptively account for various learning scenarios and incorporate robust mechanisms for troubleshooting and guidance, ensuring that learners receive the support they need even when direct interaction with a teacher is not feasible.

2.6.3 Other Systems in CALL/ICALL

Beyond *SMILLE* and *Revita*, several other systems have significantly influenced the development of CALL and ICALL, illustrating both the commercial reach of language learning technologies and their research-driven evolution.

On the commercial side, Duolingo (von Ahn, 2013) has become one of the most widely used platforms worldwide, popularizing gamification as a pedagogical strategy. By combining short adaptive lessons with a reward-based structure, it has successfully engaged millions of learners, though questions remain about its effectiveness in fostering advanced proficiency (Loewen et al., 2019). Similarly, *Rosetta Stone*, one of the earliest large-scale CALL applications, adopted an immersion model that relied heavily on contextualized input supported by visuals. While criticized for its limited focus on grammar and overemphasis on repetition, it was instrumental in making CALL accessible to a global market. Other commercial platforms such as Babbel and *Busuu* extend this tradition, offering CEFR-aligned curricula and opportunities for peer-to-peer interaction, echoing Long's (1996) interaction hypothesis in their design.

In academic and research contexts, more specialized systems have advanced the integration of intelligent technologies. *IntelliText* (Meurers, 2015) exemplifies the use of natural language processing to create interactive reading environments, supporting comprehension through adaptive input enhancement. *TELL Me More* (Auralog) pioneered the integration of speech recognition into CALL, enabling learners to practice pronunciation based on phonetic feedback aligned with the principles of Ladefoged and Johnson (2014). Similarly, experimental ICALL systems such as ELLIPSIS and ALICE applied the principles of Intelligent Tutoring Systems (ITS) to grammar and vocabulary learning, laying the groundwork for modern intelligent feedback mechanisms.

Taken together, these systems demonstrate the breadth of CALL and ICALL innovation. While commercial applications have expanded access to language learning globally, research-driven projects have deepened the field's understanding of how speech recognition, NLP, and adaptive algorithms can support second language acquisition. Both strands continue to inform each other, shaping a landscape where learner engagement, pedagogical alignment, and technological sophistication are increasingly interwoven.

2.6.4 Comparative Discussion

The overview presented in the previous sections reveals a clear correlation between technological advancements and their application in the field of second language acquisition. A continuum can be observed, beginning with the advent of computer technology and interface-based applications, which facilitated early forms of computer-assisted language learning (CALL). These initial technologies enabled basic input/output interactions and laid the foundation for more complex forms of digital learning. Over time, developments in string manipulation techniques allowed for more refined text processing, enhancing the ability to analyze and generate language patterns. This was followed by advances in natural language processing (NLP), which introduced the capacity to handle large datasets of linguistic input, offering learners more authentic, data-driven exposure to language use. NLP-powered systems enabled more personalized and context-aware learning environments, leading to adaptive feedback and intelligent tutoring systems. As these technologies evolved, they set the stage for the integration of artificial intelligence (AI), where machine learning algorithms can now model learner behavior, predict learning trajectories, and generate adaptive content in real time. These AI-driven applications represent the cutting edge of second language acquisition technologies, providing increasingly dynamic, interactive, and individualized learning experiences.

However, despite the significant advancements in technology for second language acquisition, many systems face challenges related to availability and maintenance. A common issue is the discontinuation or lack of regular updates to language learning platforms, which can render them obsolete or less effective over time. In many cases, the initial enthusiasm surrounding new technological tools is not matched by long-term support, leading to diminished user engagement and eventual abandonment. This is often due to the resource-intensive nature of maintaining such systems, as ongoing updates, technical support, and integration of new features require substantial investments. Without consistent maintenance, these systems may fail to keep pace with evolving technologies or learner needs, thereby losing their relevance.

Furthermore, the unavailability of certain tools, either due to licensing restrictions, proprietary limitations, or geographic access barriers, exacerbates the gap between potential technological innovations and their widespread application in educational contexts. This highlights the importance of sustainable development and support for second language acquisition technologies to ensure their continued effectiveness and accessibility.

In the intersection of didactics and technology within second language acquisition, the success of such systems is not solely dependent on their instructional design, but also on the continual technological updates and effective data management practices. As these platforms increasingly rely

on vast amounts of learner data to deliver personalized learning experiences, the ability to manage, store, and analyze this data becomes crucial. Efficient data management ensures that learner progress can be tracked accurately and that adaptive algorithms can make informed adjustments to instructional content. Moreover, the rapid pace of technological development demands that these systems stay updated to incorporate new features, improve user interfaces, and remain compatible with other evolving tools and platforms. Without regular technological updates, language learning systems risk becoming outdated, limiting their effectiveness and appeal to users. Therefore, the dynamic interplay between pedagogical strategies, technological advancements, and robust data management is essential to the long-term success and relevance of these educational technologies. Efforts to integrate emerging technologies into second language acquisition have been evident over the years, with continuous attempts to harness the best features of each new advancement to enhance the learning process and support acquisition. The academic community has extensively documented these efforts, as seen in numerous articles that highlight the intention behind technological innovations, the projects undertaken, and the processes of creation. These studies collectively underscore the academic validity and transformative potential of intelligent language learning tools. However, what clearly emerges is a persistent gap between theoretical promises and practical implementation. While conceptual frameworks and pilot projects in CALL and ICALL often attract substantial research interest, the resulting systems are frequently unavailable, discontinued, or fail to achieve broad adoption (Chapelle, 2001; Colpaert, 2010; Heift & Schulze, 2015). This discontinuity can be attributed to factors such as limited funding, technical complexity, and the absence of sustainable maintenance and scalability plans, which hinder the transition from research prototype to enduring educational product (Hubbard, 2009; Levy & Stockwell, 2006). Consequently, the field is left with a wealth of promising theoretical advances but relatively few robust, long-term tools that can be fully integrated into institutional and individual learning contexts. Often, technological products appear to be the culmination of extensive theoretical development, aimed at demonstrating their practical applicability and effectiveness. However, there seems to be a notable lack of interest in systematically evaluating the impact of these products on the learning process once they are implemented. While the initial phases of development focus on theoretical frameworks and conceptual validity, subsequent attention to the practical outcomes and effectiveness in real educational settings tends to diminish. This gap raises concerns about the long-term sustainability and relevance of such innovations in the field of second language acquisition, as the emphasis on empirical testing and user feedback is crucial for refining and validating these technologies.

Conversely, in the commercial sphere, applications designed for language learning often receive robust implementation and ongoing maintenance, particularly when they meet specific educational objectives and achieve economic targets. Companies in this sector are motivated by market demand and the necessity to remain competitive, leading to continuous updates, user engagement strategies, and enhancements based on user feedback. Successful commercial language learning platforms often prioritize user experience and measurable outcomes, ensuring that their products are not only functional but also appealing to a broad audience. This focus on maintaining relevance and effectiveness results in a dynamic interplay between technological innovation and user needs, allowing these applications to thrive and evolve in ways that many academic projects may not. As a result, commercial language learning tools often provide learners with reliable resources that are regularly updated and aligned with contemporary pedagogical practices.

However, even within the commercial sector, certain limitations arise due to the substantial amount of work required to develop and implement these applications effectively. The complexity of creating a user-friendly interface, integrating advanced technologies such as artificial intelligence and natural language processing, and ensuring ongoing technical support demands significant resources, both in terms of time and financial investment. Pedagogically, the need for continuous content updates and pedagogical adjustments to reflect current best practices can place considerable strain on development teams. While commercial entities are often better equipped to handle these challenges compared to academic projects, the pressures of rapid technological advancement and shifting market demands can lead to compromises in quality or functionality. This highlights the need for a balanced approach that not only focuses on the initial development of language learning applications but also prioritizes sustainable practices that support long-term success and effectiveness.

Returning to the realm of academic applications, it becomes evident that factors such as data availability, platform management, and the costs involved can significantly hinder the realization of innovative intentions. Limited access to high-quality data can restrict the development of adaptive learning algorithms and impede the effectiveness of personalized learning experiences. Furthermore, managing complex educational platforms often requires specialized knowledge and resources that may not be readily available within academic settings, leading to operational inefficiencies. Additionally, the financial constraints commonly faced by educational institutions can limit their ability to invest in state-of-the-art technologies and ongoing maintenance, further stalling progress. These challenges underscore the need for a strategic approach that considers not only the theoretical aspects of technological integration but also the practical realities of resource allocation, data management, and platform sustainability in order to fully harness the potential of technological advancements in language learning.

It is commendable to recognize the extensive efforts made over the past two decades in advancing computer-assisted language learning. Each initiative, whether successful or not, has contributed invaluable insights that have shaped the current technological landscape. The cumulative knowledge gained from these experiences has been instrumental in reaching the present state of the art in language learning technologies. By critically analyzing the successes and shortcomings of previous projects, we have endeavored to extract the best practices while remaining mindful of potential pitfalls. This reflective approach has guided the development of our product, ensuring that we incorporate proven methodologies while addressing the challenges faced by earlier systems. In doing so, we aim to create a more effective and sustainable learning tool that enriches the lessons learned from the rich history of CALL innovations.

Modern computer-assisted language learning (CALL) systems integrate natural language processing (NLP) to provide automated feedback, facilitating language acquisition by identifying errors and suggesting improvements (Chapelle, 2001). Advances in speech recognition technology have further enhanced pronunciation training, enabling the analysis of learner speech and the delivery of targeted feedback (Dybkjær et al., 2005). The ability to process multilingual data allows comparative linguistic analysis, aiding researchers in understanding cross-linguistic variations in language acquisition (De Bot et al., 2005). Computational models simulate language transfer effects, offering insights into how a learner's first language (L1) influences the acquisition of a second language (L2) (Odlin, 1989). Additionally, dynamic language models, continuously refined with new data, contribute to the development of increasingly accurate language learning tools (Laufer & Goldstein, 2004). Longitudinal studies, supported by computational tools, provide valuable perspectives.

The review of *SMILLE*, *Revita*, and other CALL/ICALL systems highlights both the diversity of approaches in the field and the convergence toward certain shared goals. At a broad level, these systems can be distinguished along two axes: commercial versus research-driven development, and rule-based versus adaptive, data-driven design.

Commercial platforms such as Duolingo, Rosetta Stone, Babbel, and Busuu¹ have been instrumental in popularizing language learning technologies and making them widely accessible. Their strengths lie in engaging user interfaces, mobile availability, and gamification strategies that sustain motivation.

¹ Commercial platforms such as Duolingo, Rosetta Stone, Babbel, and Busuu have been instrumental in popularizing language learning technologies and making them widely accessible.

Duolingo — <https://www.duolingo.com>

Rosetta Stone — <https://www.rosettastone.com>

Babbel — <https://www.babbel.com>

Busuu — <https://www.busuu.com>

However, their scientific validation has often been limited, and they tend to prioritize general usability over nuanced alignment with SLA theory. By contrast, research-driven systems like *SMILLE*, *Revita*, *IntelliText*, *WERTi*, and *FLAIR* were designed with explicit reference to linguistic theory and pedagogy, often focusing on specific domains such as vocabulary recycling, input enhancement, or pronunciation training. These tools, while less widely disseminated, contribute valuable insights into how SLA principles can be operationalized in technology.

Another point of comparison is the evolution from rule-based CALL to adaptive ICALL. Early tools such as Hot Potatoes or Rosetta Stone provided structured but static exercises, offering little variation in response to individual learner needs. Systems like *SMILLE* and *Revita* represent a shift toward dynamic personalization, where learners receive adaptive content and context-sensitive feedback. *Revita* also exemplifies the integration of machine learning techniques, from error detection with fine-tuned BERT models to morphological disambiguation and gamification, showing how advances in artificial intelligence can extend CALL beyond fixed rulesets.

In terms of focus areas, different systems privilege different aspects of SLA. *SMILLE* emphasizes grammatical awareness and explicit instruction, aligning with Ellis's (2006) arguments for form-focused learning. *Revita*, while also attentive to grammar, prioritizes vocabulary acquisition, feedback reliability, and learner engagement through gamified design, echoing Nation's (2001) emphasis on lexis. *IntelliText* and *WERTi* focus on input enhancement and reading comprehension, consistent with Krashen's (1985) input hypothesis, while *TELL Me More* and related tools foreground phonetics and pronunciation, drawing from the descriptive traditions of Ladefoged and Johnson (2014). Commercial platforms, meanwhile, often attempt to cover multiple skill areas but may sacrifice depth in favor of breadth.

Taken together, these comparisons reveal a field that is increasingly moving toward integration. The most promising systems, such as *Revita*, combine elements of SLA theory, data-driven adaptivity, and engaging design, while also addressing issues of accessibility and scalability. This integrative trend suggests that the future of CALL and ICALL lies not in privileging one linguistic domain or one theoretical framework, but in creating holistic environments where input, output, interaction, and feedback are all scaffolded in ways consistent with CEFR descriptors and empirical SLA research.

2.7 Rethinking CALL and ICALL: Challenges and Critiques

While the development of Computer-Assisted Language Learning (CALL) and its intelligent successor (ICALL) has introduced unprecedented opportunities for innovation in second language

education, both fields have also faced sustained critique. These challenges are not merely technical but also pedagogical, theoretical, and ethical, reminding us that technological solutions must remain grounded in the complex realities of language learning.

One of the most persistent issues concerns the limitations of technology itself. Despite advances in speech recognition and natural language processing, ICALL systems still struggle with spontaneous, accented, or non-standard language. Learners whose speech deviates from the training models used by these systems often receive inaccurate or inconsistent feedback. Similarly, NLP tools may handle well-formed sentences but fail to interpret pragmatics, irony, or indirectness, areas that are central to communicative competence (Searle, 1969; Blum-Kulka, House, & Kasper, 1989). These technological gaps raise questions about whether current ICALL platforms can fully support higher-order aspects of language use, especially at advanced CEFR levels (C1–C2).

A second critique focuses on pedagogical alignment. CALL and ICALL, at times, risk overemphasizing form-focused instruction, particularly when error detection and corrective feedback become the primary functions of the system. While form-focused attention has its place (Ellis, 2006), an imbalance may undermine communicative goals and reduce learning to a cycle of correction rather than authentic interaction. Likewise, heavy reliance on adaptive algorithms can narrow the learner's path to those items the system prioritizes, potentially restricting exposure to the rich, varied input that Krashen (1985) argued is essential for acquisition.

A third concern lies in the danger of fossilization. Without carefully designed tasks and feedback loops, CALL systems may allow learners to practice errors repeatedly, reinforcing incorrect patterns. Swain's (1995) emphasis on pushed output and metalinguistic reflection highlights the importance of carefully structured opportunities for production; if ICALL feedback is superficial or too tolerant of approximations, learners may plateau in their development. Similarly, Long's (1996) interaction hypothesis suggests that true progress comes from negotiation of meaning, something that automated systems can only approximate.

Beyond pedagogical concerns, there are issues of equity and accessibility. CALL and ICALL require technological infrastructure, reliable internet access, and in many cases, expensive proprietary software. Learners from under-resourced contexts risk being excluded from these innovations, exacerbating educational inequalities. Even when access is available, differences in digital literacy between teachers and learners can limit effective use. The promise of personalization, therefore, risks being offset by disparities in opportunity.

Finally, ethical considerations are increasingly salient. ICALL systems collect and analyze large amounts of learner data to personalize instruction. While this enables fine-tuned feedback, it also raises questions of privacy, consent, and data ownership. Moreover, the use of AI in educational

contexts prompts debates about transparency: to what extent should learners understand how adaptive decisions are being made on their behalf?

These critiques do not undermine the potential of CALL and ICALL but rather serve as reminders of the need for balance. Technology should not be seen as a replacement for teachers but as a complement that enhances instruction when integrated thoughtfully. ICALL is most effective when grounded in SLA theory, combined with communicative pedagogy, and situated within equitable and ethical learning frameworks.

In short, CALL and ICALL offer remarkable tools, but they are not panaceas. Their success depends on how well they address challenges of technological accuracy, pedagogical balance, accessibility, and ethical responsibility. Recognizing these limitations prepares the way for a more nuanced discussion of the opportunities and future directions in the field, which will be the focus of the next section.

2.7.1 Identified Gaps in Current CALL/ICALL

Despite the substantial progress made in Computer-Assisted Language Learning (CALL) and its intelligent extension, Intelligent CALL (ICALL), significant gaps remain that hinder their full pedagogical potential. These shortcomings become particularly apparent when evaluated against the theoretical benchmarks outlined earlier. Many existing systems succeed in offering interactivity, multimedia enrichment, or a degree of adaptivity, yet they often fall short in addressing the deeper cognitive and linguistic principles that underpin second language acquisition.

A recurring limitation of traditional CALL tools is their reliance on static content and fixed instructional pathways, which prevent them from responding dynamically to learner progress. While platforms such as Hot Potatoes or WordChamp introduced valuable innovations in exercise design and vocabulary support, they typically provide the same material regardless of learner proficiency or development. This static model contrasts sharply with Krashen's concept of input at the $i+1$ level and Nation's emphasis on lexical thresholds. Without adaptive mechanisms that regulate input difficulty in alignment with learner growth, many users either plateau due to unchallenging content or disengage when materials become overly complex.

Even among more advanced ICALL systems, such as *SMILLE* and *Revita*, a central challenge persists in balancing personalization with linguistic transparency. These systems make impressive use of natural language processing and adaptive algorithms, yet their internal decision-making processes often remain opaque, leaving learners and instructors uncertain about the rationale behind specific text selections or task assignments. While personalization can enhance motivation and engagement, it risks becoming pedagogically superficial when not explicitly grounded in SLA theory. Adaptivity

based solely on general performance metrics may improve user satisfaction but does not necessarily guarantee alignment with key linguistic targets such as collocational development, phraseological competence, or pragmatic appropriateness.

Another persistent gap lies in the underrepresentation of vocabulary and phraseology as organizing principles. Although vocabulary is widely acknowledged as central to communicative competence, many CALL systems continue to treat it as ancillary to grammar-focused instruction. Tools that emphasize vocabulary often restrict their focus to single words or frequency lists, neglecting Nation's insights into multiword units and collocations or the phraseological depth required for fluent, natural expression. Consequently, learners may accumulate lexical items without acquiring the combinatorial knowledge necessary for authentic language use.

Cognitive considerations also tend to be underdeveloped. Few systems explicitly address cognitive load management, despite strong evidence from cognitive psychology and neuroscience that optimal learning occurs when new material is introduced incrementally and within manageable cognitive limits. Systems that inundate learners with glosses, annotations, or unfamiliar vocabulary risk overloading working memory, thereby reducing learning efficiency. Likewise, limited attention is often given to the dual modes of implicit and explicit processing described in Dual Process Theory. Many platforms favor either incidental exposure or conscious analysis, rarely achieving the balance that SLA research identifies as crucial for long-term retention and fluency.

A further shortcoming lies in the limited integration with established proficiency frameworks such as the Common European Framework of Reference for Languages (CEFR). Although the CEFR has become the global standard for defining language competence, many CALL and ICALL systems do not systematically align their progression models or assessment criteria with their levels. This lack of alignment diminishes both the comparability of learning outcomes and the institutional applicability of these tools within CEFR-based curricula. As a result, learners often find it difficult to situate their progress within a coherent developmental trajectory.

Taken together, these issues reveal a disconnect between technological innovation and the theoretical principles of SLA. While CALL and ICALL have succeeded in expanding access, interactivity, and engagement, they frequently overlook the theoretical foundations that ensure sustained acquisition. The outcome is a proliferation of tools that are visually appealing and motivating in the short term but pedagogically inconsistent, capable of fostering surface-level learning without guiding learners through the full continuum of language proficiency.

Emerging approaches in adaptive language learning seek to address these shortcomings by centering lexical organization, embedding dynamic adaptivity grounded in Krashen's $i+1$ and Nation's lexical thresholds, incorporating productive tasks inspired by Swain's Output Hypothesis, simulating

interactional feedback in line with Long's Interaction Hypothesis, and aligning progression with the CEFR. Such systems aim to move beyond superficial personalization toward a principled integration of SLA theory and technological innovation, offering a more transparent, cognitively balanced, and pedagogically coherent framework for second language learning in the digital age.

2.8 Implementing Digital Language Tools: Pedagogical and Technological Hurdles

Drawing on insights from Second Language Acquisition (SLA), behavioral sciences, mental lexicon research, and computer-assisted education, contemporary second language learning tools increasingly place personalization at the core of the learning experience. Traditional CALL (Computer-Assisted Language Learning) environments often rely on static interventions, overlooking the dynamic and evolving nature of learner progress. In contrast, next-generation adaptive systems are designed to respond continuously to the learner's evolving lexical profile, thereby enabling a more nuanced and meaningful evaluation of lexical competence. This ambition, however, highlights a range of challenges spanning both didactic and technological dimensions.

From a didactic perspective, effective digital tools must align closely with established theories of language acquisition. Krashen's (2002) Input Hypothesis underscores the importance of comprehensible input at the 'i+1' level, while Swain's Output Hypothesis and Long's Interaction Hypothesis emphasize that input alone is insufficient without opportunities for production and negotiation of meaning. Complementing these models, Dual Process Theory (DPT) (Birdsong, 1994; Kahneman, 2011) highlights the interplay between implicit, automatic processing (System 1) and explicit, analytical learning (System 2), both of which must be supported in digital learning environments. Neuroscientific findings, such as the N400 effect (Rugg, 1985), further suggest that repetition and contextualized exposure reduce cognitive load, facilitating lexical access and fluency. Designing systems that balance these processing modes while maintaining learner motivation and engagement remains one of the central pedagogical challenges.

Another key didactic concern involves cognitive load management. Learners must encounter novel material without being overwhelmed. Following Nation's (2014) principle of keeping unfamiliar vocabulary below 2% of the text, adaptive systems aim to optimize input difficulty through dynamic adjustment. Cognitive Load Theory thus provides a critical framework for structuring learning sequences that are challenging yet achievable, avoiding both overload and underload while sustaining motivation and progress.

From a technological perspective, CALL systems face equally demanding challenges. Integration with existing Learning Management Systems (LMS) requires compatibility, reliability, and scalability, while data privacy and security demand adherence to regulatory frameworks and robust safeguards to protect user information. Scalability remains crucial: systems must preserve performance and responsiveness even as user numbers expand. Furthermore, interface design must prioritize accessibility and inclusivity, ensuring that learners from diverse backgrounds and abilities, including those with disabilities, can participate fully. Ongoing technical maintenance and support are also essential to address emerging issues, update linguistic resources, and sustain long-term user engagement.

The successes and shortcomings of earlier CALL tools underscore the importance of addressing these challenges holistically. Systems such as *SMILLE* and *Revita* illustrate how personalization and adaptive learning can be advanced through the innovative use of NLP and machine learning. Conversely, less successful implementations reveal common pitfalls: insufficient attention to engagement, lack of theoretical grounding, or technical instability that undermines pedagogical intent.

In the design of advanced language learning systems, these didactic and technological challenges are best treated not as obstacles but as design imperatives. Effective solutions integrate SLA theory with robust linguistic analysis, balance personalization with transparency in AI processes, and sustain engagement through careful management of cognitive load. The convergence of theoretical, pedagogical, and technological considerations ultimately defines the quality and sustainability of such adaptive learning environments.

2.8.1 Theoretical underpinnings

The design of adaptive second language learning systems is grounded in a synthesis of key theories from Second Language Acquisition (SLA) and cognitive science, forming a coherent framework for understanding how learners acquire and consolidate linguistic knowledge. Rather than treating technology as an end in itself, these systems draw from established theoretical traditions to remain pedagogically rigorous, cognitively informed, and linguistically transparent.

At the foundation lies Krashen's Input Hypothesis, which holds that language acquisition is most effective when learners encounter input that is comprehensible yet slightly beyond their current level of competence ($i+1$). This notion of incremental challenge is operationalized in adaptive text selection and sequencing, ensuring that input difficulty evolves in step with a learner's growing proficiency. By maintaining the proportion of unfamiliar words within Nation's (2014) recommended 2%

threshold, such systems aim to provide input that stretches the learner without overwhelming them, thereby fostering sustained engagement and progress.

Yet input alone is insufficient for the development of productive competence. Swain's Output Hypothesis demonstrated that the act of producing language compels learners to notice gaps in their knowledge, test hypotheses about form, and reflect consciously on their interlanguage. Adaptive CALL environments integrate this insight by embedding productive tasks, such as gap-fills, paraphrasing, and collocational completions, within their design. These tasks strengthen lexical recall and encourage learners to refine their grammatical and phraseological awareness through active production.

Long's Interaction Hypothesis adds an interactional dimension, emphasizing that negotiation of meaning and corrective feedback play critical roles in language acquisition. While digital systems cannot replicate human conversation entirely, adaptive feedback loops can simulate these processes by providing contextualized cues, glosses, and prompts that guide learners toward clarification and self-correction in real time.

In parallel, Rod Ellis's focus on form within meaningful communication stresses that attention to linguistic detail accelerates learning when it emerges from authentic communicative contexts. CALL systems apply this principle by embedding linguistic annotations, such as part-of-speech tags, morphological cues, and phraseological structures, directly into texts. This approach scaffolds learner awareness, ensuring that attention to form develops organically within comprehension and production activities.

Paul Nation's lexical perspective provides another essential pillar. His research established vocabulary knowledge, both breadth and depth, as the cornerstone of communicative competence, emphasizing mastery of high-frequency words and collocations. Adaptive systems therefore monitor lexical diversity, frequency bands, and collocational strength to ensure steady vocabulary expansion alongside acquisition of multiword units that underpin fluency and naturalistic expression.

The Common European Framework of Reference for Languages (CEFR) offers a unifying structure within which these theoretical elements converge. By aligning adaptive progression mechanisms with CEFR levels (A1–C2), digital learning tools ensure that development is internationally recognizable, pedagogically coherent, and easily integrated into institutional curricula. The CEFR's emphasis on communicative competence across modalities, listening, speaking, reading, and writing, resonates strongly with the goal of balancing receptive input, productive output, and reflective learning.

Finally, Dual Process Theory (DPT), distinguishing between automatic (System 1) and analytical (System 2) processing, provides a valuable cognitive lens. Language learning requires both the automatization of patterns for fluent performance and the conscious analysis of structures for explicit

knowledge. Effective systems support this duality by combining implicit learning through repeated, controlled exposure with explicit reinforcement via metalinguistic tasks and annotations. Neuroscientific evidence, including the N400 effect (Rugg, 1985), underscores how repetition and contextual priming reduce cognitive load, informing design choices that manage the introduction and recycling of vocabulary.

Taken together, these theoretical underpinnings form a comprehensive foundation for the development of adaptive language learning technologies. By synthesizing insights from input, output, interactionist, lexical, and cognitive models, anchored within the CEFR framework, modern CALL environments embody a holistic approach to language learning. Rather than privileging a single theory, they integrate complementary perspectives to create systems that dynamically respond to learners' evolving linguistic profiles. This theoretical grounding distinguishes current adaptive approaches and positions them as a bridge between established SLA research and the next generation of intelligent learning technologies.

2.9 Looking Ahead: Opportunities in Digital Language Learning

Despite the critiques and limitations discussed above, Computer-Assisted Language Learning (CALL) and Intelligent CALL (ICALL) continue to offer rich opportunities for innovation in second language acquisition. As digital technologies advance, the potential to integrate theoretical insights from SLA into adaptive, scalable, and interactive platforms grows stronger, opening new pathways for pedagogy and research alike.

One of the most promising opportunities lies in the integration of ICALL with Mobile-Assisted Language Learning (MALL). Smartphones and tablets extend language learning beyond the classroom, allowing learners to practice in authentic contexts, on the bus, at work, or during travel. When combined with ICALL's adaptive algorithms, mobile learning can provide learners with personalized micro-lessons, vocabulary review via spaced repetition, or pronunciation practice aligned with Ladefoged and Johnson's (2014) phonetic principles. This convergence enables continuity between formal instruction and informal learning, a long balance emphasized in applied linguistics.

Another opportunity is the incorporation of gamification and immersive environments. By leveraging game mechanics such as rewards, progress tracking, and competition, CALL systems can enhance motivation and learner engagement. Virtual reality (VR) and augmented reality (AR) environments further expand this potential, allowing learners to practice pragmatics in situationally rich contexts. For example, VR simulations of workplace meetings or intercultural encounters can operationalize

Searle's (1969) theory of speech acts and Blum-Kulka et al.'s (1989) work on pragmatics, making learners aware of the subtleties of register, politeness, and indirectness.

Big data and learning analytics represent a further frontier. By collecting and analyzing large volumes of learner interaction data, ICALL platforms can identify patterns of acquisition, error frequency, and task effectiveness across populations. This not only supports adaptive feedback for individual learners but also allows researchers to test SLA theories on a scale. For instance, questions about the role of input frequency (Krashen, 1985; Nation, 2001), output (Swain, 1995), or interaction (Long, 1996) can be investigated using real-world learner data collected in ICALL environments. In this sense, ICALL functions not only as a teaching tool but also as a research laboratory for applied linguistics. Moreover, ICALL creates opportunities for alignment with the CEFR (Council of Europe, 2001, 2020), ensuring that learner progress can be benchmarked against widely recognized standards. Adaptive platforms can map performance on vocabulary, syntax, pronunciation, and pragmatics directly to CEFR descriptors, providing transparent pathways for learners and instructors. This connection enhances both the accountability of ICALL systems and their usability across educational contexts worldwide.

The future also points toward hybrid teacher–AI collaboration. Rather than replacing instructors, ICALL can relieve them of repetitive assessment and error-correction tasks, freeing time for higher-order teaching such as facilitating discussions, guiding intercultural awareness, or fostering learner motivation. Teachers remain central mediators of learning, but ICALL can extend their reach and precision, ensuring that instruction is both human-centered and data-informed.

Finally, there is growing potential for ICALL to foster equity and inclusion in language learning. By making high-quality instruction accessible on low-cost mobile devices, ICALL could help bridge gaps in educational opportunities, particularly for learners in under-resourced regions. Open-source initiatives and publicly available platforms may play a vital role here, ensuring that the benefits of intelligent technologies extend beyond elite educational settings.

In summary, the opportunities for CALL and ICALL lie not only in technological sophistication but in their ability to embody and extend the principles of SLA theory. From Krashen's input-rich exposure to Swain's output-driven noticing, from Long's interactional negotiation to Nation's lexical frameworks and Searle's pragmatic competence, intelligent technologies can operationalize decades of research within adaptive, engaging, and accessible learning environments. The challenge for the future is to ensure that these opportunities are pursued ethically, inclusively, and with pedagogical integrity, so that innovation remains firmly grounded in the realities of language acquisition.

2.10 Conclusion

The trajectory of Computer-Assisted Language Learning (CALL) reflects the broader evolution of language pedagogy itself. From the early drill-and-practice programs shaped by behaviorist principles, through the communicative and integrative phases aligned with the rise of CLT, TBLT, and the CEFR, CALL has consistently mirrored the shifting priorities of second language education. The most recent emergence of Intelligent CALL (ICALL) represents a qualitative leap: by leveraging artificial intelligence, natural language processing, and adaptive algorithms, ICALL operationalizes decades of SLA theory within responsive, learner-centered environments.

The discussion in this chapter has shown that CALL and ICALL are not merely technological innovations, but pedagogical tools deeply intertwined with the insights of leading scholars. Krashen's emphasis on comprehensible input, Swain's call for pushed output, Long's interactionist framework, Ellis's task-based perspective, Nation's lexical principles, Ladefoged and Johnson's phonetic foundations, and Searle's pragmatics all find echoes in contemporary CALL design. At their best, ICALL platforms embody these theoretical principles, turning abstract models of acquisition into practical learning pathways. At the same time, the critiques of CALL and ICALL, technological limitations, pedagogical imbalances, issues of equity, and ethical concerns, remind us that innovation must be pursued with caution and responsibility.

Looking forward, the opportunities for CALL and ICALL lie in their ability to integrate seamlessly with broader educational ecologies. The future points toward hybrid teacher–AI collaboration, mobile and immersive learning, and learning analytics that link individual growth to global research agendas. Yet the ultimate measure of success will remain whether these technologies enhance learners' ability to use language meaningfully in diverse real-world contexts.

This sets the stage for the next chapter, which turns from theory and critique to practical applications of CALL and ICALL. The following discussion will examine concrete case studies, platforms, and design principles, demonstrating how innovations in technology can be harnessed to create adaptive, scalable, and pedagogically sound environments for second language acquisition.

3 Methodology: Didactics and Computer Science

3.1 Introduction

The present research has been conceived and developed at the intersection of two traditionally distinct domains: didactics and computer science. On the one hand, it draws upon pedagogical theories from The present research sits at the intersection of pedagogy and technology, combining insights from Second Language Acquisition (SLA), cognitive psychology, and instructional design with computational frameworks and data-driven methodologies. By balancing these perspectives, the project ensures that technical solutions align with pedagogical objectives, while didactic principles remain compatible with computational constraints. This dual orientation has guided both the design and implementation of the system, requiring the researcher to assume the combined role of educator and developer.

Building on the theoretical insights and opportunities outlined in Chapter 2, the methodology translates SLA principles and ICALL innovations into a practical, adaptive language learning system. Key pedagogical frameworks, from Krashen's comprehensible input to Swain's output-driven tasks, inform the design, while adaptive, mobile, and immersive technologies enhance learner engagement. In this way, the approach bridges theory and practice, demonstrating how intelligent technologies can support meaningful, evidence-based language learning.

3.2 Introduction to Didactic Methodology

Recent advancements in Second Language Acquisition (SLA) research, combined with the growing influence of computer-assisted linguistic analysis, have significantly transformed modern language learning methodologies. These developments arise from a deeper understanding of the cognitive mechanisms underlying language acquisition, alongside continuous refinement of pedagogical strategies aimed at enhancing learner engagement and long-term retention.

Research in lexical-semantic networks has shown that the structure of a learner's mental lexicon is profoundly influenced by the learning environment, particularly in non-traditional language learning contexts. Studies by Beckage et al. (2011), Bilson et al. (2015), and Kenett et al. (2013) demonstrate that external factors can shape the organization of lexical-semantic representations, while more recent

work, such as Borodkin et al. (2025), explores how second language learners' lexical networks differ from those of native speakers. Synthesizing these theoretical insights with cutting-edge computational tools has led to the emergence of innovative language learning systems tailored to diverse learner profiles.

The alignment effect in language production extends beyond isolated lexical items to include word frequency, ranking, collocation, and morphology, elements unified under the broader concept of word usage in context (H. Chen, 2021). Language production tasks in second language assessment typically evaluate a learner's ability to produce meaningful and coherent language in response to specific prompts or scenarios, measuring vocabulary, grammar, cohesion, and discourse organization. The input language also exerts a strong influence on written production: high-intensity, interactive tasks often result in more natural and self-organized output compared to low-intensity tasks, suggesting that the facilitating effects of alignment are closely tied to the depth and frequency of interaction (H. Chen, 2021). Similarly, the content of the input itself plays a crucial role, as variations in input type can lead to distinct outcomes in learner production.

Building upon these insights, contemporary adaptive language learning systems draw on sophisticated computational models to personalize instruction dynamically. Such systems analyze learner data to tailor the learning process to individual needs, proficiency levels, and learning preferences. This personalization not only enhances learning efficiency but also sustains motivation by aligning instructional materials with learners' interests and cognitive styles. Real-time analysis of user interactions enables continuous optimization of learning pathways, ensuring that learners receive targeted feedback and appropriately scaffolded instruction throughout their progression.

Through this seamless integration of SLA theory and computational innovation, adaptive, data-informed learning environments represent a paradigm shift in second language education. Moving beyond traditional one-size-fits-all approaches, these systems provide engaging and responsive learning experiences that empower users to achieve linguistic goals with greater autonomy and efficacy. As technology continues to evolve, such applications hold the potential to redefine language education, making it more personalized, accessible, and effective than ever before.

The development of these adaptive systems is guided by a commitment to bridge the gap between SLA theory and real-world practice. Central to this effort is the integration of computational linguistics, which allows vast quantities of linguistic data to be processed and analyzed for pedagogical use. This data-driven approach facilitates the creation of highly individualized learning pathways, tailored to each learner's proficiency level, cognitive profile, and areas of interest.

A defining feature of these systems is the implementation of adaptive algorithms that dynamically adjust content and difficulty in real time. By continuously calibrating instructional input to match the

learner's evolving competence, such mechanisms maintain engagement while preventing both cognitive overload and stagnation. This adaptability directly addresses long-standing challenges in second language learning, such as fluctuating motivation, retention difficulties, and the plateau effect commonly observed at intermediate levels.

Moreover, modern adaptive learning environments place strong emphasis on contextualized and interest-driven instruction. By aligning materials with learners' personal goals and real-world communicative contexts, these systems foster deeper cognitive processing and more meaningful engagement. The result is not only improved linguistic performance but also enhanced learner investment and autonomy.

Following initial stages of theoretical modeling and data validation, the design of such systems typically progresses into the technical and didactic implementation phase. At this stage, theoretical insights from SLA research are systematically translated into practical design features, informing the development of learning modules, feedback mechanisms, and adaptive processes. Iterative testing and refinement play a central role, with both educational content and computational algorithms undergoing continuous evaluation to enhance engagement, relevance, and learning efficacy.

This iterative, evidence-based approach ensures that adaptive systems are pedagogically grounded and technically robust. By aligning theoretical advancements with computational methodologies, developers are able to create user-centered learning environments that combine educational soundness with technological precision. This synthesis of academic research and practical design exemplifies how technology-enhanced learning can advance second language education in impactful ways.

Early-stage deployment among pilot groups of learners is an essential component of system validation. Pilot studies typically encompass diverse proficiency levels, learning styles, and personal interests to ensure broad applicability. Feedback from these controlled trials informs targeted refinements in instructional design, user experience, and algorithmic adaptivity. Special attention is given to the system's responsiveness to learner trajectories, the clarity of instructional content, and the effectiveness of feedback in maintaining motivation.

Empirical insights from these trials play a crucial role in shaping subsequent iterations of adaptive systems, ensuring both usability and pedagogical soundness prior to large-scale deployment. Such empirically informed refinement highlights the importance of integrating continuous evaluation with innovation in digital education.

A central technical component in these systems is the adaptive difficulty algorithm, which modulates task complexity and text difficulty in real time according to user performance. This personalization optimizes comprehension, retention, and progression by aligning instructional input with learners'

evolving capacities. Parallel attention to interface design ensures accessibility and intuitive navigation, enabling efficient and meaningful interaction with learning materials.

For practical reasons, many such systems adopt asynchronous learning formats, allowing learners to proceed at their own pace. This flexibility accommodates diverse schedules and learning preferences while supporting greater autonomy, key factors in maintaining engagement over the long term.

Through the integration of adaptive algorithms, data-driven curriculum design, and user-centered interface development, modern language learning systems exemplify a commitment to evidence-based practice in second language acquisition. Their evolution reflects a growing synthesis between cognitive theory, pedagogical principles, and computational capability, an integration that continues to redefine the landscape of digital language education.

To provide a structured overview of this development process, the following discussion is organized into two main sections. The first focuses on the didactic design of adaptive learning systems, addressing pedagogical principles, curriculum structures, and instructional strategies informed by SLA theory. The second examines the technical implementation of these systems, detailing architecture, interface design, backend infrastructure, and the integration of adaptive learning technologies. This dual focus reflects the inherently interdisciplinary nature of language learning innovation, rooted in both educational theory and technological advancement, and illustrates how conceptual frameworks are systematically transformed into effective, functional learning environments.

3.2.1 Promoting Self-Regulated Learning

As attention turns to the self-regulated learning dimension of adaptive web-based language learning environments, it becomes essential to understand how this component integrates with the broader educational approach. Self-regulated learning constitutes a central pedagogical principle, empowering learners to take ownership of their educational process. This involves setting personal goals, monitoring progress, and adapting strategies to meet individual needs.

To support learner autonomy, such systems are designed to provide virtually limitless opportunities for exploration, leveraging the extensive resources of openly available repositories such as Wikipedia. This vast database enables learners to select topics aligned with their personal interests and areas of relevance, thereby enhancing motivation and engagement. The following sections examine how these self-regulated learning elements are embedded in the course design, detailing the mechanisms that facilitate self-assessment, goal-setting, and reflective practice, and how these mechanisms contribute to fostering independent, motivated, and effective language learners.

An essential feature of modern language learning systems is the integration of written resources with speech synthesis technology. Written resources, such as the comprehensive repository of Wikipedia content, provide rich textual material that serves as the foundation for linguistic input. Speech synthesis technology transforms these texts into spoken language, enabling learners to hear correct pronunciation, intonation, and rhythm. This integration bridges the gap between reading and listening skills, catering to diverse learning styles by reinforcing comprehension through auditory as well as visual channels. The combination of textual and auditory modalities creates a more immersive and comprehensive learning experience, aiding pronunciation, listening, and reading fluency while reinforcing understanding and retention.

3.2.1.1 Challenges in Data Selection and Integration

During the development of such systems, significant challenges arise concerning the availability, suitability, and licensing of data repositories for language learning. The process of identifying datasets that are both pedagogically relevant and sufficiently diverse in content is complex. Ensuring data quality while maintaining alignment with educational best practices requires careful curation, particularly when covering a wide range of themes and contexts.

A primary obstacle lies in finding materials that satisfy both pedagogical and legal requirements. Many existing repositories are governed by restrictive licenses that limit their use in educational applications, especially those emphasizing self-regulated learning and large-scale resource integration. To address these limitations, openly licensed resources such as Wikipedia present an effective solution. Its Creative Commons Attribution–ShareAlike (CC BY-SA 4.0) license allows flexible adaptation and integration, ensuring compliance with copyright laws while providing access to an extensive and continuously updated corpus of texts.

This accessibility makes Wikipedia an invaluable foundation for educational tools that aim to provide learners with an expansive and varied selection of topics. Its open structure and thematic diversity foster an engaging and autonomous learning environment, while its continuous community-driven updates ensure reliability and accuracy.

3.2.1.2 Ensuring Infinite Choice and Learner Autonomy

A core design goal of adaptive web-based learning systems is to offer learners an infinite potential for choice, a condition that supports self-regulation, curiosity, and engagement. Wikipedia's encyclopedic range of subjects provides the breadth necessary to achieve this objective. By granting

learners the freedom to explore virtually any topic of personal or academic interest, such systems encourage deeper involvement and intrinsic motivation.

Selecting Wikipedia as the principal input source is advantageous not only for its licensing flexibility but also for its breadth of coverage and up-to-date information. This diversity ensures that learners can pursue topics that align with individual goals and proficiency levels, accommodating a wide spectrum of interests while maintaining pedagogical rigor.

3.2.1.3 Content Curation and Pedagogical Alignment

The identification and integration of appropriate input content are critical to the pedagogical effectiveness of adaptive language learning environments. Even before a complete didactic framework is implemented, it is vital to establish a systematic strategy for sourcing and structuring content to ensure coherence with educational objectives.

Content selection must satisfy several key criteria, including quality, to guarantee that texts are accurate, contextually appropriate, and pedagogically sound; quantity, to provide a sufficient volume of material for sustained practice; thematic diversity, to expose learners to a broad range of vocabulary, grammatical structures, and cultural references; and licensing compliance, to ensure the lawful and ethical use of open-access materials.

Through rigorous evaluation and careful integration of openly licensed resources, developers can ensure a legally compliant and pedagogically robust foundation for adaptive learning systems.

3.2.1.4 Designing Adaptive Algorithms for Learner-Centered Content

To support personalized learning, adaptive algorithms must be supplied with detailed, structured input data. Each text is analyzed and categorized based on linguistic variables such as lexical complexity, sentence length, syntactic structure, and thematic relevance. These linguistic annotations serve as metadata, enabling the system to match texts with learners' evolving proficiency levels.

By dynamically adjusting text complexity in response to user performance, adaptive systems maintain the delicate balance between challenge and support, ensuring optimal cognitive engagement without overloading working memory. This approach facilitates smooth, individualized progression in language learning and helps prevent stagnation or frustration.

3.2.1.4.1 Content Management and Asynchronous Learning

Given the web-based and often asynchronous nature of these environments, effective content management is essential. A robust content management system (CMS) ensures that learning materials

are logically organized, easily retrievable, and accessible on demand. This structure allows learners to navigate at their own pace, reinforcing autonomy while maintaining curricular coherence.

A well-designed CMS also supports iterative content updates based on learner feedback and performance data. Continuous monitoring and refinement ensure that educational materials remain relevant, accurate, and aligned with learner needs, creating a dynamic cycle of evaluation and enhancement that sustains both engagement and pedagogical integrity.

3.2.1.4.2 Integration of Linguistic Linked Data

Beyond Wikipedia, the incorporation of linguistic linked data further enriches the learning experience by connecting texts to broader semantic networks. By linking words and expressions to databases such as WordNet and other Linguistic Linked Open Data (LLOD) resources, learners are exposed to the relational nature of language. This allows exploration of synonyms, antonyms, and semantic associations, deepening understanding of lexical patterns and contextual meaning.

This approach shifts learning from isolated vocabulary acquisition to relational, network-based understanding, reflecting the interconnectedness of real-world language use. Integrating such linguistic data enhances both the academic rigor and the pedagogical depth of digital learning environments, encouraging inductive reasoning and fostering a deeper awareness of language as a dynamic system.

3.2.1.4.3 Benefits of Open, Encyclopedic Content

The encyclopedic nature of Wikipedia offers unparalleled advantages for language learning. Its vast and diverse body of content ensures exposure to a wide array of subjects, ranging from culture and science to technology and everyday life, broadening both linguistic and general knowledge. This variety sustains engagement by allowing learners to align language practice with their personal interests.

The constant updating and fact-checking performed by Wikipedia's global community provide additional educational benefits, ensuring that learners encounter current, accurate, and reliable information. This continuous renewal of content maintains the relevance of materials and supports the principle of learning through authentic, up-to-date language.

By leveraging openly licensed and continuously evolving content, adaptive language learning systems are able to offer differentiated, personalized, and contextually rich experiences. The resulting environments combine technological flexibility with pedagogical soundness, empowering learners to

progress at their own pace while engaging with authentic, meaningful language in real-world contexts.

3.2.1.4.4 Open Licensing, Collaborative Knowledge Sharing, and Multimodal Integration in Adaptive Language Learning Systems

A defining principle of modern digital language learning environments is their commitment to open access and collaborative knowledge sharing. This philosophy is reflected in the adoption of open licensing frameworks that encourage transparency, accessibility, and community engagement. In alignment with openly licensed content such as that found on Wikipedia, many educational platforms adopt the Creative Commons Attribution–ShareAlike 4.0 International License (CC BY-SA 4.0). This license enables users to freely access, share, and adapt educational materials, provided that appropriate attribution is given and derivative works are shared under the same terms. By adhering to this open model, such systems promote a culture of collaboration, empowering educators and learners alike to contribute improvements and insights back to the community. This approach not only enhances credibility and transparency but also aligns digital education with the broader goals of equitable access to knowledge and participatory learning. The licensing advantages of using Wikipedia and Wikimedia Commons content, both available under the CC BY-SA 4.0 license, have proven particularly significant in the design of adaptive educational systems. This framework allows for free use, adaptation, and redistribution of textual and visual resources, removing barriers imposed by traditional copyright restrictions. As a result, educational platforms can seamlessly integrate open materials, modify them to fit specific didactic needs, and ensure that the resulting resources remain fully compliant with legal standards. Open licensing also supports the customization of content to meet pedagogical objectives. Developers can adapt Wikipedia articles for specific language levels or curricular themes, enriching them with annotations, exercises, or linked linguistic data. This flexibility fosters continuous refinement of educational materials, enabling responsive updates based on user feedback and performance data. Moreover, the vast and constantly updated pool of Wikipedia content provides a scalable foundation for digital platforms seeking to expand their linguistic and thematic offerings.

3.2.1.4.5 Integration of Visual Resources and Multimodal Interaction

The open licensing model extends to the vast array of images available through Wikimedia Commons, offering a powerful means to enhance multimodal learning. Freely accessible images, diagrams, and

infographics enrich text-based materials, providing learners with visual context that reinforces understanding and retention. Visual resources are especially valuable in language education, where imagery can aid comprehension of abstract or culturally specific concepts, support vocabulary learning, and improve pronunciation through visual associations. Incorporating visual materials also accommodates diverse learning styles. For example, visual learners may benefit from illustrations that clarify grammatical structures, while infographics can simplify complex linguistic data. The inclusion of such multimodal elements fosters engagement by making lessons more interactive, dynamic, and relatable. Because these images are distributed under flexible Creative Commons licenses, they can be adapted and reused without legal restrictions, ensuring both pedagogical richness and compliance with copyright law. Furthermore, the thoughtful integration of images encourages interactive learning behaviors. Learners can analyze, describe, and discuss visual content, linking textual and visual modes of expression to deepen comprehension and promote critical thinking. This multimodal approach strengthens the connection between language and meaning, leading to a more holistic and immersive learning experience.

3.2.1.4.6 Licensing as a Foundation for Open and Scalable Design

The adoption of the Creative Commons Attribution–ShareAlike 4.0 license plays a crucial role in enabling the core functionality and scalability of adaptive language learning environments. The research and alignment of licensing frameworks often represent a complex but essential stage of development, ensuring that all content, textual, visual, and interactive, can be integrated seamlessly and lawfully. By operating within the CC BY-SA 4.0 framework, educational systems gain access to a vast reservoir of open resources while preserving the ability to expand and update content continuously. This model supports sustainable growth, reduces reliance on proprietary materials, and ensures that the learning platform remains adaptable to evolving educational needs. Open licensing thus serves as both a legal safeguard and a mechanism for innovation, facilitating collaboration and continuous improvement across the learning community. Importantly, aligning the learning platform under an open license reinforces the same principles embodied in its content. Making the system reproducible and shareable under the CC BY-SA 4.0 license ensures consistency between pedagogical philosophy and technological implementation. This transparency strengthens community trust and encourages educators, researchers, and learners to build upon and contribute to the system's evolution.

3.2.1.4.7 Challenges of Attribution and Licensing Management

Despite its benefits, open licensing introduces certain operational challenges, particularly in environments that allow users to freely modify and adapt materials. When learners are empowered to edit or remix content, maintaining accurate attribution and license compliance becomes complex. To address this issue, advanced systems have implemented automated attribution mechanisms. These tools detect content modifications, apply appropriate credits to original authors, and automatically assign the correct license information to derivative works. Such systems ensure that all user-generated adaptations remain compliant with open-access regulations, while reducing administrative burden and safeguarding intellectual property rights. Implementing automated attribution requires careful design and technical sophistication. Identifying modified content, tracking its origins, and ensuring accurate credit assignment are intricate tasks. However, achieving this balance between flexibility and legal compliance is essential for sustaining an ethical and transparent open-access ecosystem. As users navigate and interact with materials, visible attribution features reinforce transparency and intellectual integrity. Learners can easily view the source and licensing information associated with each resource, fostering awareness of copyright practices and ethical content use. This approach not only fulfills legal requirements but also cultivates digital literacy and respect for collaborative authorship, key competencies in today's interconnected educational landscape.

3.2.1.4.8 Linked Open Data and the Challenge of Linguistic Resources

The discussion surrounding Linked Open Data (LOD), and specifically Linguistic Linked Open Data (LLOD), highlights a critical challenge in the field of language education: the limited availability of comprehensive, openly accessible linguistic datasets. Despite growing interest in leveraging LLOD to enhance educational technologies, resources covering linguistic phenomena such as phonetics, syntax, and semantics remain scarce and fragmented. This underrepresentation restricts the potential of linked data to support rich, contextually informed language learning. The lack of cohesive datasets across multiple languages and proficiency levels limits opportunities for integrating semantic, morphological, and syntactic insights into digital learning tools. Consequently, addressing these data gaps is vital to advancing the next generation of educational applications. Developing robust LLOD frameworks would allow learning systems to connect textual and lexical data to broader semantic networks, thereby enriching the pedagogical experience. Integrating linguistic resources into adaptive learning environments can enhance vocabulary acquisition, semantic awareness, and cross-linguistic understanding, ultimately fostering a more comprehensive grasp of language structure and use.

3.2.1.5 Word-Unit as the Primary Level of Analysis

In line with our dual focus on both word-unit and text-level analyses, we prioritize the word-unit as the primary level of analysis for language learning. While recognizing the importance of multi-word expressions and their impact on overall language comprehension, our approach emphasizes the detailed examination of individual words. This method involves processing each word separately to ensure a thorough understanding of its meaning, usage, and contextual relevance. By focusing on the word-unit, we provide learners with targeted and precise enhancements, laying a solid foundation for vocabulary acquisition. Although multi-word expressions are also considered, our detailed attention to single-word analysis ensures that learners grasp the fundamental elements of language before integrating them into broader textual contexts.

This method ensures that each word is thoroughly analyzed and enhanced with sufficient context and support. By concentrating on individual words, we provide detailed and targeted feedback that facilitates deeper understanding and retention. This focused approach allows learners to build a strong vocabulary foundation and gain clarity on word usage before integrating multi-word expressions. We believe that this step-by-step enhancement of word-units offers the greatest efficacy in language acquisition, as it enables learners to grasp the nuances of each word and apply them effectively within broader linguistic contexts.

Each word is meticulously processed to ensure comprehensive understanding and effective learning. This process begins with analyzing the word's meaning, followed by its meaning within various contexts to address potential ambiguities. The application employs sophisticated disambiguation techniques to clarify multiple interpretations of a word, ensuring that learners encounter accurate and relevant information. Once the word is fully understood in its various contexts, it is then processed to determine the appropriate level of enhancement or lack thereof, based on the learner's prior knowledge. This tailored approach allows the application to provide precise and contextualized support, offering the right amount of enhancement or reinforcing existing knowledge as needed. By addressing each word at this detailed level, we ensure that learners receive a clear and contextually relevant understanding, facilitating more effective vocabulary acquisition and application.

The primary objective of the project is to guide learners from an absolute beginner level to a comprehensive vocabulary base of over 4,000 words, corresponding approximately to the B1 level of proficiency. This progression is intended to provide learners with a solid lexical foundation, enabling effective engagement with a wide range of texts and communicative contexts. The learning trajectory, however, does not conclude at this milestone. The application is designed to foster ongoing engagement and continuous learning, allowing users to extend their vocabulary acquisition beyond

the initial target. Learners are encouraged to explore new texts, participate in additional activities, and pursue topics of personal interest to further their language development. This open-ended approach ensures opportunities for continual expansion of linguistic repertoire and advancement toward higher levels of proficiency according to individual goals and motivations.

To support the advancement from a foundational level to a comprehensive vocabulary base of over 4,000 words, a large and dynamic lexical database has been developed. This database initially comprises a core set of 1,000 foundational words, serving as the starting point for learners. As learners interact with the application and conduct independent exploration, the database expands dynamically to include new words and expressions encountered during the learning process. This expansion is achieved through cross-referencing with user searches and activities, creating an implicit collaborative environment that adapts to individual interests and learning needs. By continuously updating and enriching the database based on learner interactions, the system maintains vocabulary coverage that is both relevant and personalized, thereby promoting a more engaging and effective learning experience.

An initial focus in enhancing vocabulary learning concerns the provision of visual representations for each word to support comprehension and retention. For this purpose, each lexical item is linked to an image that visually conveys the concept associated with the term. This visual enhancement constitutes a fundamental stage in the learning process, as it facilitates the formation of clear mental associations and promotes a more intuitive understanding of meaning. By associating words with images, the system fosters a multimodal learning experience that connects linguistic and visual information. This approach not only aids vocabulary acquisition but also reinforces conceptual understanding, accommodating diverse learning styles and enhancing overall learning effectiveness.

In addition to visual supports, precise and contextually appropriate definitions are provided for each word to ensure accurate comprehension. Emphasis is placed on the value of contextualization and lexical interconnectivity. Accordingly, the application incorporates knowledge graphs to represent semantic relationships among words. These graphs visually illustrate the interconnections between lexical items, enabling learners to explore and understand the relationships within the vocabulary. Engagement with these semantic networks allows learners to infer or retrieve meanings by recognizing links between related terms. This approach facilitates not only the acquisition of individual vocabulary items but also the development of deeper lexical understanding within a broader linguistic framework. The integration of knowledge graphs thus promotes a more interconnected and cognitively rich learning process, encouraging learners to make associations and apply their knowledge flexibly across contexts. To ensure robust and accurate representation of semantic relationships, our application draws on WordNet as a key reference. WordNet, a

comprehensive lexical database, provides a rich source of information on word meanings, synonyms, antonyms, and various semantic connections. By leveraging WordNet, we are able to create detailed knowledge graphs that map out the intricate relationships between words, enhancing learners' understanding of vocabulary through a well-established linguistic resource. These semantic networks, informed by WordNet, enable learners to see how words are related in terms of meaning and usage, facilitating deeper comprehension and more effective retrieval of word meanings. This integration of WordNet supports a sophisticated approach to vocabulary learning, ensuring that the connections drawn between words are both accurate and meaningful.

By integrating WordNet into our application, we not only provide learners with accurate semantic relationships but also facilitate a more interactive and engaging learning experience. The detailed knowledge graphs created using WordNet enable learners to visualize and explore the connections between words, fostering a deeper understanding of vocabulary. These semantic networks allow learners to engage in activities such as guessing meanings based on word relationships or exploring related terms, thereby reinforcing their knowledge and enhancing retention. This approach encourages active learning and critical thinking, as learners are invited to interact with the vocabulary in a meaningful way. Additionally, the ability to trace and analyze semantic connections supports the development of advanced language skills, as learners become adept at understanding and using words within various contexts. Overall, the use of WordNet's semantic networks enriches the learning experience by providing both a comprehensive and interactive framework for vocabulary acquisition. While the integration of WordNet provides a robust foundation for understanding semantic relationships, the automatic recognition of the correct sense of a word in context requires further implementation beyond WordNet's capabilities. The complexity of natural language often necessitates additional algorithms and contextual analysis to accurately determine the appropriate meaning of a word based on its usage within specific texts. To address this challenge, our application employs advanced computational techniques and contextual algorithms that complement WordNet's data. These additional implementations analyze the surrounding text to disambiguate word meanings and ensure that learners receive the most relevant and precise information. By combining WordNet with these sophisticated methods, we enhance the accuracy of word sense recognition and provide learners with more contextually appropriate vocabulary enhancements, ultimately supporting a more effective and nuanced learning experience.

Before processing a word as a unit, we acknowledge that its meaning can vary significantly depending on the sentence or text level. Words are inherently dynamic, and their meanings can shift based on contextual cues and surrounding linguistic elements. Therefore, our application is designed with an acute awareness of how the same word can assume different meanings within various contexts. This

contextual variability necessitates a nuanced approach to word processing, where the initial analysis of a word is informed by its potential meanings in different sentences or texts. By considering these contextual factors, we ensure that the word is interpreted accurately, taking into account the specific usage and meaning it conveys in each instance. This approach allows for a more precise and contextually relevant processing of words, enhancing the overall effectiveness of vocabulary learning and comprehension.

Accurate word sense disambiguation is crucial for effective vocabulary learning, particularly in understanding words within their contextual usage. To address this, our application utilizes the Lesk algorithm, a well-established method for word sense disambiguation, in conjunction with WordNet. The Lesk algorithm disambiguates word meanings by comparing the context in which a word appears with the definitions of its possible senses provided in a lexical database like WordNet.

The Lesk algorithm operates by calculating the overlap between the words in the context of a target word and the words in the definitions of its potential senses. The sense with the highest overlap is selected as the most likely meaning of the target word in that context. This approach leverages WordNet's comprehensive lexical entries, which include detailed definitions and examples, to inform the disambiguation process.

By applying the Lesk algorithm, our application can effectively resolve ambiguities and provide learners with the most contextually appropriate sense of a word. This method ensures that the vocabulary presented to learners is relevant and accurate, enhancing their comprehension and retention of new terms. The combination of the Lesk algorithm and WordNet supports a sophisticated approach to word sense disambiguation, contributing to a more precise and contextually aware learning experience.

Addressing the variability in word meanings based on sentence or text level has necessitated advanced technical implementations beyond standard word processing techniques. To effectively manage these contextual changes, our application incorporates sophisticated algorithms and contextual analysis methods. These technical implementations are designed to dynamically interpret word meanings within different contexts, ensuring accurate and relevant vocabulary processing. The intricacies involved in handling contextual variability require a detailed technical approach, which will be further elaborated in the following sections. This additional layer of implementation is crucial for maintaining the precision and relevance of word sense disambiguation, ultimately enhancing the learning experience by providing contextually appropriate vocabulary insights.

3.2.1.6 Text-Level Processing and Complexity Analysis

In addition to word-unit analysis, our application emphasizes a comprehensive approach to text-level processing. Recognizing that effective language learning extends beyond individual words, we analyze entire texts to assess their complexity and facilitate a deeper understanding of vocabulary within broader contexts. To achieve this, we have developed a custom algorithm designed to evaluate the overall complexity of a given text. This algorithm considers various factors, such as sentence length, syntactic structures, and vocabulary diversity, to provide an average complexity score for the text.

The text-level analysis within our application is performed through a specialized feature known as the *text input area*, situated on the learning page. This *text input area* functions as an advanced analytical tool where texts are processed to evaluate their complexity, identify key vocabulary and grammatical structures, and generate contextual insights. The area integrates various algorithms and processing techniques to provide a detailed breakdown of the text, including syntactic and semantic analysis, dependency parsing, and coreference resolution. By utilizing this interactive and dynamic component, learners receive a comprehensive evaluation of texts in real-time, which enhances their understanding and engagement with the material. The *text input area* is integral to connecting text comprehension with personalized learning activities, thereby improving the overall educational experience. Learners have the flexibility to arrive at the learning page with a text selected from home, bringing their personal choice of reading material into the educational environment. This feature allows users to integrate texts they are interested in or are already familiar with, facilitating a more personalized and engaging learning experience. By starting with a text of their choice, learners can immediately apply the application's analytical tools, including the *text input area*, to their preferred content. This approach not only enhances motivation and relevance but also supports a seamless transition from independent reading to structured language learning activities. The ability to bring in texts from home ensures that the learning process is closely aligned with the learner's interests and needs, promoting a more effective and enjoyable educational experience.

Once a text is submitted for processing, the algorithms in the covert modality perform a comprehensive set of operations behind the scenes. These covert processes include syntactic and semantic analysis, dependency parsing, coreference resolution, and the extraction of key vocabulary and grammatical patterns. The algorithms work discreetly to assess the text's complexity, identify relevant linguistic features, and map out semantic connections. This extensive processing ensures that all necessary information is accurately captured and analyzed without interrupting the user experience. The results of these covert operations are then used to tailor the text's presentation and

the associated learning activities, providing learners with detailed insights and targeted enhancements based on the processed text.

During the processing phase, all the didactic and second language acquisition principles that have been carefully integrated into the application come into practice. This phase involves applying various pedagogical strategies and theoretical frameworks to ensure that the text is enhanced in a way that supports effective learning. These principles guide the algorithms in providing contextualized definitions, generating relevant examples, and tailoring feedback to the learner's needs. However, from the learner's perspective, the process is seamless, and the only visible outcome is the enhanced text that appears at the end of the process. This enhanced text reflects the application of these principles, offering a richer and more supportive learning experience without disrupting the learner's engagement with their chosen material.

Once the text has been processed, it is securely stored in the learner's personal memory within the application. This stored text serves as a valuable reference for future use. Additionally, words identified as *to enhance* during the processing phase are transformed into interactive flashcards. These flashcards are designed for subsequent practice and reinforcement on a dedicated page within the application. By converting enhanced words into flashcards, the application provides learners with an effective tool for vocabulary retention and review. This approach ensures that the newly acquired language elements are revisited and practiced, supporting long-term learning and mastery of the language.

In our approach, we metaphorically steer the wheel towards achieving an optimal balance of Krashen's *input +1*. By empowering learners to choose texts that align with their interests, we ensure they engage with language that is just beyond their current proficiency level, thereby maintaining an effective challenge without overwhelming them. This strategic navigation helps to fine-tune the complexity of the input, steering learners towards a balanced learning curve where each new text contributes to their linguistic growth while remaining accessible and engaging. This balance is crucial for fostering continuous development and preventing stagnation, thereby facilitating a dynamic and progressive learning experience.

The principle of free choice in text selection is crucial for aligning with Krashen's input hypothesis, particularly the concept of *input +1*. This phase is integral to maintaining an appropriate balance in text complexity, ensuring that learners are consistently exposed to language that is slightly beyond their current proficiency level. By allowing learners to select texts of personal interest, the application fosters engagement and motivation while also adhering to the principle of providing language input that challenges them just enough to promote growth. This method ensures that learners encounter new vocabulary and grammatical structures within a context that is both meaningful and manageable.

Thus, the process of balancing text complexity through free choice not only aligns with theoretical frameworks but also enhances the overall effectiveness of the language learning experience.

By assessing text complexity, our application can tailor learning experiences to match the proficiency levels of learners, ensuring that texts are appropriately challenging and supportive of their language development. This holistic approach to text processing allows for a more nuanced understanding of how vocabulary functions within complete passages, enhancing learners' ability to comprehend and utilize language in diverse contexts. The following sections will delve into the specifics of our algorithm and its role in text-level analysis.

Being text enhancements strategically provided on both an individual and text basis optimizes the learning experience. On an individual basis, enhancements are tailored to the specific needs and proficiency levels of each learner. This personalized approach ensures that vocabulary and grammatical support are aligned with the learner's current knowledge and learning objectives, allowing for targeted and effective improvements. Enhancements include contextualized definitions, usage examples, and interactive activities designed to reinforce understanding. On a text basis, enhancements focus on improving the overall readability and comprehension of the texts being studied. This includes providing additional contextual information, simplifying complex sentences, and highlighting key vocabulary and grammatical structures. By addressing the text as a whole, our application ensures that learners can engage with content that is appropriately challenging yet accessible, facilitating a more holistic approach to language acquisition. Together, these dual levels of enhancement contribute to a richer and more supportive learning environment, helping learners to progress effectively through their language studies.

On a text basis, enhancements focus on improving the overall readability and comprehension of the texts being studied. This includes providing additional contextual information, simplifying complex sentences, and highlighting key vocabulary and grammatical structures. By addressing the text as a whole, our application ensures that learners can engage with content that is appropriately challenging yet accessible, facilitating a more holistic approach to language acquisition. Together, these dual levels of enhancement contribute to a richer and more supportive learning environment, helping learners to progress effectively through their language studies.

Our first concern in text enhancements involves the implementation of dependency parsing and coreference resolution to build a structural skeleton represented in a graph. Dependency parsing analyzes the grammatical structure of sentences, identifying the relationships between words and how they depend on each other. Coreference resolution, on the other hand, determines when different words or phrases in a text refer to the same entity. By integrating these techniques, we construct a syntactic graph that visually represents the relationships and dependencies within the text. This

graphical representation allows learners to gain a clearer understanding of how textual elements are interrelated, facilitating better comprehension of complex sentences and enhancing their ability to navigate and interpret nuanced language structures. The graph-based approach supports learners in visualizing and managing syntactic relationships, thereby improving their grasp of how language functions at a deeper level.

A second layer of assistance in our application is provided through text color-coding, which visually distinguishes between known words, enhanced words, and named entities (NER). This color-coding system helps learners quickly identify and differentiate between various types of vocabulary and textual elements. Known words, which learners are already familiar with, are highlighted in one color, while enhanced words, which are new or have additional contextual information, are marked in another color. Named entities, such as people, places, or organizations, receive a distinct color to facilitate their recognition. This visual differentiation aids in comprehension by making it easier for learners to focus on and interact with different aspects of the text. By incorporating color-coding, we enhance the readability of the text, support effective learning, and provide learners with immediate visual cues that guide their understanding of both familiar and new information.

Each chosen and processed text within our application is systematically stored in a personal memory, creating a unique and individualized corpus for each learner. This personal corpus consists of all texts the learner has interacted with, including those analyzed for vocabulary, grammar, and contextual understanding. By storing texts in this manner, learners have access to a tailored collection of materials that reflect their learning journey and progress. This personal corpus not only serves as a reference for reviewing previously encountered content but also supports ongoing learning by allowing learners to revisit and engage with texts they have previously processed. The stored texts contribute to a dynamic learning environment, where the personal corpus evolves based on the learner's interactions and preferences, thereby providing a personalized and continuously relevant resource for further study and practice.

Stored texts in the personal corpus can be dynamically retrieved and reprocessed according to the learner's current proficiency level. This feature allows the application to revisit and re-evaluate previously encountered texts, adapting them to match the learner's evolving skillset. As learners progress and their language abilities improve, texts that were initially processed at an earlier stage can be revisited with enhanced complexity and additional focus on advanced vocabulary and grammar. This dynamic reprocessing ensures that learners continually engage with content that is appropriately challenging and relevant to their current level. By adjusting the level of difficulty and incorporating new learning objectives, the application supports ongoing development and reinforces previously learned material, contributing to a more robust and adaptive learning experience.

Following text processing, our application generates various interactive activities to reinforce learning and engagement. Specifically, each text automatically produces memory and cloze games based on the analysis of vocabulary and grammatical units within the text. The memory game focuses on helping learners recall and reinforce their understanding of key terms and phrases identified during text processing. It uses flashcards and matching exercises to strengthen vocabulary retention. The cloze game, on the other hand, involves filling in blanks within sentences or passages, targeting specific vocabulary or grammatical structures highlighted during the text analysis. This game aids in contextualizing and applying learned concepts. Both activities are dynamically generated to align with the processed text, ensuring that they directly reflect the language elements that need attention. By integrating these automated activities, our application provides targeted practice that complements the learner's engagement with the text, enhancing overall comprehension and retention.

3.2.1.7 Accommodating the choices: delivering the right content

In this adaptive language learning system, the vast repository of Wikipedia content is not presented as static information but is actively processed through advanced algorithms to address specific learning needs and track learner progress. Sophisticated computational models analyze and modify the input dynamically, tailoring materials to match each learner's proficiency level, interests, and goals. These adaptive mechanisms adjust language complexity, identify areas of difficulty, recommend personalized topics, and continuously recalibrate content to maintain optimal challenge and engagement. By leveraging these algorithms, the system not only personalizes the learning experience but also supports effective self-regulated learning, helping users achieve their educational objectives efficiently and with sustained motivation.

The system must also accommodate diverse learner choices in terms of text selection and evaluation, functioning with a level of adaptability comparable to that of a human instructor. Much like a teacher who modifies feedback in response to a student's progress, the system dynamically evaluates learner's performance and adjusts assessments accordingly. Pre-configured tools and linguistic metrics allow for consistent and objective evaluation across a wide range of content while retaining flexibility to handle varied topics, styles, and complexity levels. This adaptability ensures that assessments remain accurate and meaningful regardless of learner choice, reinforcing both personalization and rigor in the educational process.

A central design intention is the creation of a covert learning environment, one in which users engage deeply with content without being overtly aware of the instructional mechanisms at work. The focus extends beyond linguistic forms to encompass meaningful interaction with the encyclopedic knowledge embedded in the texts. By integrating educational elements seamlessly into materials that

learners find intrinsically interesting, the system fosters natural and enjoyable language acquisition. Learners are motivated to explore and interact with material, thereby internalizing linguistic knowledge as part of their engagement with authentic content rather than perceiving learning as a separate task.

In alignment with Krashen's Input Hypothesis, the system manipulates content to ensure that learners are consistently exposed to input that lies just beyond their current level of competence, the 'i+1' principle. The algorithms dynamically adjust text difficulty to provide material that challenges learners while remaining comprehensible. This balance supports incremental development, prevents cognitive overload, and promotes continuous linguistic growth. Such dynamic adaptation reflects established SLA theory and enhances the effectiveness of self-regulated learning by aligning exposure with learners' evolving abilities.

Vocabulary acquisition is managed through a structured framework inspired by Paul Nation's lexical research, in which the lexicon is divided into ten progressive levels of frequency and complexity. This organization ensures that learners encounter vocabulary appropriate to their proficiency, gradually expanding from high-frequency core terms to more advanced lexical items. Each stage builds systematically on prior knowledge, enabling learners to consolidate vocabulary through controlled exposure and repetition. The clear progression path provided by this model supports efficient and cumulative lexical development.

Grammatical instruction within the system follows a scaffolded progression that builds on prior knowledge while promoting incidental learning through contextual exposure. New grammatical forms are introduced progressively and reinforced through authentic examples within meaningful texts. Learners thus internalize structures both explicitly, through targeted tasks, and implicitly, through repeated exposure. This dual approach, structured scaffolding coupled with incidental learning, facilitates deeper grammatical understanding and long-term retention while maintaining learner engagement.

The system employs a dual analytical focus encompassing both word-level and text-level processing. At the word-unit level, algorithms examine meaning, context, and disambiguation to provide precise lexical enhancement. Concurrently, text-level analysis considers how words function in larger syntactic and discourse contexts. This combined approach supports comprehensive linguistic understanding and facilitates the transition from vocabulary learning to fluent text interpretation.

To guide learners effectively, assistance is structured through a funnel-based support model. The system begins with a holistic overview of the text and gradually narrows its focus to individual linguistic elements. This approach helps learners develop contextual comprehension before

concentrating on specific vocabulary and grammatical details, ensuring a balanced progression from general understanding to focused mastery.

The effective management of Wikipedia content represents a critical element of the system's design. Given the asynchronous and self-paced nature of the course, it was essential to ensure that materials remained accessible, well-organized, and pedagogically coherent. A dedicated content management system (CMS) enables seamless integration, retrieval, and presentation of texts, allowing learners to engage with lessons at their own pace and according to personal preference.

Regular content updates are guided by learner feedback, performance analytics, and continuous evaluation of instructional effectiveness. This iterative development process ensures that learning materials remain relevant, responsive, and engaging. By analyzing user behavior and outcomes, the system can refine its resources in real time, maintaining both pedagogical rigor and learner motivation.

The identification, evaluation, and integration of Wikipedia content were foundational to the system's adaptive framework. Careful curation of this open corpus provided high-quality, contextually relevant input materials while ensuring compliance with Creative Commons licensing. Beyond Wikipedia, the system incorporates Linguistic Linked Open Data (LLOD) to enhance semantic depth. By connecting textual material to external linguistic databases such as WordNet, learners gain exposure to the relational nature of language through semantic associations, synonyms, antonyms, and conceptual networks.

This integration supports inductive and cognitively rich learning, enabling learners to discover language patterns and semantic structures autonomously. The result is a deeper understanding of language as an interconnected system rather than a collection of isolated forms. The use of linked data also enhances academic rigor, situating learning within a transparent, research-informed framework that encourages exploration and critical engagement with language.

The encyclopedic scope of Wikipedia offers major pedagogical advantages, providing learners with exposure to an unparalleled range of topics across disciplines such as science, history, and culture. This diversity not only enriches linguistic input but also stimulates curiosity and intrinsic motivation. By engaging with content that aligns with their personal interests, learners are encouraged to pursue knowledge while developing language proficiency.

Wikipedia's collaborative model ensures that content is continually updated and fact-checked, offering learners access to current and relevant information. The system capitalizes on this dynamic quality through automatic synchronization with updated entries, exposing users to the latest developments and real-world contexts. This ongoing refreshment maintains relevance, sustains engagement, and fosters authentic language use.

While the open-editing nature of Wikipedia means that occasional inaccuracies may occur, its overall reliability and breadth make it an invaluable resource for language learning. The variety and richness of topics allow learners to engage with material that reflects their personal passions, thereby deepening motivation and cognitive engagement.

A sophisticated adaptive algorithm underpins the system's personalization capabilities. Wikipedia texts are linguistically analyzed and categorized according to lexical difficulty, syntactic complexity, and thematic relevance. These metadata annotations inform the adaptive model, enabling it to recommend appropriate materials, adjust text complexity, and scaffold learning progression dynamically.

A comprehensive database organizes this metadata, facilitating continuous tracking of learner performance and adaptive content delivery. As users advance, the algorithm incrementally increases task complexity to sustain progress and reinforce previous learning. This cyclical process of challenge and consolidation mirrors effective instructional design principles, ensuring both mastery and advancement.

Through the integration of adaptive algorithms, linked linguistic resources, and openly licensed encyclopedic content, this system exemplifies a holistic, data-driven approach to second language acquisition. It merges cognitive theory, pedagogical structure, and computational intelligence to create a personalized, self-regulated, and continuously evolving learning environment capable of meeting diverse learner needs.

3.2.1.8 Readability Index definition

To assess the analytic validity of a weighted lexical-semantic similarity index, we further implemented a computational pipeline for analyzing the linguistic characteristics of a multidisciplinary corpus composed of approximately 3,000 abstracts from the fields of Medicine, Second Language Acquisition, and Computer Science. The corpus was deliberately composed of texts extracted from open-access sources to ensure reproducibility and compliance with ethical data usage guidelines. The pipeline calculates the distribution of lexical items across three distinct categories: stopwords, words from a general-purpose metrics dictionary, and words from the AWL.

The selection process involved searching for relevant documents within Springer Nature's Open Access collection². Open Access works published by Springer Nature were prioritized, as they are

² 'Open access works published by Springer Nature are published under Creative Commons licences. These provide an industry-standard framework to support re-use of OA material.' <https://www.springernature.com/gp/open-science/about/the-fundamentals-of-open-access-and-open-research> (Accessed on 28/04/2025)

made available under Creative Commons licenses, supporting the reuse of the material with proper attribution³. The search for article abstracts was further refined to include results published between April 2015 and 2025, focusing on Open Access articles written in English. Open Access works published by Springer Nature, as they are made available under Creative Commons licenses, which offer a widely recognized framework for the reuse of OA content. We respectively applied keyword searches for the following areas: *Biomedicine*, with results filtered specifically for biomedicine; *Artificial Intelligence*, with results filtered for the Computer Science field and its general subfields; and *Second Language Acquisition*, with results filtered for Education. The search criteria also included Open Access articles published between April 2015 and 2025 and written in English.

This structured approach ensured the inclusion of open-access, peer-reviewed articles published between April 2015 to April 2025, covering a diverse range of subfields and research methodologies within each discipline. The set of abstracts processed consisted in 1584 valid and processable abstracts divided into three subgroups with the following distribution: Biomedicine 197, Computer Science 701 and Second Language Acquisition 686. It is important to note that all articles were subjected exclusively to linguistic preprocessing steps, such as tokenization, stopword removal, and stemming, without any alteration to their original content. The processing was designed solely to support lexical and semantic analysis, ensuring that the integrity and meaning of each abstract remained intact throughout the computational workflow.

Prior to semantic analysis, each abstract undergoes a standardized preprocessing pipeline designed to prepare the text for lexical and structural assessment. This process involves converting the input text to lowercase, followed by tokenization into individual word units. Crucially, stopwords, commonly occurring grammatical function words such as articles, prepositions, and conjunctions, are retained during this stage. Unlike traditional natural language processing workflows that often discard stopwords to focus on content-bearing vocabulary, this study preserves them in order to identify and quantify grammatical elements that contribute to the syntactic scaffolding of scientific writing. The inclusion of stopwords enables a fuller representation of linguistic structure and supports downstream analysis aimed at distinguishing grammatical words from technical and academic terminology. Additionally, each word is automatically stemmed using the Porter stemming algorithm (Porter, 2006; van Rijsbergen et al., 1980) to reduce inflected or derived words to their root forms, thereby facilitating more consistent matching across lexical resources.

To systematically assess the lexical and semantic characteristics of each abstract in the corpus, we implemented a computational pipeline that performs multi-level linguistic analysis. This pipeline

³ Creative Commons Attribution 4.0 International: <https://creativecommons.org/licenses/by/4.0/deed.en>

extracts and quantifies various textual features including lexical distribution, semantic similarity, and structural complexity. The following subsections describe the derivation and theoretical justification of each metric, followed by the integrated formulas used in the analysis.

The first stage involves computing lexical coverage proportions by classifying all tokens in the abstract into three distinct categories: stopwords, general-purpose vocabulary, and words from the AWL. This categorization allows for a fine-grained lexical profiling of the text. Given a total token count TT in the abstract, the relative proportion of each category is calculated as follows:

$$P_{\text{stopW}} = \frac{N_{\text{stopW}}}{T}, \quad P_{\text{general}} = \frac{N_{\text{general}}}{T}, \quad P_{\text{AWL}} = \frac{N_{\text{AWL}}}{T}$$

Formula 1 - Normalized frequency metrics representing the proportion of stop words, general vocabulary, and AWL terms within a given text corpus.

Here, N_{stopW} , N_{general} , and N_{AWL} represent the counts of stopwords, general-purpose words, and AWL words, respectively. This lexical categorization is essential not only for understanding the distribution of linguistic elements within a text, but also for weighing the contribution of each semantic domain in the subsequent analysis.

The second stage of the pipeline evaluates semantic similarity based on lexical stems and their associated similarity scores. These scores are retrieved from two database resources: a general-purpose metrics lexicon and a domain-specific AWL table, each containing precomputed similarity values for known word-stem pairs. For each abstract, the average semantic similarity is computed independently for the metrics and AWL vocabulary:

$$\text{Sim}_{\text{general}} = \begin{cases} \frac{1}{N'_{\text{general}}} \sum_{i=1}^{N'_{\text{general}}} S^{\text{general}}(w_i, s_i) & \text{if } N'_{\text{general}} > 0 \\ 0 & \text{otherwise} \end{cases}$$

Formula 2 Average semantic similarity between words w_i and stems s_i , using a general English vocabulary. Computed over general valid N matches; returns 0 if none found

$$\text{Sim}_{\text{AWL}} = \begin{cases} \frac{1}{N'_{\text{AWL}}} \sum_{i=1}^{N'_{\text{AWL}}} S^{\text{AWL}}(w_i, s_i) & \text{if } N'_{\text{AWL}} > 0 \\ 0 & \text{otherwise} \end{cases}$$

Formula 3 Same as above but based on the Academic Word List. Uses $S^{\text{AWL}}(w_i, s_i)$ over N'_{AWL} pairs; returns 0 if no valid matches.

To assess domain-specific vocabulary alignment, we compute separate semantic similarity scores for general ($\text{Sim}_{\text{general}}$) and academic (Sim_{AWL}) lexical domains. Each score is derived as the mean-similarity between words w_i and their stems s_i , using domain-specific semantic similarity functions $S_{\text{general}}(w_i, s_i)$ and $S_{\text{AWL}}(w_i, s_i)$, respectively. These are averaged over the number of valid matches N_{general} and N_{AWL} identified in the corresponding lexical resources. In these formulas, we calculate the mean semantic similarity between the original words w_i in the abstract and their corresponding stemmed forms s_i , where similarity scores $S_{\text{general}}(w_i, s_i)$ are computed with reference to a general English vocabulary database. In Formula 1, the term N'_{general} denotes the number of valid word-stem pairs identified in the general vocabulary set. In the absence of any valid pairs, the similarity is defined to be zero. Formula 2 mirrors the general similarity metric but relies on the AWL as the reference database. The resulting value captures the degree to which the vocabulary in the abstract aligns semantically with academic language, with N'_{AWL} representing the count of successfully matched word-stem pairs within the AWL.

Subsequently, to derive a single integrated measure of semantic similarity for each abstract, we calculate a weighted semantic similarity score. This score reflects the degree to which each vocabulary domain contributes to the text, based on its relative lexical density. The weighting ensures that semantic similarity scores have greater influence when their corresponding lexical category appears more frequently in the text. The weighted similarity index is thus formulated as:

$$\text{Sim}_{\text{weighted}} = \frac{(\text{Sim}_{\text{general}} \cdot P_{\text{general}}) + (\text{Sim}_{\text{AWL}} \cdot P_{\text{AWL}})}{P_{\text{general}} + P_{\text{AWL}}}$$

Formula 4 Computes a weighted similarity score by combining general and AWL vocabulary similarities, scaled by their respective lexical densities, to reflect the prominence and impact of different vocabulary types within the text.

In this expression, P_{general} and P_{AWL} refer to the relative lexical densities of the general and AWL word categories within the abstract, respectively. By weighing each similarity score by the frequency of its lexical category, this composite index ensures that categories with greater lexical presence have proportionally more influence on the final similarity measure. If no lexical items from either category are present, the similarity is again defined as zero. This formulation ensures a proportionally balanced score that integrates both general and academic semantic knowledge. The final value is bounded in the interval $[0,1]$, consistent with the normalized nature of the input similarity scores. In addition to lexical and semantic measures, the pipeline computes surface-level textual complexity metrics such as word density, unique word count, and sentence count. These indicators offer further insight into the structure and richness of the abstract. The formulas for these metrics are as follows:

$$D = \frac{T}{S}, \quad U = |\{w_i\}_{i=1}^T|$$

Formula 5 Calculates textual density (D) as the ratio of total words to sentences, and lexical uniqueness (U) as the count of distinct word forms, capturing structural and lexical diversity in the text.

where D is the average number of tokens per sentence, S is the total number of sentences, and U is the count of unique word forms (types) in the abstract.

Altogether, this computational framework yields a multi-dimensional profile of each abstract, capturing lexical richness, semantic informativeness, and syntactic complexity. These metrics can be used not only to assess individual texts but also to draw comparative conclusions across disciplinary domains or publication periods.

3.2.1.9 Network analysis

Building on the assessment of text complexity through interlanguage similarity and cognate transparency, a deeper understanding of lexical structures can be achieved through network analysis. By examining the structural relationships between words in both L1 and L2 lexicons, network analysis provides valuable insights into how vocabulary is organized and processed across languages. This approach allows for a more nuanced exploration of lexical connectivity, revealing patterns of word association, frequency distribution, and the influence of cognates on language comprehension. In our study, we adopt network analysis as a key methodological framework, utilizing tools from network science to investigate the structural properties of lexicons and their implications for language learning and processing.

In our study we strongly adopt the network analysis approach to the structural aspects of L1 and L2 lexicons by utilizing tools from network science. The cognitive revolution, which emerged in response to the behaviorist perspective, has significantly shaped the interdisciplinary field of cognitive science. This field centers on understanding how humans think by analyzing cognitive representations and structures, as well as the computational procedures that operate on these representations. Early work challenging behaviorist principles and focusing on the internal properties and structure of the mind includes George Miller's research on short-term memory (Miller, 1956), John McCarthy, Marvin Minsky, Allen Newell, and Herbert Simon's studies on artificial intelligence (Chrisley & Begeer, 2000; Simon & Newell, 1971), and Noam Chomsky's research on universal grammar (Chomsky & Skinner, 1959). These scholars laid the groundwork for new metaphors of the human mind, one of the most prominent being the idea of the mind as a computer, a symbol-processing, information-processing machine. This metaphor underpins computationalism approaches

to cognition, where information is symbolically represented, and cognitive processes are described through algorithms that operate on these symbols. Such an approach often depends on the implementation of cognitive models through computer programs.

This computational approach has been further integrated into more intricate systems, such as the mental lexicon, by conceptualizing them as networks. In these networks, individual words are represented as nodes, and their relationships are illustrated as edges. Network-based models, applied across various disciplines like biology, social sciences, and technology (Barabási, 2009; Baronchelli et al., 2013), have proven successful in capturing the complexities of cognitive structures. Networks are characterized by their local and global connectivity patterns, which significantly influence how information is transferred within them. For instance, random networks typically feature low local clustering (indicating sparse connections between neighboring nodes) and short global distances (meaning only a few steps are required to travel between distant nodes). In contrast, small-world networks are characterized by high local clustering and short global distances, enabling efficient communication at both local and global levels. These networks, often compared to random networks with identical node and edge counts, are marked by a relatively high clustering coefficient (CC) and a comparable average shortest path length (ASPL) to that of a random network. The CC measures the likelihood that two neighboring nodes are also connected, while the ASPL reflects the average number of steps required to connect two randomly selected nodes. Small-world networks optimize information exchange by balancing these two parameters.

In the domain of bilingualism, previous research applying this network framework has suggested that the L2 lexicon's structure is less intricate and sparser compared to the L1 lexicon (Wilks & Meara, 2002). However, the authors later reconsidered this view, acknowledging that their assumption of numerous direct word connections might have been overly simplistic, potentially undermining the robustness of their conclusions (Meara, 2009; Wilks et al., 2005). Building on this earlier work, the current study extends the field by incorporating alternative computational network techniques, focusing specifically on the small-world properties and community structures of both L1 and L2 lexical networks. This approach promises to provide deeper insights into the complexities of bilingual lexical organization, allowing for a more nuanced understanding of how these networks evolve and operate.

In this context, Wolter's (2001) *Depth Individual Word Knowledge Model* becomes a relevant framework for examining how word knowledge develops over time in bilingual speakers. Wolter's model emphasizes the depth of linguistic understanding, moving beyond the mere acquisition of vocabulary, and aligning with the notion that word knowledge in both L1 and L2 is not static but rather evolves in depth through experience, social interaction, and contextual exposure. Just as

bilingual lexicons are dynamic and complex, so too is the *space* in which words evolve and acquire deeper meanings. In the realm of second language acquisition (SLA), this model highlights the gradual deepening of word knowledge as learners interact with their linguistic environment, engaging with words in varying contexts and expanding their understanding over time.

By integrating the insights from both network-based studies of bilingualism and Wolter's model of word knowledge, the current research extends our understanding of bilingual lexical networks. It underscores the idea that L2 word knowledge is not merely a *sparser* or *simpler* version of L1 knowledge, but rather that both lexicons evolve through a process of deepening understanding shaped by cognitive, social, and cultural contexts. This expanded view allows for a more comprehensive understanding of how bilinguals navigate and develop their lexicons in both languages.

Wolter's (2001) *Depth Individual Word Knowledge Model* offers an insightful framework for understanding word knowledge, emphasizing the depth of linguistic understanding rather than the mere superficial acquisition of vocabulary. This model, rooted in linguistic and psychological theories, can be reinterpreted within the context of second language acquisition (SLA), where the focus is on how learners deepen their comprehension of words as dynamic entities over time.

In the context of SLA, Wolter's model aligns with the idea of *space* as a dynamic, evolving context in which word knowledge develops. Word knowledge is not simply about memorizing static definitions, but rather about engaging with words as living, evolving entities that are shaped by experience, social interaction, and the cultural context in which they are used. In this regard, the *space* of language acquisition can be seen as both a cognitive and social environment, where words transform, evolve, and acquire new meanings as learners interact with them in various contexts.

Wolter suggests that understanding words operates on different levels: from basic, surface-level knowledge, focused solely on definitions, to a deeper understanding that involves connotations, associations, and contextual meanings. In this model, words are not just static symbols; they evolve into more complex entities that are influenced by the learner's exposure to the language. This approach is particularly applicable to second language acquisition, where learners progress through stages of understanding as they encounter words in different contexts. At first, a learner may only grasp the basic definition of a word, but over time, they develop a richer, more nuanced understanding of the word as an entity with multiple facets, based on context, usage, and experience.

In second language acquisition, the learning process can be viewed as the movement through various *spaces* of meaning. Each word, treated as a dynamic entity, evolves as learners become more proficient, and as their exposure to the language broadens. These *spaces* are not only cognitive but also social, as learners engage with native speakers, texts, and cultural practices that shape the way words are used and understood. The interaction between language and its context, whether through

communication, reading, or media, contributes to the continuous evolution of a learner's word knowledge and their ability to use words more effectively in different situations.

Wolter's (2001) *Depth Individual Word Knowledge Model* provides a valuable framework for understanding the evolving nature of word knowledge. This model emphasizes not only the space in which word knowledge develops (the varying contexts in which a word is used) but also the time aspect, how word knowledge deepens and expands as learners encounter the word repeatedly across different temporal stages of learning. Unlike traditional views of word knowledge, which often focus on rote memorization of word definitions, Wolter's model underscores the dynamic, contextual, and cognitive elements that contribute to a richer understanding of words.

The diagram typically shows how learners' understanding progresses from surface-level recognition to a richer, more contextually nuanced comprehension, evolving over time as learners interact with words and engage with them in diverse contexts. This progression can be represented as a developmental pathway, where learners' knowledge of words expands in complexity and depth as they encounter them in new situations.

The first layer of the diagram often represents the basic or surface-level knowledge of a word. This would include the basic dictionary definition and common associations. At this stage, a learner may know how to recognize and use the word in common contexts, but they may not have a full understanding of its nuances or broader implications. For example, a learner might know the word 'dog' refers to a four-legged animal, but may not yet grasp the range of meanings and connotations the word carries, such as the emotional attachment people have to their pets or the cultural significance of dogs in various societies.

In the diagram, this stage may show a simple connection to a definition, perhaps without much further elaboration or connection to other words. This is the foundational stage where learners begin to build a basic framework for understanding, which will later be expanded.

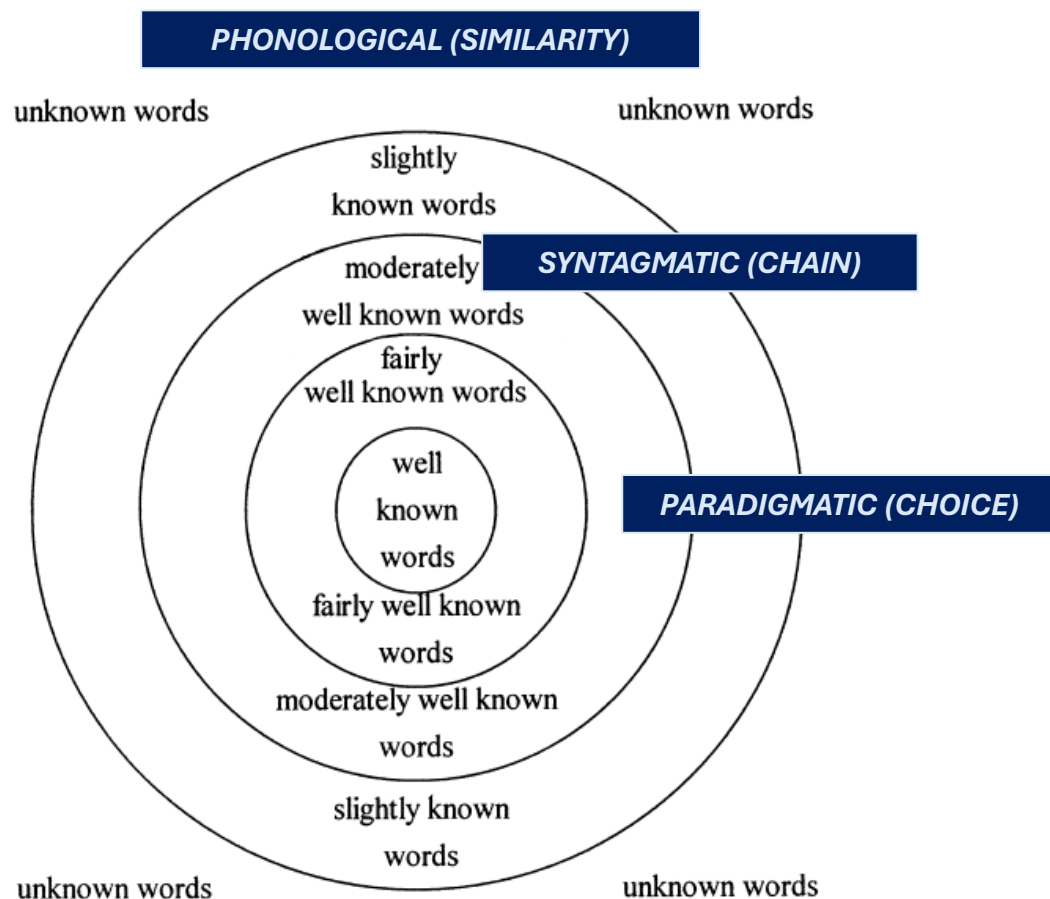


Figure 1 The Depth Individual Word Knowledge model Wolter's (2001)

Moving deeper into the diagram, we would likely encounter layers representing the intermediate stages of word knowledge. At this point, the learner begins to understand the connotations and associations that a word carries. These layers reflect how the word begins to take on meanings shaped by social, cultural, and experiential factors. For instance, the word *dog* might not just refer to the animal anymore, but also evoke ideas of loyalty, companionship, or even symbolism in art and literature.

These layers indicate how words become enriched with emotional reactions, symbolic meanings, and relationships to other words in the lexicon. The diagram may show the word branching out into different *spaces*, indicating the range of contexts in which the word can be used or understood differently. For example, 'dog' might be used in various contexts such as *dogged determination* (meaning persistence) or *every dog has its day* (meaning that everyone has their moment of success). At the deepest levels of the diagram, we find the richest and most nuanced understanding of a word. This includes full appreciation of how a word functions in different contexts, how it relates to other words in subtle or complex ways, and how it is integrated into the larger cognitive and cultural system of language. In the case of 'dog,' a learner with deep knowledge might recognize how the word can vary depending on cultural perceptions, historical significance, and even the way it plays a role in

different social settings (e.g., *therapy animals, service dogs, or the various breeds and their characteristics*).

In the diagram, this could be represented by extensive network connections to various contexts, examples, and even related concepts, illustrating the word's multifaceted meanings and uses. This stage demonstrates how words are linked to broader cognitive and social networks, encompassing a wide range of possible interpretations and implications based on individual experiences and cultural frameworks.

One of the most important elements of Wolter's model is that word knowledge is not static. The diagram may illustrate the dynamic process of acquisition, with arrows or paths showing how learners move from a basic, superficial understanding to a deeper, more integrated knowledge of words. The model may also represent this progression as a circular or iterative process, indicating that word knowledge can evolve as learners interact more with language and use words in diverse contexts.

This process is iterative in nature; learners do not necessarily progress in a linear fashion from surface-level knowledge to deep knowledge. Instead, as they are exposed to new contexts and usage, their understanding of words can be revised or refined. Moreover, as they encounter variations of the same word across different situations, they continuously reframe and adapt their understanding of its meaning, which reflects the evolving nature of language acquisition itself.

In the context of second language acquisition (SLA), this diagram can highlight the difference between L1 and L2 word knowledge. The structure of the L2 lexicon may initially appear sparser or less interconnected compared to the L1 lexicon, as learners have less experience with the word in various contexts and may lack the cultural and social knowledge associated with the word. However, over time, as learners engage with the L2 in meaningful ways, whether through immersive experiences, social interaction, reading, or media, they will build deeper and more complex networks of word knowledge.

This idea of deepening understanding can be represented by expanding connections or layers in the diagram, illustrating how words are understood at various depths across both languages. For example, in SLA, learners might initially know the word 'car' in English to refer to a vehicle, but as they engage more with the culture and language, they might learn connotations such as *road trip, car culture*, or the symbolic role of cars in different societies.

Moreover, the L2 learner's path may not mirror the L1 acquisition process in every way. For example, the learner's first encounters with a word may be highly influenced by their L1 vocabulary and concepts, which can sometimes lead to interference or overgeneralization. However, as the learner's L2 knowledge deepens, these words start to acquire more nuanced meanings and connections that are more aligned with the L2's linguistic and cultural context.

Wolter's *Depth Individual Word Knowledge Model* offers a valuable perspective for understanding the process of second language acquisition. It highlights that word knowledge is not a static entity but a dynamic, evolving process, where words deepen in meaning through experience, context, and interaction. By treating words as entities that evolve in depth and complexity, this model underscores the importance of context in language learning and offers a more holistic view of how learners acquire, process, and understand new words over time. It also emphasizes that word knowledge, particularly in SLA, is not confined to memorization but evolves in parallel with real-world usage and cultural understanding, offering a richer and more adaptable language-learning experience.

In our learning application, we take special care to address both the *time* and *space* dimensions of word knowledge as outlined in Wolter's (2001) *Depth Individual Word Knowledge Model*. Understanding that word knowledge evolves over time, we have structured our application to guide learners through various stages of word acquisition, ensuring that they gain both a broad understanding of a word's basic meaning and a deep, nuanced comprehension as they continue to interact with it.

The time aspect of learning is central to our approach. We recognize that learners do not acquire word knowledge in a single moment but instead progress through distinct stages. Our application incorporates gradual progression, where learners initially encounter words at a surface level, focusing on basic definitions and simple usage. As learners move through the stages, the application provides opportunities to revisit and engage with words in more complex contexts.

This approach allows learners to deepen their understanding over time, similar to how Wolter's model shows how word knowledge expands as learners interact with words across different stages of their learning journey. Each stage in our application is designed to reinforce and build upon previous encounters, ensuring that learners' word knowledge evolves iteratively, becoming more sophisticated with each exposure.

In addition to the time aspect, our application emphasizes the *depth* of word knowledge, how a learner's understanding grows in complexity as they encounter new contexts, connotations, and associations of words. As learners progress, the application introduces them to words in increasingly varied and meaningful contexts, allowing them to develop a deeper, richer understanding.

For example, a learner who starts with a basic understanding of a word like *run* will eventually encounter its many connotations, like *run a business*, *run for office*, or *run errands*. They will also explore how the word is used in idiomatic expressions, cultural references, and specialized fields. This ensures that, by the time they reach deeper levels of understanding, learners have a broad and nuanced comprehension of the word, as represented in Wolter's model by extensive network connections to multiple contexts.

By carefully addressing both the *time* and *space* aspects of word knowledge, our application creates a dynamic learning experience. Learners are not only exposed to words in isolation but are guided through a holistic process that mirrors the real-world progression of language acquisition. Just as Wolter's model illustrates the evolving nature of word knowledge, our approach allows learners to continuously refine and deepen their understanding, making the acquisition of new words both meaningful and enduring.

To track learners' progress and ensure that their knowledge deepens over time, our application incorporates a combination of metrics and feedback as part of the assessment process. These assessments are designed to reflect both the *depth* and *time* aspects of word knowledge, allowing us to measure how learners' understanding of words evolves and deepens through the stages of learning. The metrics offer quantitative data on learners' progress, such as the accuracy of word usage, the range of contexts in which words are applied, and the frequency of interactions with specific words. These metrics help us assess whether learners are advancing through the stages of knowledge, from surface-level recognition to a deeper, more integrated understanding.

In addition to metrics, learners receive timely and personalized feedback. This feedback serves as a guide, helping learners understand where they might need to focus their attention, offering insights into the nuances of word meanings, and reinforcing how the words they are learning function in different contexts. The feedback provides a qualitative layer to the assessment, ensuring learners are not only aware of their progress but also informed on how to continue deepening their knowledge.

We will describe these assessment processes in greater depth in the following sections, highlighting how the combination of metrics and feedback offers a comprehensive and dynamic approach to learning, supporting learners as they progress through the stages of word knowledge.

3.2.1.10 Dynamic exercise generation

In our application, the dynamic exercise generation feature is designed to continuously enhance learners' practice and consolidate language concepts. Learners can access a dedicated section with exercises directly tied to the topics covered in their lessons, offering them the flexibility to engage with exercises as often as necessary until they feel confident and ready to progress. This feature ensures that learning is not limited to passive exposure; it promotes active, personalized practice tailored to each learner's evolving needs.

As learners move through the content, the system adapts to their progress by dynamically modifying exercises based on the texts and word units they interact with. These personalized activities include flashcard games, cloze exercises, memory association games, crosswords, and even the Game of the Goose. By leveraging rich data from learners' previous choices and interactions, the system creates

practice opportunities that remain relevant and engaging, providing a responsive learning experience that supports continuous skill development and reinforces language acquisition.

In our view, incorporating content that learners are already familiar with can significantly lower the barriers to understanding and increase their engagement with the material. By integrating topics and themes learners are interested in or already have background knowledge of, the dynamic exercise generation system bridges the gap between new content and existing knowledge. This approach not only fosters a more meaningful learning experience but also allows learners to focus more on the nuances of language, such as grammar and vocabulary, rather than struggling with entirely new concepts.

When learners encounter familiar contexts or subjects, they can build on their prior knowledge, reducing cognitive load and making it easier to absorb new language forms. This creates an opportunity for learners to develop a deeper understanding of language usage and application, improving their overall proficiency and confidence. Engaging with material that resonates with their interests boosts motivation, increases the likelihood of sustained engagement, and contributes to more effective language acquisition. By reducing the complexity of new content through familiar topics, our system encourages learners to explore language forms and structures in more depth, ultimately fostering more robust and confident language use.

Additionally, our approach aligns with the Affective Filter Hypothesis, which posits that emotional factors significantly impact language learning success. According to this hypothesis, learners with a low affective filter, characterized by positive emotions such as motivation, interest, and confidence, are more likely to acquire language effectively. By integrating familiar and personally engaging content into our dynamic exercises, we aim to create a positive emotional environment that lowers the affective filter.

This strategy reduces anxiety and builds learners' confidence, allowing them to approach language tasks with a more relaxed and receptive mindset. As a result, learners are better positioned to absorb and internalize new language skills. By providing an environment that supports motivation and emotional engagement, we facilitate a more enjoyable and effective learning experience, ultimately enabling learners to progress more successfully in their language acquisition.

In our application, we ensure that exercises are not only contextually relevant but also tailored to reinforce the specific lexical core that learners engage with during their exploration of freely selected materials. By aligning exercises with the vocabulary and themes present in the learner's chosen texts, the system creates a cohesive learning experience that enhances both comprehension and retention.

When learners select texts that spark their interest, these texts serve as the foundation for their initial lexical exposure. The system then uses the lexical core, words and phrases from these texts, to

generate targeted exercises that reinforce these terms in varied contexts. This ensures that learners encounter and practice the same vocabulary multiple times through different exercise formats, such as cloze tests, flashcards, and contextual games. This repeated exposure helps solidify learners' ability to recognize and use these words in meaningful contexts, promoting deeper lexical acquisition.

The iterative engagement with the same lexical core across diverse exercise formats promotes retrieval practice, which is essential for long-term retention. By encountering the same vocabulary in multiple contexts, learners are better able to recall and use words meaningfully, strengthening both their comprehension and production skills. The system's adaptive nature ensures that exercises evolve in response to learners' progress, refining the lexical focus as they continue to improve. This continuous cycle of review and reinforcement ensures that learners internalize the language and can apply it flexibly in real-world situations.

This approach is grounded in principles from lexical processing theory, which highlights the importance of repeated exposure and context-based usage in vocabulary acquisition. By providing learners with numerous opportunities to engage with and apply key vocabulary, our system fosters a robust understanding of how words function in various contexts. This approach not only enhances lexical retention but also supports learners in transferring their acquired vocabulary into practical language use, contributing to their overall linguistic proficiency and confidence.

One key element of our application is the integration of gamification, which has been shown to significantly enhance vocabulary acquisition and retention. According to Lee (2022), game-enhanced learning can serve as an effective alternative to traditional vocabulary teaching methods. The study found that learners' interest in and attention to the game content played a crucial role in boosting vocabulary outcomes. Games that engage students with their content lead to more positive learning outcomes, reinforcing the effectiveness of gamified experiences in language acquisition.

In our application, the dynamic exercises, such as flashcard games, memory association games, and contextual games like the Game of the Goose, encourage learners to interact with vocabulary in a fun and engaging way. These game-like activities hold learners' attention and foster a deeper connection to the words they are learning, which in turn supports both their vocabulary retention and ability to use words meaningfully in different contexts. The incorporation of gamified elements not only enhances the learning experience but also motivates learners by making vocabulary practice enjoyable and rewarding.

Lee (2022) highlights the importance of attention and salience in vocabulary learning. Words that stand out due to their context, repetition, or emotional engagement have a greater impact on learners' vocabulary retention. Our application integrates this concept by dynamically adjusting the difficulty and variety of exercises based on each learner's progress, ensuring that the vocabulary they are

exposed to remains engaging and contextually relevant. By incorporating game mechanics and interactive formats, learners are encouraged to engage with words in ways that increase their salience, thus maximizing their vocabulary acquisition.

Additionally, the system's ability to generate exercises related to themes and topics that learners are already familiar with ensures that the words encountered are not only engaging but also highly relevant to the learner's existing knowledge. This fosters an environment where learners can connect new vocabulary to personal experiences and interests, further enhancing salience and retention.

Supporting Motivation through the Affective Filter Hypothesis:

Our approach also aligns with the Affective Filter Hypothesis, which posits that emotional factors significantly impact language learning success. As Lee (2022) suggests, learners are more likely to acquire language effectively when they are motivated, engaged, and have a positive emotional connection to the content. The gamified elements in our system are designed to create an environment that lowers the affective filter by making learning enjoyable and reducing anxiety. By encouraging positive emotions such as motivation and satisfaction, our system ensures that learners are more receptive to language tasks and more likely to internalize new vocabulary.

By combining gamification with a focus on engagement and salience, our application not only supports the acquisition of vocabulary but also fosters a learning environment where learners are motivated to continuously improve. The dynamic exercise generation feature, along with game-based activities, ensures that vocabulary practice remains both effective and enjoyable, contributing to long-term retention and the development of language proficiency.

3.3 Introduction to Computer Science Methodology

The translation of the didactic framework into a functioning web-based learning environment required a comprehensive, multi-phase development approach, carefully balancing pedagogical requirements with robust technical solutions. While the previous chapters examined the theoretical and cognitive foundations of the project, this section turns to the technological decisions and implementation processes that transformed those principles into a responsive and scalable platform.

Building upon this foundation, a key innovation of the system lies in the internal representation of linguistic knowledge. While data integration and adaptive algorithms enable the delivery of content tailored to individual learner profiles, the deeper challenge involves structuring and interpreting meaning in a way that reflects the complexity of human language. To address this, the system employs network-based models of linguistic knowledge representation, in which lexical and semantic relationships are captured and analyzed through interconnected structures. This approach allows for

a more dynamic understanding of how words, phrases, and concepts interact, supporting richer adaptivity and more meaningful feedback in the learning process.

3.3.1 Knowledge Graphs in Knowledge Representation and Reasoning

In the field of knowledge representation and reasoning, a knowledge graph is a specialized type of knowledge base that employs a graph-structured data model to represent and manipulate data. These graphs capture the semantics, or relationships, between entities, which can encompass objects, events, situations, or abstract concepts. By establishing a network of interconnected entities, knowledge graphs create a framework that not only links descriptions of real-world phenomena but also enables reasoning and the derivation of new insights from these relationships. At the core of a knowledge graph, nodes represent the entities of interest, such as objects, events, concepts, or situations, while edges encode the relationships that connect these entities. The relationships between the entities are typically labeled to specify their nature, such as 'is a part of' or 'caused by' which defines the semantic connections in a graph. One of the distinctive features of knowledge graphs is their ability to serve as a shared substrate, a continuously evolving foundation that provides a common reference framework for integrating data and supporting reasoning across various applications and domains. In this way, they function similarly to the mental lexicon in human cognition, where the brain organizes words, concepts, and their semantic connections into a network of interrelated nodes, facilitating the retrieval and combination of related information. Just as the mental lexicon holds words and their meanings, a knowledge graph holds entities and their relationships, enabling efficient retrieval and comprehension. This approach to representing knowledge has similarities with the methodology of Network Science, which uses a graph structure to analyze how nodes (entities) are interconnected through edges (relationships) (Borgatti, Everett, & Johnson, 2013). The network science approach has been extensively applied in diverse fields, including social media networks (Chen & Pan, 2009), epidemic spreading (Cui & Li, 2020), friendship networks (McKay, Grygiel, & Karwowski, 2017; Ivaniushina & Alexandrov, 2018), and information-sharing networks (Pooranachandran & Balasubramanian, 2010). Similarly, knowledge graphs function as dynamic, evolving structures that continuously update, refine, and expand to incorporate new knowledge, akin to how the brain's mental lexicon changes over time through experience and exposure to new information (Molano & Polo, 2015; Ergün & Usluel, 2016).

Knowledge graphs also mirror the structure of the mental lexicon by offering a rich representation of the interrelatedness of concepts, where associations between nodes reflect semantic and syntactic relationships that can be exploited for tasks like disambiguation, contextualization, and inference. The continuous evolution of knowledge graphs involves several key processes: extraction, which

gathers information from various sources; enrichment, which adds context and detail to enhance the data's value; quality assessment, ensuring accuracy and consistency; and refinement, which updates and corrects the graph to reflect new information. These processes align with how the brain constantly updates and refines the mental lexicon through experience and exposure to new information. Knowledge graphs are dynamic structures that grow over time, much like how the mental lexicon evolves as we learn new words and meanings, and as we encounter novel associations between concepts. The development of extended knowledge graphs introduces new capabilities by integrating diverse data types, including textual content from web sources, images, audio, video, and sensor data, capturing a more holistic view of the world. This evolution mirrors the way the human mental lexicon adapts to include not only linguistic information but also multimodal knowledge, allowing individuals to associate words with images, sounds, and sensory experiences. Knowledge graphs also reflect the principles of network science by representing the interrelatedness of concepts. In these graphs, nodes and their relationships can be described using parameters such as degree, clustering coefficient, and centralization, allowing for detailed analysis of how entities are interconnected (Scott, 2000). The continuous evolution of knowledge graphs involves several key processes: extraction, which gathers information from various sources; enrichment, which adds context and detail to enhance the data's value; quality assessment, ensuring accuracy and consistency; and refinement, which updates and corrects the graph to reflect new information. These processes align with how the brain constantly updates and refines the mental lexicon.

Extended knowledge graphs integrate diverse data types, including textual content, images, audio, video, and sensor data, offering a more holistic view of the world. This integration mirrors how the human mental lexicon not only encompasses linguistic knowledge but also includes multimodal experiences, allowing individuals to associate words with images, sounds, and other sensory data.

Additionally, extended knowledge graphs introduce reasoning capabilities that enable the generation of new knowledge from existing relationships, improving their ability to answer complex queries and support decision-making, much like how human cognition can infer meanings and relationships beyond direct experience. In the Semantic Web, ontologies play a foundational role in representing knowledge in a machine-readable format. These formal descriptions of domain concepts and their relationships enable interoperability between disparate datasets, providing a shared understanding of data semantics.

In the Semantic Web, ontologies play a foundational role in representing knowledge in a machine-readable format. These formal descriptions of domain concepts and their relationships enable interoperability between disparate datasets, providing a shared understanding of data semantics. Similar to how linguistic structures and word meanings are encoded in the mental lexicon, ontologies

use standardized languages like the Web Ontology Language (OWL) to define classes, properties, and axioms, supporting a web of interconnected and meaningful information. Ontologies thus help facilitate the vision of a web of data that mirrors the human ability to understand and navigate complex networks of knowledge through semantic connections and inferential reasoning.

Knowledge Graphs (KGs) have emerged as powerful tools for structuring and linking information in ways that closely mirror human cognition. In language learning, they offer a promising approach to enhancing vocabulary acquisition, improving contextual understanding, and resolving semantic ambiguities. By explicitly representing semantic relationships between words, KGs enable learners to explore connections among synonyms, antonyms, hypernyms, and hyponyms. This interconnected structure not only aids in vocabulary retention but also enhances learners' ability to use words accurately in different contexts, fostering a richer and more versatile lexical framework.

When enriched with multimodal data such as text, images, and audio, KGs provide a more comprehensive learning experience. This multimodal approach integrates example sentences, visual content, and real-world usage scenarios, promoting a deeper understanding of words in varied contexts. Additionally, KGs support adaptive learning by personalizing content based on a learner's proficiency level, tailoring recommendations, grammar rules, and practice exercises accordingly.

For semantic disambiguation, KGs address polysemy by mapping words to their specific meanings based on context, which is especially beneficial in machine translation and other language-learning applications. They also contribute to the development of automated assessment tools that analyze learner responses, identify errors, and provide meaningful feedback.

The hierarchical structure proposed by Collins and Quillian (1969) represents concepts as nodes in a tree-like network, with connections defined by class-inclusion relationships. This model effectively reflects aspects of human semantic memory, illustrating the connection between general and specific concepts. However, it has limitations, particularly when dealing with concepts that do not fit neatly into taxonomic categories.

To address these limitations, Steyvers and Tenenbaum (2005) introduced an alternative model that emphasizes the dynamic process of concept acquisition and network growth. In this model, new concepts are linked to a subset of existing ones within a predefined neighborhood, with the probability of connection proportional to the neighborhood's size. Densely connected areas are more likely to expand, leading to scale-free small-world networks characterized by hubs with dense interconnections and efficient pathways between nodes.

This growth process can be seen as semantic differentiation, where new concepts emerge as specific variations of existing ideas. Concepts with greater connectivity tend to evolve and diversify over time. The model also predicts the timeline of semantic acquisition, suggesting that earlier-learned concepts

exhibit higher connectivity, a pattern supported by age-of-acquisition norms (Gilhooly & Logie, 1980; Morrison, Chappell, & Ellis, 1997). This insight helps explain behavioral phenomena in lexical decisions and naming tasks, particularly those influenced by the age of acquisition.

Integrating this model into language learning through KGs enhances our understanding of how concepts evolve and how acquisition order and complexity shape language comprehension. This framework supports the structured growth of a learner's semantic network, ultimately improving personalized learning experiences.

Network analysis has increasingly been used to examine the lexicon as a complex system, providing insights into the structural organization and interconnectivity of lexical items. This methodology systematically investigates how nodes, representing words, concepts, or linguistic units, are interconnected through various edges capturing relationships such as semantic similarity, co-occurrence, and syntactic dependencies (Borgatti, Everett, & Johnson, 2013). Beyond linguistics, network science has been applied in fields such as social media analysis (Chen & Pan, 2009), epidemic modeling (Cui & Li, 2020), friendship network studies (McKay, Grygiel, & Karwowski, 2017; Ivaniushina & Alexandrov, 2018), and information-sharing structures (Pooranachandran & Balasubramanian, 2010). In language acquisition research, complex network methodologies have been used to examine student interactions within learning communities (Molano & Polo, 2015; Ergün & Usluel, 2016).

Network science provides a systematic framework for analyzing the topological properties of intricate systems by evaluating relationships among nodes. A system is conceptualized as a graph, where nodes represent discrete entities, such as words in a lexical network or individuals in a social network, and edges denote their connections. The structural properties of these networks are characterized by well-defined parameters that describe connectivity patterns (Yan & Zhang, 2007). Key metrics in network analysis include:

Metric	Description
Size (n)	The total number of nodes within the network.
Degree (k)	The number of direct connections a node possesses.
Average degree ($\langle k \rangle$)	The mean number of connections per node.
Density (D)	The ratio of observed edges to the total possible edges.
Clustering Coefficient (CC)	The probability that two nodes connected to a common node are also directly connected, indicating tightly interlinked subgroups.

Metric	Description
Average shortest path length (L)	The mean shortest distance between all pairs of nodes.
Centralization (C)	A measure of the extent to which a network is structured around central nodes, with higher values indicating greater concentration of edges around specific nodes (Scott, 2000).

Table 2 Knowledge Graphs Key Network Metrics and Their Definitions

By applying network science to linguistic and lexical studies, researchers can uncover structural patterns, enhance models of language acquisition, and deepen our understanding of how lexical units interact within a language system. This methodological approach offers a powerful tool for analyzing both linguistic networks and broader complex systems, paving the way for more sophisticated models of language learning and cognitive representation.

3.3.2 Ontologies as the Backbone of Linked Open Data in Linguistics

Linked Open Data (LOD) in linguistics is intricately structured around ontologies, which function as conceptual frameworks that underpin the organization, integration, and dissemination of linguistic data. These ontologies define entities, attributes, and relationships among linguistic concepts, ensuring consistent and interoperable representation of linguistic knowledge across diverse datasets. By providing structured vocabulary, they enable seamless linking and integration of linguistic data from various sources, creating an interconnected web of linguistic resources that can be explored from multiple perspectives and across languages. This structured approach is not merely a technical necessity but a foundational element in advancing linguistic research, natural language processing (NLP), and cross-linguistic studies.

As Hirst (2009) highlights, ontologies standardize the representation of linguistic concepts, facilitating the alignment and merging of data from disparate sources. This standardization is critical in the development of comprehensive linguistic repositories such as WordNet, *BabelNet*, and *DBnary*, which serve as essential resources for both theoretical linguistics and computational applications. Without ontologies, establishing common ground among these vast linguistic databases would be challenging, leading to fragmentation and limiting the potential for broader applications of linguistic data. Ontologies act as bridges, ensuring that linguistic knowledge is not siloed but rather interconnected and accessible across different platforms and disciplines.

Beyond mere data organization, ontologies empower sophisticated querying and reasoning over linguistic data. By defining essential semantic relationships, such as synonym, antonymy, hypernymy,

and hyponymy, they create structured networks of meaning, allowing users to navigate complex linguistic landscapes with greater ease. These relationships extend beyond single-language corpora, forming multilingual mappings that enhance translation accuracy, cross-linguistic research, and knowledge extraction. Through these structured semantic connections, ontologies provide a dynamic mechanism for exploring language in a way that mirrors human cognitive processes, facilitating a deeper and more nuanced understanding of meaning across languages and contexts.

Ontologies in LOD also play a pivotal role in the realization of the Semantic Web, where data is not only linked but also enriched with meaning through semantic annotations (Berners-Lee et al., 2021). This semantic layer enables machines to process and interpret linguistic concepts more intelligently, vastly improving applications in NLP, machine translation, and automated text analysis. By embedding deeper contextual understanding into linguistic data, ontologies enhance the capabilities of AI-driven systems, allowing for more precise language modeling, better disambiguation of polysemous terms, and improved contextual inference in machine learning applications.

Recent developments in LOD, such as the Linked Data models for lexicography (e.g. *OntoLex-Lemon*), underscore the indispensable role of ontologies in linking linguistic data with other knowledge domains (McCrae et al., 2012). These models provide a robust framework for representing multilingual lexical data in ways that are both interoperable and extensible, ensuring that linguistic datasets can be seamlessly connected with other datasets on the web. This cross-domain connectivity is particularly significant as language is deeply embedded in various aspects of human knowledge, from historical and cultural studies to artificial intelligence and education. By facilitating these connections, ontologies enrich linguistic data while simultaneously expanding the broader web of knowledge.

The integration of ontologies in LOD extends beyond the boundaries of linguistics. As linked data initiatives continue to evolve, linguistic ontologies increasingly serve as crucial junctions between language and other domains, linking linguistic data with geographical, cultural, and social datasets. This cross-disciplinary enrichment has far-reaching implications, supporting diverse real-world applications such as intelligent tutoring systems, automated translation services, and data-driven cultural analysis. The ability to interconnect linguistic resources with broader knowledge structures ensures that language-related data is not only more accessible but also more contextually relevant in interdisciplinary research and real-world applications.

Ontologies serve as the structural foundation of Linked Open Data in linguistics, enabling the organization, integration, and reasoning over linguistic knowledge. By standardizing the representation of linguistic concepts and their interrelations, they foster a rich and interconnected linguistic web that enhances both research and practical applications across multiple fields. This

interconnectedness is not only vital for academic inquiry but also for advancing real-world technologies, from AI-driven language processing to multilingual information retrieval. Moreover, ontologies represent the crucial contact point between raw data and structured knowledge, transforming fragmented linguistic information into a coherent, semantically enriched system that machines and humans can effectively interpret. As the Semantic Web continues to evolve, ontologies will remain essential in ensuring that linguistic data is not just linked but meaningfully structured, ultimately positioning it as a key pillar in a more semantically aware and interconnected digital world.

3.3.3 Network analysis in linguistic knowledge representation

One of the most compelling innovations in contemporary approaches to linguistic modeling is the use of network analysis for knowledge representation. Language is inherently relational: words derive much of their meaning from the connections they maintain with other words, concepts, and contexts. Traditional models of word sense disambiguation (WSD), which rely on statistical co-occurrence or dictionary definitions, often struggle to capture this relational complexity. Network-based models, by contrast, transform text into semantic knowledge graphs that make explicit the intricate web of relationships that structure language.

In such a graph-based approach, words are not treated as isolated surface forms but are instead represented as nodes linked to their possible senses and semantic relations. These relations, many of which can be derived from resources such as WordNet, include synonymy, hyponymy, hypernymy, meronymy, holonymy, antonymy, and broader associative links. By interconnecting words and senses in this way, the graph produces a multi-layered semantic space in which meaning emerges from patterns of relationships rather than from single, context-free entries. For example, a lexical item may connect both to its synonyms, which indicate meaning similarity, and to its hypernyms, which place it within a larger conceptual category.

The advantages of this network perspective are significant. By mapping meaning relationally, semantic graphs facilitate context-sensitive disambiguation, as the surrounding connections help determine which sense of a polysemous word is most appropriate in a given context. Unlike models that treat disambiguation as a choice between discrete dictionary entries, a network approach enables a dynamic representation in which meaning is continuously adjusted according to context. This makes the model particularly well-suited for language learning, where learners must not only decode unfamiliar words but also situate them within larger conceptual frameworks.

Moreover, knowledge graphs can be conceived not only as computational models but also as cognitive analogues of how humans' structure lexical information. Research in psycholinguistics and cognitive science suggests that the mental lexicon operates through associative networks, with activation

spreading across related concepts during comprehension and production. By mirroring this structure, graph-based models resonate with theories of the mental lexicon, providing both a descriptive and explanatory account of how words relate to one another.

Beyond the immediate task of WSD, network analysis has broader implications for linguistic inquiry and pedagogy. As a visual and relational map of word meanings, knowledge graphs offer an accessible way to explore how words function in relation to one another and how their meanings evolve across different contexts. For learners, this means language can be approached not simply as a linear sequence of rules but as a constellation of interconnected meanings. For researchers, network structures open new possibilities for investigating phenomena such as semantic density, polysemy, collocational strength, and even diachronic change.

Finally, the complementarity of different graph types enhances the richness of representation. On the one hand, contextual graphs, built from dependency parsing and semantic role labeling, capture text-specific relationships (such as subject–predicate–object links). On the other hand, lexical graphs, constructed from resources like WordNet, anchor words to more general semantic relations that transcend individual texts. Together, these perspectives provide both micro- and macro-level analyses of meaning, situating local textual relationships within the broader architecture of the lexicon.

In sum, network analysis provides a powerful theoretical framework for representing linguistic knowledge. By rendering meaning as a structure of interdependent relations, semantic graphs not only advance computational models of language but also resonate with cognitive theories of the mental lexicon and open new avenues for pedagogy. For second language learners in particular, the ability to visualize and interact with these relationships offers a promising pathway to deeper comprehension and more robust vocabulary acquisition.

3.3.4 Integrating Open Data and AI for Educational Systems

The translation of the didactic framework into an operational platform required not only robust architectural and programming choices but also a systematic approach to data management and integration. Without carefully selected and curated resources, no digital system can provide meaningful input or reliable feedback. Consequently, the second stage of implementation was devoted to the identification, evaluation, and integration of linguistic and textual data repositories, ensuring that the resulting platform rested on a foundation that was both pedagogically sound and technically sustainable.

From the outset, it was evident that resource acquisition posed both technical and didactic challenges. While textbooks and traditional curricula provide vocabulary lists, glossaries, and structured input, a

digital environment demands corpora, lexical databases, and linked resources that can be dynamically processed. The integration process therefore focused on four essential categories:

Resource Type	Description	Pedagogical Function
Linguistic Repositories	Contain annotated corpora and structured data for syntax, morphology, and semantics, providing the foundation for automated linguistic analysis.	Enable the generation of personalized feedback and adaptive exercises informed by learner performance.
Textual Collections	Include graded materials ranging from simple dialogues and narratives for beginners to domain-specific texts for advanced learners.	Support CEFR-aligned progression and comprehensible input tailored to proficiency level.
Vocabularies and Thesauri	Integrate open dictionaries, lexical databases, and semantic networks to represent word meanings and relationships.	Facilitate contextualized learning, thematic expansion, and precise lexical profiling.
Didactic Instruments	Incorporate open-access exercises, multimedia materials, and instructional frameworks aligned with current SLA methodologies.	Ensure pedagogical coherence with SLA principles and contemporary teaching approaches.

Table 3 Core resource components supporting adaptive language learning systems and their pedagogical functions.

This integration effort was shaped by a commitment to open-access resources. Many high-quality linguistic datasets remain proprietary, either hidden behind paywalls or restricted by licensing terms. For a scalable and sustainable educational system, reliance on openly licensed or public-domain resources was therefore a non-negotiable requirement. Although this constraint significantly narrowed the available resource pool, it also encouraged innovation in preprocessing and harmonization. Resources were normalized and unified into a single framework that could be dynamically queried by the system. For example, corpora were segmented and annotated according to CEFR proficiency levels, while thesauri were restructured to highlight semantic and morphological relationships relevant to language learning.

To enhance interoperability across heterogeneous datasets, the project explored the use of Linked Open Data (LOD). By adopting open standards such as OntoLex and leveraging knowledge graphs including WordNet, *BabelNet*, and *ConceptNet*, the backend was designed to interconnect disparate

data sources in meaningful ways. This approach aligns with broader European Union initiatives promoting open data governance, transparency, and reusability. In practice, it allowed the system to integrate lexical, semantic, and pragmatic information from multiple sources while maintaining structural and ontological coherence, an essential step for providing accurate, context-sensitive feedback.

In parallel with data integration, the implementation also engaged with artificial intelligence (AI) technologies. Rather than using AI as a generative mechanism, it was employed selectively for analysis and adaptation. Natural Language Processing (NLP) modules, implemented in Python (Python Software Foundation, 2022) and connected to the backend through RESTful APIs (Fielding, 2000; Pautasso, 2021; Richardson et al., 2013), managed key computational tasks such as tokenization, lemmatization, disambiguation, and named-entity recognition. These processes enabled real-time monitoring of learner input, alignment with reference corpora, and the provision of personalized linguistic support.

At the same time, the integration of AI was approached critically, with awareness of the risks inherent in automated language technologies. Bias in training data, underrepresentation of dialects, and uneven treatment of non-standard varieties can all lead to inequitable outcomes. To mitigate these issues, AI components were grounded in curated, diverse datasets, with transparency, inclusivity, and fairness serving as guiding principles.

Ultimately, this phase of implementation ensured that the learning platform was not merely an empty framework, but a comprehensive learning environment enriched with high-quality linguistic data and intelligent processing capabilities. The emphasis on open-access resources guarantees scalability, accessibility, and sustainability, while the incorporation of Linked Open Data and AI components provides adaptability and long-term growth potential. By intertwining didactic priorities, legal compliance, and computational innovation, the data integration process established a foundation for an adaptive language learning environment that can evolve alongside its users and remain aligned with the most recent advances in SLA research and educational technology.

3.3.4.1 Integration of Linked Open Data and Linguistic Linked Open Data in Educational Systems

In the modern digital landscape, the demand for robust, interoperable, and openly accessible data has become increasingly critical for both research and education. The Linked Open Data (LOD) movement represents a paradigm shift in how information is structured, shared, and connected across domains. Built on the principles of openness and interoperability, LOD employs the Resource

Description Framework (RDF) to encode information in machine-readable triples that explicitly define relationships between entities. This relational model enables not only the storage of information but also its integration into a global knowledge network, where datasets from diverse sources can be federated, queried, and recombined to generate new insights.

The value of LOD lies in its capacity to guarantee openness, availability, and interlinking, qualities that are particularly advantageous for the development of digital educational tools. By assigning Uniform Resource Identifiers (URIs) to data points and ensuring their retrievability via HTTP protocols, LOD establishes a standardized framework through which linguistic and educational data can be seamlessly discovered, accessed, and cross-referenced. This framework fosters the creation of dynamic learning ecosystems in which data is not siloed but continuously enriched through links to external resources. In the context of language learning, such interconnectivity allows learners and educators to access extensive, up-to-date linguistic information, benefiting from the richness and diversity of global knowledge networks.

A more specialized branch of this movement, Linguistic Linked Open Data (LLOD), applies these principles to the representation of linguistic resources such as lexicons, ontologies, morpho-syntactic databases, and multilingual corpora. Initiatives such as OntoLex-Lemon have formalized models for describing lexical and semantic information within the RDF framework, allowing words, senses, and grammatical categories to be represented as interconnected nodes within a global linguistic graph. These representations enable fine-grained modeling of linguistic phenomena, from word sense disambiguation and cross-lingual alignment to the mapping of semantic fields across languages. Importantly, LLOD supports pedagogical aims by providing learners with structured, discoverable, and contextually enriched resources that extend far beyond static dictionary definitions.

For adaptive language learning platforms, the integration of LLOD ensures that their linguistic backbone remains scalable, transparent, and adaptable. Unlike closed or static datasets, which may become outdated or restricted in scope, LLOD resources can be continuously expanded and updated through community contributions. This openness enhances the reliability and relevance of educational tools while aligning with didactic imperatives to expose learners to authentic, evolving language use. Embedding a learning system within the broader LLOD ecosystem also allows it to evolve alongside ongoing advances in both computational linguistics and educational methodology. In this sense, the incorporation of LOD and LLOD represents more than a technical strategy, it embodies a pedagogical philosophy of interconnected learning. Just as language itself constitutes a network of interrelated forms, meanings, and contexts, linked data situates learners within a digital knowledge environment that mirrors this complexity. Rather than encountering isolated lexical items or grammatical rules, learners engage with language as a living network of semantic and syntactic

relationships. This approach aligns with cognitive models of the mental lexicon and contemporary theories of Second Language Acquisition (SLA), emphasizing connection, association, and contextual learning as drivers of proficiency development.

Moreover, the foundational principles of LOD align closely with the FAIR principles: Findability, Accessibility, Interoperability, and Reusability, as articulated in The FAIR Guiding Principles for Scientific Data Management and Stewardship (Wilkinson et al., 2016). By adhering to these standards, LOD ensures that integrated data is discoverable, accessible under open terms, interoperable with other datasets, and reusable across educational and research contexts. As Wilkinson and colleagues emphasize, these principles provide essential guidance for data producers and publishers seeking to navigate the complexities of digital scholarship, ultimately enhancing the value, longevity, and reach of open educational resources.

Given these advantages, the adoption of Linked Open Data frameworks represents a strategic choice for ensuring that educational technologies are grounded in data structures that support quality, transparency, and long-term sustainability. Through the use of LOD, developers and educators commit to creating open, scalable, and interoperable learning environments that empower both learners and institutions to engage with shared, evolving knowledge networks.

The integration of standardized data structures, such as ontologies, plays a pivotal role in ensuring consistency, interoperability, and semantic coherence across diverse datasets. Ontologies establish a shared vocabulary for defining entities, relationships, and categories within a given domain, creating a logical framework for organizing and connecting knowledge. By formalizing how data is represented and related, ontologies support efficient data retrieval, semantic reasoning, and resource alignment across multiple systems.

In educational contexts, this ontological structuring enhances both flexibility and scalability, enabling platforms to accommodate diverse learner profiles, content types, and linguistic resources. Ontology-driven data modeling ensures that language-related information, from morphological tags to pragmatic annotations, can be seamlessly integrated and repurposed across applications. Such interoperability is critical for sustaining adaptive learning ecosystems capable of evolving with pedagogical and technological advancements.

The dynamic nature of LOD allows for continuous growth and enrichment as new data sources are incorporated through the Web of Data. This expansion is both automated and community-driven, reflecting the collaborative ethos of open data. The resulting repositories are human-readable and machine-processable, ensuring transparency for learners and educators while enabling computational processing through algorithms that support personalization and real-time data integration.

A defining strength of LOD lies in its heterogeneous composition, exemplified by the diversity of domains represented within the Linked Open Data Cloud (LLOD). As illustrated in Figure Figure 2, this interconnected ecosystem spans fields as varied as geography, life sciences, linguistics, government, media, and social networks. The breadth of this network demonstrates LOD's capacity to bridge disciplines and facilitate interdisciplinary exploration, providing opportunities to link linguistic information with cultural, scientific, or historical data.

Through these interconnections, LOD and LLOD exemplify how the global web of data can underpin the next generation of educational technologies, platforms that are open, collaborative, contextually aware, and capable of evolving in tandem with both technological progress and linguistic diversity.

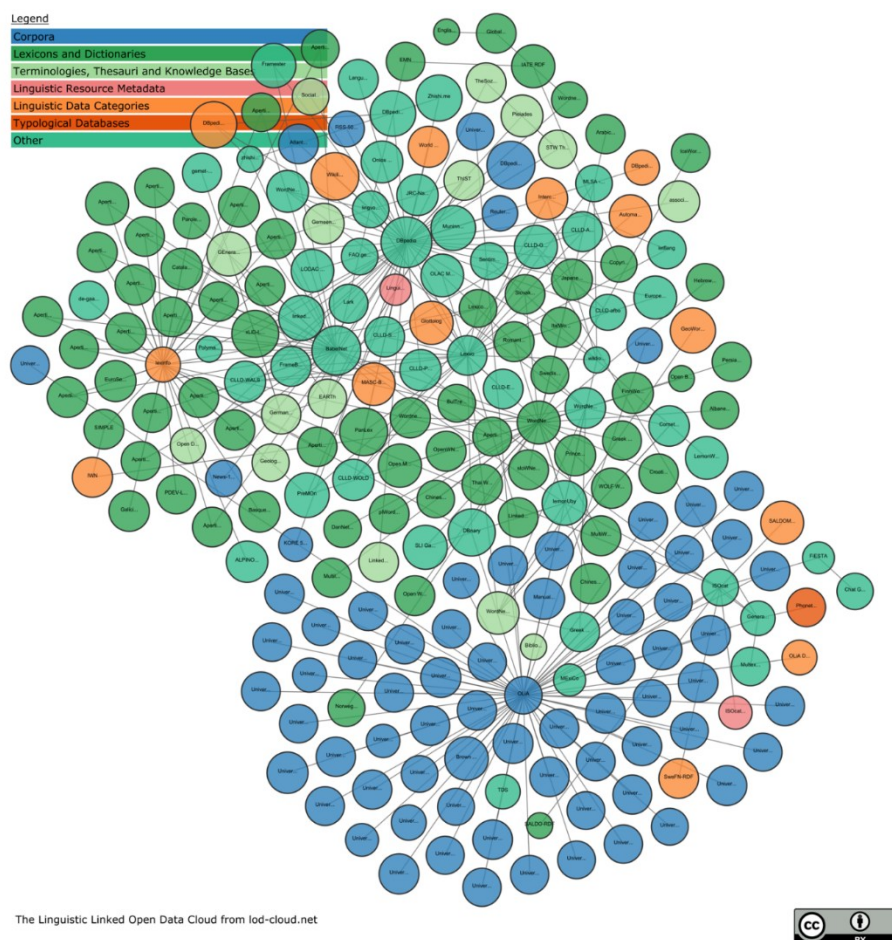


Figure 2 LLOD cloud September 2023 - <https://linguistic-lod.org/lod-cloud>

As we narrow our focus to the linguistic domain within the LLOD Cloud, it becomes clear that the linguistic resources available represent a rich and diverse set of knowledge graphs. The LLOD movement, championed by the Open Linguistics Working Group, has seen remarkable growth since its inception. The integration of a large number of linguistic datasets into the Linked Data Cloud has led to substantial improvements in the quality, structure, and breadth of the data. Notably, the datasets have evolved significantly over time, with many repositories being updated regularly to enhance their

coverage and functionality, while others continue to serve as important milestones in the progression of the field.

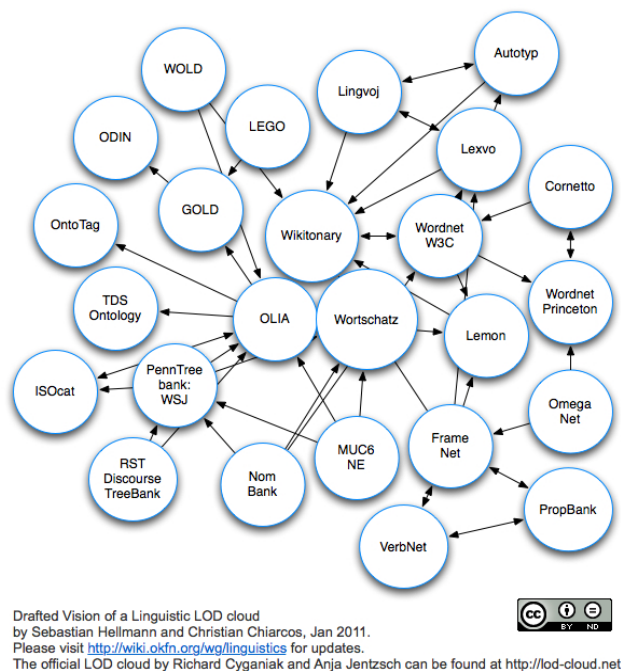


Figure 3 LLOD Cloud, January 2011

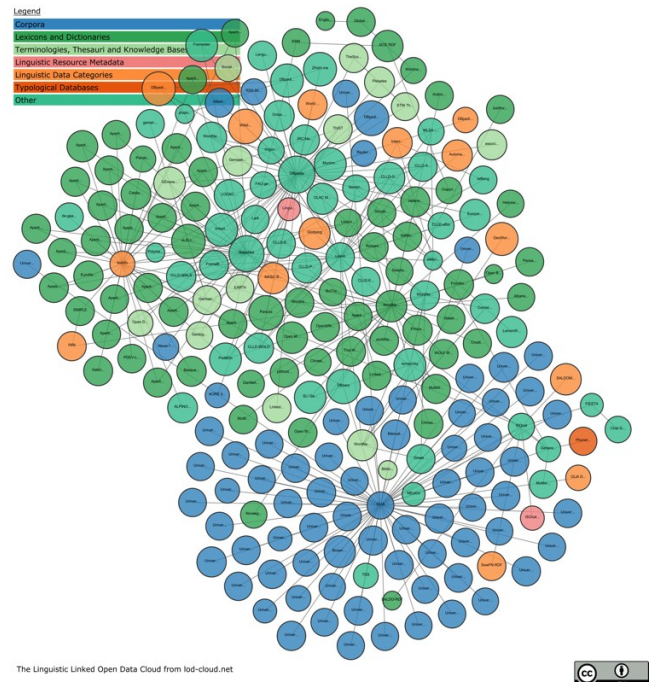


Figure 4 LLOD Cloud September 2023

The evolution of Linguistic Linked Open Data (LLOD), from its early integration into the Semantic Web community in 2011 to the present, reflects the rapid growth in both the quantity and quality of linguistic datasets available for research and education. Although some resources have become dormant or are no longer frequently updated, their historical contributions remain invaluable in shaping the current LLOD landscape. These earlier datasets, whether actively maintained or not, constitute significant milestones in the ongoing development of structured, openly accessible linguistic resources, providing essential groundwork for subsequent innovation in the representation and sharing of linguistic knowledge.

In light of these developments, the integration of LLOD into contemporary language-learning environments is driven by the need for high-quality, interoperable, and freely accessible linguistic data capable of supporting diverse and adaptive educational experiences. The LLOD cloud, with its extensive collection of linguistic knowledge graphs and adherence to Linked Data principles, offers an unparalleled opportunity to enrich digital learning tools with contextually relevant, real-time linguistic information. By leveraging this interconnected infrastructure, language-learning platforms can provide learners with immersive, adaptive, and data-driven experiences that mirror the evolving dynamics of human language and communication.

3.3.5 Systematic Literature Review: Linked Open Data

To achieve a comprehensible overview of the current state of Linguistic resources in the Linked Data environment and available resources, we explored peer-reviewed articles that document the effort in publishing and maintaining Linguistic KGs, while the performed assessment focuses on the lexical and semantic features coverage of LLOD. We moved from more than 2.000 to 97 peer-reviewed articles. By deeply exploring those contributions, we narrowed it down to six relevant contributions that adopted a cross-domain perspective and referred to linguistic LD through a linguistic lens. Moreover, regardless of the format in which data are released, we also included two resources that are considered as a reference by linguistics for their substantial role in defining the synset notion, a foundational aspect in linguistics. WordNet is an example of this direction. Therefore, the current research not only deals with the structure and features of linguistic KGs but also offers a qualitative overview of crucial aspects in shaping and conceiving LD resources relevant to linguistics. Linguistic linked data and ontologies have been designed to recreate a bond between linguists and technical perspective and shape information around accurate ontological models such as OntoLex-Lemon . The innovative component LD can certainly represent a relevant advancement in the field of digital lexicography, but they require following some linguistics criteria to achieve high quality standards. As demonstrated in recent articles, efforts have been made to offer systems and directives which are convenient for linguistic research (Khan et al., 2022; Spahiu et al., 2022) in metadata enrichment (Di Buono et al., 2022), dictionary generation from encyclopedic knowledge (Bajcetic & Declerck, 2022). Data have been also kept machine readable and exploitable for computer science's common task, such as, Large Language Models (LLMs) and word embeddings improvement (Celikyilmaz et al., 2015), word-sense disambiguation (AlMousa et al., 2022; Gangemi et al., 2016) meaning representation, cross language linking (Chen et al., 2021; Miao et al., 2013; Narducci et al., 2013), Natural Language Processing operations and question answering (Li & Xu, 2016; Skandan et al., 2023). The utility of LLOD in LLMs and, in general, Artificial Intelligence (AI) applications require a deep focus on lexical resources and the way they are structured to be fully and properly exploited in natural language related tasks.

This section investigates to what extent the LLOD modeling approaches satisfy linguistics' expectations. More in detail, it qualitatively compares features expected by linguists and if and to what extent they are covered by ontologies used by the LLOD as a reference. It is worth noting that it embraces the following observations mainly through linguistics' lens proposing LLOD ontologies extensions that might be beneficial for learning applications.

The review was performed with the aim of analyzing the current state of LLOD from a technical and linguistic perspective. Hence, it aims to achieve a comprehensive understanding and perform a critical assessment concerning the interest of the Semantic Web community in designing and publishing (high-quality) KGs accurately modeling information that is relevant to linguistics and to what extent lexicographic or linguistic areas can take advantage of Semantic Web technologies in their research and applications. The RQs at the basis of this literature review follow:

RQ1 How many linguistic KGs are available and working?

RQ2 Do working linguistic KGs properly cover linguistics' expectations in terms of lexicographic information completeness?

The presented literature review was conducted by in-depth reading, interpreting, and categorizing papers that describe, present, and review any KG contributing to the LLOD and the use of Semantic Web technologies in either (both) information technology or (and) lexicographic or linguistic areas. The review focuses on peer-reviewed contributions indexed by the Scopus database as it is one of the most comprehensive database sources (Austin, John Langshaw, 1962) and offers the broadest document coverage over other databases (Mongeon & Paul-Hus, 2016), indexes the widest number of peer-reviewed literature sources (Martín-Martín et al., 2018) concerning the Semantic Web community, and includes interdisciplinary literature across all research fields, so the probability of missing key research information is greatly reduced. However, the procedure is fully detailed to make it possible to repeat it systematically on other databases. We used KG, linguistics, and different variations of these terms as keywords. Specifically, we carried out the following query:

```
TITLE-ABS-KEY((knowledge graph* OR semantic network* OR
linked data OR linked open data) AND (linguistic* OR
encycloped* OR lexi* OR dictionar*))
```

We considered only English contributions published in the last 10 years (from 2013 up to July 2023). A total of 2.347 met these criteria. Figure 3 summarizes and schematically reports the exclusion and inclusion criteria applied to both the abstract and the full-text revision, which will be detailed in the following schema:

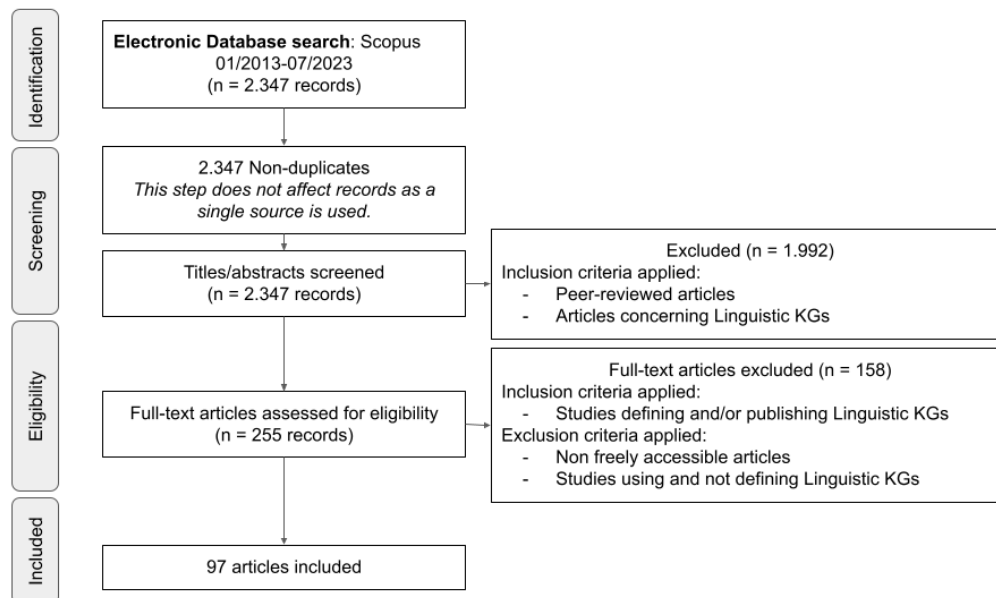


Figure 5 PRISMA chart LLOD research

To be included in this review, papers needed to propose a KG modeling information and relevant contribution to the expansion of English language digital lexicographic resources, the process of creation of a resource, and its development through time. The articles included in the review also refer to other resources that have not been included in the following table, being them translation-oriented, non-English, etymological, mainly encyclopedic, or domain-specific. Considering these criteria, the number of articles included was narrowed down to 97. We excluded all non-peer-reviewed and non-accessible papers as well as papers with a topic not relevant to this review. For instance, we left out articles that refer to KGs exclusively to implement or improve NLP tasks such as word-embeddings creation, training, and sentiment analysis tasks. Our search initially yielded a broader range of articles, including those from sub-fields like psycholinguistics and clinical linguistics, which were deemed irrelevant for our purposes. This refinement led us to focus on linguistic Knowledge Graphs (KGs) discussed in 97 peer-reviewed articles. It's important to note that while there are 97 articles, the number of unique KGs is smaller due to overlapping references.

Figure 5 provides a comprehensive overview of the linguistic information and lexical-semantic features offered by various linguistic resources. These resources span encyclopedic databases, linked open data initiatives, and specialized linguistic repositories.

The table is organized into three main categories:

CATEGORY	DESCRIPTION	EXAMPLES / KEY ELEMENTS
Basic Linguistic Properties	Contains fundamental lexical information used to describe the core attributes of a word.	Word sense, definition, example usage, phonetic transcription, translation, derivation, etymology.
Lexical–Semantic Features	Focuses on the semantic relations that define connections among lexical items.	Synonyms, antonyms, hypernyms, hyponyms, meronyms, holonyms.
Additional Features	Encompasses supplementary elements tailored for advanced linguistic and computational applications.	Polysemy disambiguation, compound formation, derived forms.

Table 4 Overview of lexical data categories and their descriptive scope in language resource design.

By analyzing this table, researchers and developers can gain valuable insights into the specific linguistic information and semantic relationships available from different resources, enabling them to select the most suitable tools for their specific needs.

	BabelNet	OpenEnglish WordNet	WikiData	DBnary	DBpedia**	Princeton WordNet 3.0*	ConceptNet	Wiktionary*
Active	✓	✓	✓	✓	✓	✓	✓	✓
Word sense	✓	✓	✓	✓	✓	✓		✓
Definition	✓	✓	✓		✓	✓		✓
Example	✓	✓	✓			✓		✓
IPA	!				✓			✓
Translation	✓						✓	✓
Derivation	✓	✓			✓			✓
Etymology	!				✓		✓	✓
Lexical semantic features								
Synonyms	✓	✓	✓	✓	✓	✓	✓	✓
Antonyms	✓	✓		!	✓	✓	✓	✓
Hypernyms	✓	✓		!	✓	✓	✓	✓
Hyponyms	✓	✓			✓	✓	✓	✓
Meronyms	✓	✓		!	✓	✓	✓	✓
Holonyms	✓	✓		!	✓	✓	✓	✓
Polysemic disambiguation	✓	✓				✓	✓	✓
Compounds	✓	✓		✓	✓	✓	✓	✓
Derived forms	✓	✓		✓	✓	✓	✓	✓

Table 5 Overview of linguistic information and lexical-semantic features reported by the analyzed resources.

The survey conducted as part of this study played a crucial role in identifying eight repositories that are well-suited to serve as potential candidates for achieving the objectives of this research. Through a thorough examination of these repositories, we were able to assess their relevance, coverage, and potential contributions to the overall scope of the study. In addition to the identification of repositories, the survey facilitated the discovery of a wide range of pertinent scholarly articles and papers that address related works in the field. These articles were subsequently reviewed and analyzed in depth to provide a comprehensive understanding of the current landscape of research on Linguistic

Linked Open Data (LLOD) and its applications. This process not only enabled us to map out the existing knowledge but also allowed us to identify gaps and opportunities for further investigation. The datasets and resources that emerged from this process are presented in the following section, offering a detailed overview of the materials that form the foundation of this study's analysis.

3.3.5.1 Princeton WordNet (Miller, 1995)

Princeton WordNet represents a groundbreaking shift in the way lexical information is organized, moving away from traditional dictionary structures toward a more cognitive and conceptual model. Developed under the guidance of George A. Miller, WordNet was designed with the aim of reflecting psycholinguistic theories of human lexical memory, offering insights into the way the human brain organizes and retrieves words. Unlike conventional dictionaries, which tend to focus on alphabetical ordering and etymological roots, WordNet organizes English words, nouns, verbs, adjectives, and adverbs, into sets of synonyms, referred to as synsets. Each synset represents a distinct lexical concept, allowing for a more nuanced and holistic view of word meaning.

In his seminal paper, Miller (1995) described WordNet as a 'dictionary based on psycholinguistic principles, highlighting its innovative approach to structuring lexical knowledge. This methodology reflects an effort to replicate the semantic relationships that exist within human cognition, distinguishing it from the traditional, formal approaches used in lexicography. WordNet does not simply categorize words based on superficial characteristics but rather focuses on the underlying conceptual connections between them. The system allows for more meaningful organization, facilitating the identification of related words and their semantic properties.

The most recent version of WordNet includes approximately 117,000 synsets, each accompanied by gloss, short definitions that describe the meaning of a word within a synset, along with example sentences that demonstrate how the word is used in context. This rich dataset serves not only as a linguistic resource but also as a model for understanding the cognitive processes involved in word recognition and retrieval.

Rather than adhering to a rigid ontological structure, WordNet operates within a flexible system that captures a wide array of lexical-semantic relations. These relations include synonymy (the relationship between words that share the same meaning within a synset), as well as hypernymy (the hierarchical isA relationship), meronymy (part-whole relationships), antonymy (opposites), and hyponymy (subordinate relationships). This intricate web of relations mirrors the way words are connected in the human mind and provides a dynamic tool for navigating the complexity of the English lexicon.

Additionally, WordNet has been adapted for use within the context of Linked Open Data, transforming it into a valuable resource within the Semantic Web. By integrating WordNet with the lemon (Lexicon Model for Ontologies) and RDF (Resource Description Framework) formats, it has become interoperable with other datasets, facilitating richer and more flexible applications in natural language processing, artificial intelligence, and computational linguistics. This transition to linked data formats has greatly expanded WordNet's utility, allowing it to function as a critical component in various computational systems that rely on accurate, semantically-rich lexical information.

WordNet's impact extends far beyond the realm of traditional lexicography. It has become a foundational resource in linguistics, cognitive science, and computational fields, offering an invaluable tool for understanding how language functions both in the human mind and in the digital world. Its influence continues to shape the development of language processing technologies, serving as an essential reference for a wide range of research areas, from language acquisition to machine learning.

Beyond its primary role in linguistics and cognitive science, WordNet has found applications in a vast array of fields, demonstrating its versatility as a resource. In natural language processing (NLP), WordNet is widely used for tasks such as word sense disambiguation, where it helps determine the correct meaning of a word based on context. It is also employed in information retrieval, improving search algorithms by providing a semantic understanding of queries and documents. WordNet's structure enables the development of more sophisticated machine translation systems, where it facilitates the alignment of words and concepts across languages by leveraging synonymy, hyponymy, and other semantic relations. In addition, WordNet has been utilized in sentiment analysis, helping to identify and categorize the emotional tone of texts by analyzing the relationships between words and their associated meanings. It has also been applied in areas such as text classification, question answering, and even robotics, where an understanding of language semantics is crucial for interaction and decision-making. Furthermore, WordNet's integration into the Semantic Web has allowed the creation of more intelligent, interconnected systems capable of processing and interpreting data in ways that reflect human-like understanding of language. This wide range of uses illustrates WordNet's enduring value as a versatile tool for both academic research and practical applications in various domains.

3.3.5.2 *BabelNet* (Navigli & Ponzetto, 2012)

BabelNet is an expansive and highly sophisticated multilingual semantic network and knowledge base that was originally developed by integrating concepts from WordNet and Wikipedia. Since its

creation, *BabelNet* has undergone continuous development, with significant improvements in both its scope and the quality of its data. As of November 2022, *BabelNet* reached its version 5.2, marking a major milestone in its growth, further extending its database and refining its capabilities. The structure of *BabelNet* is organized as a multi-graph system that seamlessly links words and concepts across a multitude of languages through semantically rich relationships. This interconnected framework not only enhances the accuracy and comprehensiveness of its data but also provides a deep understanding of the way words and concepts relate to one another across different languages and cultures.

One of *BabelNet*'s most remarkable features is its comprehensive multilingual coverage. The learning platform supports synsets, sets of synonyms that represent the same meaning, across an impressive 520 languages in its most recent release. This extensive linguistic range positions *BabelNet* as one of the most inclusive and versatile resources for multilingual semantic information. The diversity of languages covered ensures that *BabelNet* is invaluable for both researchers and practitioners working with non-English languages, offering a wealth of data that extends far beyond traditional English-centric resources. *BabelNet*'s support for a wide array of languages makes it particularly useful for global applications, as it fosters the creation of systems that can understand and process multiple languages in a unified way.

In addition to its broad multilingual coverage, *BabelNet* embraces multimodal data, further enhancing its utility and scope. The learning platform not only provides text-based information but also allows users to link words with relevant images, offering a richer, more contextually informed understanding of words and their meanings. This multimodal capability is especially valuable for applications in areas like image recognition, where the association between words and images can aid in the development of more sophisticated, context-aware systems. By incorporating visual elements into its network, *BabelNet* creates a more immersive and holistic learning experience for users and provides a deeper understanding of language concepts.

BabelNet's latest version (version 6) has been designed with an emphasis on user accessibility, featuring an intuitive and user-friendly interface that facilitates seamless navigation through its vast and complex repository of linguistic and encyclopedic knowledge. Users can search for specific words or concepts, and *BabelNet* will return a comprehensive set of results, including corresponding synsets, translations, images, and detailed graph relations that highlight the semantic connections between concepts. The learning platform also visualizes how these concepts are linked across languages, providing an insightful perspective on their relationships and context. This rich interconnection of data makes *BabelNet* a powerful tool not only for linguistic research but also for practical applications in fields such as natural language processing (NLP), machine translation, and multilingual information retrieval. Researchers, developers, and language professionals rely on

BabelNet for its extensive, structured knowledge, which supports more accurate translations, improved search functionality, and enhanced comprehension of complex multilingual content.

In summary, *BabelNet* is a cutting-edge resource that transcends traditional lexical databases, offering a robust, multilingual, and multimodal network that serves a wide array of practical and academic purposes. Its continued development and expansion ensure that it remains at the forefront of language resources, making it an indispensable tool in the field of computational linguistics and beyond.

3.3.5.3 Open English WordNet (McCrae et al., 2019):

Open English WordNet (OEW) is a linguistically rich, open-source resource that serves as an alternative to the renowned Princeton WordNet, one of the most widely used lexical databases for the English language. OEW is distinct in that it is a ‘fork’ of the original Princeton WordNet, meaning it was developed as an independent project, retaining compatibility with its predecessor while introducing its own unique features and improvements. This forking approach allows OEW to function as a drop-in replacement for Princeton WordNet, ensuring seamless integration into applications and systems that were initially designed to rely on the original version. By maintaining this compatibility, OEW serves as an accessible and flexible option for users who wish to enhance their existing workflows without undergoing major modifications to their infrastructure.

The inaugural release of OEW in 2019 marked a significant milestone in the development of this open-source resource. The first version included over 2,000 synsets that were carefully modified to improve their accuracy and semantic representation, addressing some of the limitations of the original WordNet. Since then, the resource has expanded substantially, with the 2020 release growing to encompass 120,054 nodes, illustrating the project's continued growth and evolution. This expansion is a testament to the increasing demand for high-quality, open-access lexical resources, and OEW's commitment to meeting this demand through ongoing development and refinement.

OEW is designed to be versatile and accessible, supporting a variety of formats, including RDF (Resource Description Framework), XML, and database formats. This broad compatibility ensures that OEW can be integrated into a wide range of applications across different domains, including computational linguistics, natural language processing (NLP), information retrieval, and other fields that rely on rich lexical data. Its open-source nature and availability in multiple formats make it a highly flexible tool, allowing researchers, developers, and educators to adapt it to their specific needs. A standout feature of OEW is its commitment to maintaining high-quality standards through manual updates, which differentiates it from many other lexical resources. Unlike systems that incorporate automatically generated changes or updates, OEW relies on contributions from a community of linguists and developers who submit pull requests for any modifications. This manual approach

ensures that each change undergoes careful review and validation, preserving the integrity and reliability of the resource. By prioritizing accuracy and thoughtful revision, OEW ensures that its content remains stable and dependable, which is particularly important for users who require consistent, high-quality lexical data for research, teaching, or application development.

In addition to its technical rigor, OEW offers a valuable asset for those looking for a high-quality, stable lexical database that can be freely used, modified, and distributed. Its open-access nature makes it an ideal resource for researchers, educators, and developers working within the fields of linguistics, artificial intelligence, and language education. Furthermore, OEW's adaptability to various linguistic and educational applications, ranging from language teaching tools to advanced NLP algorithms, ensures its continued relevance in a wide range of contexts. As the resource continues to evolve and grow, OEW will undoubtedly remain a vital tool for those seeking to integrate lexical data to enhance our understanding of language and support the development of more sophisticated linguistic technologies.

3.3.5.4 *ConceptNet* (Speer et al., 2018)

ConceptNet is a sophisticated and highly influential semantic network designed to improve the common-sense understanding of words and concepts by computational systems. The development of *ConceptNet* was driven by the need to enhance the way machines interpret human language, providing them with a more intuitive understanding of the meanings and relationships between words. Unlike traditional lexical databases that primarily focus on formal semantic relationships, such as synonyms or definitions, *ConceptNet* places a strong emphasis on the integration of common-sense knowledge. This inclusion of conceptual relationships, derived from human experiences and everyday reasoning, enables machines to better understand language in a way that mimics human cognitive processes. *ConceptNet* synthesizes data from a wide array of sources, with *Wiktionary* being the largest contributor, providing a substantial 18.1 million edges that link words and concepts in the network. This diversity of data sources ensures that *ConceptNet* captures a broad spectrum of knowledge that is critical for applications in natural language understanding, artificial intelligence (AI), and cognitive computing.

What sets *ConceptNet* apart from traditional lexical networks is its unique structure, which does not organize knowledge into synsets as found in other resources like WordNet. Instead, *ConceptNet* represents information as a dynamic network of interconnected concepts. These concepts are linked by a diverse set of relationships that reflect the ways in which humans understand and relate to the world around them. In its latest release, *ConceptNet 5.5*, the network organizes its data through 36 distinct types of relations, which encompass both symmetric and asymmetric connections. Symmetric

relations include synonymy (words that mean the same thing), etymologicallyRelated (words with a common origin), and SimilarTo (concepts that share similar attributes). Asymmetric relations, on the other hand, describe more directional connections, such as HasProperty (a specific characteristic of an object), InstanceOf (a specific example of a broader category), IsA (subcategorization of a concept), MadeOf (material composition), MannerOf (how something is done), MotivatedByGoal (actions driven by intentions), and ObstructedBy (barriers preventing actions). These diverse relations create a nuanced framework for capturing the richness of meaning that goes beyond simple word definitions.

The richness of *ConceptNet*'s relational structure allows it to represent a wide variety of conceptual and experiential connections between words, offering a more comprehensive and contextual understanding of language. For example, *ConceptNet* links words not just by their direct meanings but by physical properties, causality, motivations, and goals, among other factors. This allows the network to represent the context in which words are used, accounting for both abstract and concrete aspects of human knowledge. This ability to capture the deeper, often implicit meanings of words makes *ConceptNet* particularly valuable for tasks that require nuanced comprehension, such as dialogue systems, sentiment analysis, and advanced machine learning applications. By providing machines with access to a broad range of common-sense knowledge, *ConceptNet* enables systems to simulate human-like reasoning, allowing for more natural and contextually aware interactions between humans and machines.

Furthermore, *ConceptNet*'s structure is designed to be highly scalable and adaptable, which allows it to be continually updated and refined as new knowledge becomes available. This scalability ensures that *ConceptNet* remains relevant in the fast-evolving fields of AI and natural language processing. Whether used for improving search engines, enhancing virtual assistants, or powering sophisticated machine learning models, *ConceptNet*'s comprehensive, context-sensitive approach to semantic relationships makes it an indispensable resource for advancing the capabilities of artificial intelligence systems.

In conclusion, *ConceptNet* stands as a pivotal tool in the field of AI and natural language processing, offering a rich, common-sense understanding of language that is crucial for developing systems capable of human-like reasoning and interpretation. Its broad range of semantic relations and diverse knowledge sources make it an essential resource for a wide array of applications, from conversational AI to machine learning and beyond.

3.3.5.5 Wiktionary

Wiktionary is a collaboratively constructed, multilingual dictionary that stands as one of the most influential and valuable resources in computer-based lexicography. As an open, crowd-sourced platform, *Wiktionary* represents a departure from traditional, centrally curated lexical databases. It follows the same collaborative principles that drive Wikipedia, allowing users from around the world to contribute and refine content in real-time. Unlike many lexical resources that are built on formal, linked graph structures, such as those found in Linguistic Linked Open Data (LLOD) networks, *Wiktionary* operates on a more decentralized model. It does not adhere to a strict semantic framework but instead focuses on the documentation of language as it exists, adopting a descriptivist approach. This means that *Wiktionary* does not attempt to prescribe or regulate language usage, but instead captures the diversity of language in all its forms, reflecting how words and phrases are used in various contexts, cultures, and regions.

What makes *Wiktionary* particularly valuable is the depth and heterogeneity of the information it contains. Each entry is typically comprehensive, including not only definitions but also etymological information, pronunciation (often in the International Phonetic Alphabet, accompanied by audio files), grammatical details, and usage notes that point out regional variations and stylistic preferences. This breadth of data makes *Wiktionary* a rich resource for linguistic researchers, language learners, and even software developers who rely on accurate and diverse linguistic information. The learning platform's structure allows it to document language from a variety of perspectives, historical, geographical, and social, ensuring a more inclusive representation of the languages it covers. Unlike more rigidly defined lexical databases, *Wiktionary*'s open-editing model fosters the continuous updating and refinement of entries, keeping the dictionary current with evolving language trends and usages.

Wiktionary stands as an exceptional resource due to its collaborative nature and open availability, covering an extensive range of languages and dialects. The sheer number of languages, including many that are underrepresented in more traditional lexical databases, makes it an invaluable tool for anyone interested in the nuances of language, whether for academic research, linguistic documentation, or personal enrichment. However, it is important to note that *Wiktionary* exists outside the Linguistic Linked Open Data (LLOD) network. While it is a rich repository of linguistic information, it does not strictly adhere to the formalized data structures of LLOD, such as those used by other lexicographical resources like WordNet or *BabelNet*. As such, *Wiktionary* functions independently, without formal semantic relationships or integration into the broader LLOD framework. Despite this distinction, *Wiktionary* remains a vital resource in its own right, offering

flexibility, accessibility, and comprehensive language data that is not constrained by the more rigid conventions of traditional linguistic datasets.

The user interface of *Wiktionary* is designed to be highly accessible, allowing easy navigation of its vast content through a simple, intuitive layout. The continuously expanding database reflects the contributions of a global community of editors, making it one of the most dynamic and up-to-date lexicographical tools available. As an open-source project, *Wiktionary* also ensures that its content can be freely used, shared, and adapted, providing an invaluable resource for both casual language enthusiasts and professionals in fields ranging from linguistics to software development.

In conclusion, *Wiktionary*'s open and decentralized model offers a wealth of linguistic information that is not only vast in scope but also flexible and constantly evolving. While it operates outside the formal structure of the Linguistic Linked Open Data network, its inclusive and descriptive approach to documenting language makes it an indispensable tool for those seeking a comprehensive, user-driven resource that reflects the full spectrum of human linguistic expression.

3.3.5.6 *DBnary* (Sérasset, 2015; Sérasset & Tchechmedjiev, 2014)

DBnary is an ontology-based dataset and lexical network that transforms and structures data extracted from various editions of *Wiktionary*, aligning it with the principles of Linked Data (LD). Developed by Gilles Sérasset, *DBnary* was conceived as a tool to enhance the accessibility and utility of *Wiktionary*'s extensive content. By organizing this information into an ontological framework, *DBnary* offers a more sophisticated approach to querying and enables seamless integration with other linguistic resources, thereby extending the practical applications of *Wiktionary* data.

The primary goal of *DBnary* is to present *Wiktionary*'s vast and diverse lexical information in a format that adheres to the standards of Linked Data, making it more interoperable with other datasets in the Linked Open Data (LOD) cloud. This structured representation allows for the easy linking of linguistic data across different sources, which significantly enhances the resource's utility for researchers, linguists, and developers who seek to integrate structured lexical data into their applications. One of *DBnary*'s key innovations lies in its ability to provide enhanced translation data. The system extracts, organizes, and refines translation entries from *Wiktionary*, offering more precise and detailed translation information than is typically available in the original resource. This additional layer of data provides greater accuracy, which is crucial for multilingual tasks such as machine translation and cross-linguistic research.

In addition to improving translation data, *DBnary*'s ontological structure supports the inclusion of a broader array of semantic relations than those typically found in *Wiktionary*. By going beyond simple word definitions and translations, *DBnary* incorporates detailed information about grammatical

properties, word senses, and interlingual relationships, creating a more comprehensive and nuanced lexical resource. This expansion of semantic relations adds depth and utility to the dataset, making it an invaluable tool for natural language processing (NLP), multilingual information retrieval (MIR), and other areas of computational linguistics where structured and interoperable data is essential for the development of sophisticated language technologies. The ontological framework also enables richer semantic queries, which improves the precision of applications that require nuanced linguistic data, such as sentiment analysis, automated content generation, and knowledge extraction.

Furthermore, *DBnary* bridges the gap between the informally constructed, crowd-sourced lexicographical data in *Wiktionary* and formal, structured ontological models. This hybrid approach makes *DBnary* a unique resource, combining the advantages of crowd-sourced lexicography with the rigor of ontological representation. As a result, *DBnary* not only enhances the usability of *Wiktionary* data but also contributes to the broader field of linguistic data interoperability, facilitating the integration of lexical knowledge across diverse domains and platforms.

DBnary's structure and adherence to Linked Data principles make it a powerful tool for advancing research and development in computational linguistics. Its ability to harmonize and enrich lexical data from a variety of sources opens up numerous possibilities for linguistic applications, from improving machine learning models to enhancing search engine accuracy and refining language resources for digital assistants. By offering an enriched, accessible, and interoperable version of *Wiktionary*, *DBnary* plays a critical role in the development of more intelligent, context-aware systems that are capable of processing language in ways that are both linguistically sophisticated and computationally efficient.

In conclusion, *DBnary* represents a significant step forward in the integration of crowd-sourced lexicographical data with formal ontological frameworks. By transforming *Wiktionary*'s content into a Linked Data-compliant structure, *DBnary* enhances the accessibility, accuracy, and utility of this valuable linguistic resource, making it indispensable for both academic and applied linguistics. The resource's capacity to facilitate detailed, semantically rich queries and its ability to support a broad range of language-related applications mark it as an essential tool for advancing the field of natural language processing and computational linguistics.

3.3.5.7 OntoLex Model and Its Extensions

As the need for structured, interoperable linguistic data grows, the Ontology-Lexicon⁴ (OntoLex) model has emerged as a crucial framework for representing lexical and linguistic information in the

⁴ <https://www.w3.org/2019/09/lexicog/>

Semantic Web. Developed within the W3C Ontology-Lexicon Community Group and finalized in May 2016, OntoLex provides a sophisticated and extensible model for linking lexical data with ontological knowledge. Its primary goal is to bridge the gap between linguistic resources and structured knowledge representation, ensuring seamless integration across diverse datasets. Designed to support applications in natural language processing (NLP), lexicography, and linked data technologies, OntoLex offers a modular structure that enables the representation of complex linguistic phenomena while maintaining interoperability with other Semantic Web technologies. At its core, the OntoLex-Lemon model serves as the foundation for structuring lexical entries, upon which multiple specialized extensions are built to capture various linguistic dimensions, such as syntax, semantics, morphology, and translation.

The modular approach ensures that linguistic data can be represented with precision and flexibility, accommodating the diverse and evolving nature of language. Each extension within the OntoLex framework addresses specific linguistic aspects, enhancing its ability to capture syntactic, semantic, morphological, and translational relationships. By integrating these modules, OntoLex provides a comprehensive and interoperable system for managing lexical data, facilitating its use in various linguistic and computational applications.

At its core, the OntoLex-Lemon model serves as the foundation for structuring lexical entries, upon which multiple specialized extensions are built. The OntoLex-SynSem module links syntactic structures with semantic representations, facilitating a richer understanding of meaning and usage. The OntoLex-Decomp module allows for the breakdown of complex lexical units into smaller components, aiding in morphological analysis. The OntoLex-VarTrans module supports the formal representation of linguistic variations and translations, crucial for multilingual applications. The OntoLex-LiMe module incorporates linguistic metadata, ensuring the inclusion of contextual information relevant to lexical resources (McCrae et al., 2020). The OntoLex-Lexicog module structures lexicographic data, improving its interoperability and accessibility across digital platforms (Bosque-Gil et al., 2016). Emerging extensions such as OntoLex-Morph, developed for modeling morphological data (Klimek & Chiarcos, 2016), and OntoLex-FrAC, which integrates corpus-based frequency and attestation information (Kernerman et al., 2021), further enhance the framework's capacity to represent complex linguistic phenomena and support computational lexicography within the broader Linguistic Linked Open Data (LLOD) paradigm (Cimiano et al., 2020).

Pragmatics plays a crucial role in understanding lexical variation, as it examines how language users employ linguistic competence in communication. The choice of words in speech and writing reflects speaker intent, social context, and interpersonal relationships. Lexical selection, shaped by an internalized lexicon, influences the weight and nuance of an utterance. The dynamic nature of

language, deeply connected to historical linguistics, underscores the evolving nature of phonetics, lexicon, semantics, and syntax. Language is not a static set of rules, but a fluid system shaped by the arbitrary choices of its speakers over time. While the OntoLex model provides a robust structure for capturing these variations, challenges remain in ensuring that it accurately reflects the complexities of regional differences, semantic shifts, and the fluidity of natural language, particularly as evidenced by ambiguities in synset-based systems such as *BabelNet*.

The Variation & Translation (VarTrans) module provides a systematic framework for representing linguistic variation across multiple dimensions, including diatopic (geographical), diaphasic (situational), diachronic (temporal), diastratic (social), and dimensional (conceptual) variants (McCrae et al., 2017; Cimiano et al., 2020). This model acknowledges the inherent dynamism of language, enabling the structured formalization of variation within multilingual and cross-domain terminologies.

However, a notable limitation of the current VarTrans model is the absence of explicit temporal markers, which are essential for tracking lexical evolution over time. Lexical change and semantic shift are time-sensitive processes through which new words emerge, existing terms acquire new meanings, and obsolete vocabulary falls out of use, phenomena that drive the continuous transformation of the lexicon (Traugott & Dasher, 2002; Blank, 1999). From a diatopic perspective, regional terminological differences further complicate representation, as conceptually equivalent terms may diverge significantly across varieties of the same language. Moreover, semantic change affects both the denotative (literal) and connotative (associative) layers of meaning, posing substantial challenges for linguistic modeling, lexical alignment, and semantic interoperability in computational applications (Grondelaers, Speelman, & Geeraerts, 2010).

To address this limitation, the LemonDIA extension (F. Khan et al., 2014) was proposed as a means of tracking diachronic shifts in word meanings. This extension introduces a temporal dimension to lexical representations, allowing for the documentation of semantic evolution over time. While the VarTrans module already facilitates the linking of lexical variants across different historical periods, it lacks chronological metadata and details regarding the duration and trajectory of these changes. LemonDIA enhances this capability by explicitly encoding temporal information, making it possible to trace the life cycle of lexical units and monitor their semantic transformations. The integration of a time-sensitive framework within OntoLex represents a significant step toward improving the ontology's ability to model real-world linguistic change, ensuring a more precise and comprehensive representation of language dynamics.

3.3.6 Meaning Representation

The shift from traditional lexicographical models to more interconnected frameworks, such as those based on Linked Open Data (LOD) and ontologies, marks a significant transformation in how lexical resources are structured. Traditional lexicons typically rely on isolated word entries or lemmas. However, synset-based representations, employed by various LLOD resources, offer a more flexible and expansive way to represent lexical knowledge. By organizing words into sets of synonyms (synsets) that represent a single concept, this model enables a richer, more nuanced understanding of word meanings, moving beyond simple definitions. This broader approach is particularly important for tasks requiring complex semantic relationships, as it supports more sophisticated querying and integration of lexical data across diverse sources. Our assessment of various linguistic resources highlights the role of ontologies in structuring these lexical relationships, emphasizing the value of synset-based models in enhancing the precision and depth of lexical representation, especially in computational linguistics and natural language processing applications.

3.3.6.1 Synset-based Representation

While traditional lexicons use lemmas as unitary representations, the Linguistic Linked Open Data (LLOD) Cloud also recognizes synset-based resources as lexicons and dictionaries, blending the characteristics of thesauri with those of descriptive dictionaries. This approach enables a broader and more flexible representation of lexical knowledge, moving beyond single-word entries to encompass networks of meaningfully related terms. Our investigation examines various linguistic resources to assess how they model formal relevance reports and register labeling, with a specific focus on the role of ontologies in structuring lexical relationships.

Synset-based representations have several advantages when it comes to processing formal and informal language in natural language processing (NLP). One of the primary benefits is the hierarchical organization of words into groups, or synsets, that are related through hypernyms (general terms) and hyponyms (more specific terms). This hierarchical structure allows for clearer semantic relationships and helps identify whether a term is being used in a broad or specific sense, which is particularly useful in analyzing formality. Furthermore, these representations aid in lexical disambiguation, an essential task in NLP, since words with multiple meanings can be resolved based on context. In formal contexts, where clarity and precision are crucial, synsets can often resolve ambiguities by grouping words with similar meanings, helping improve the accuracy of the analysis. Additionally, synsets provide a significant advantage in multilingual applications. Lexical databases like WordNet can be extended to other languages, enabling formality assessments to be conducted cross-linguistically, which is invaluable for translation and cross-cultural communication.

However, despite these advantages, synset-based representations also face certain limitations when applied to formality analysis. One of the key drawbacks is their lack of context sensitivity. While synsets can group words by their meanings, they do not account for the way meaning can change based on context. In a formal text, a word's meaning might be clear, but in an informal setting, the same word could convey a completely different nuance. This lack of adaptability is a significant challenge in analyzing formality, as the tone of language is often highly dependent on context. Additionally, synsets are typically designed with standard dictionary meanings in mind, which means they struggle to capture slang, colloquial expressions, and informal language that are critical in distinguishing between formal and informal styles, especially in modern texts like social media posts or online conversations. Another limitation of synset-based methods is their rigidity. Language evolves continuously, with words shifting in meaning over time. Synsets, by design, do not adapt easily to these changes, making them less effective for analyzing contemporary language use, where formality can fluctuate rapidly. Finally, processing text using synsets can be computationally intensive. The need for word sense disambiguation and hierarchical structuring can increase the complexity and time required for analysis, which might not be ideal for real-time applications where speed is essential.

One of the most pervasive relations within LLOD resources is synonymy, a defining feature of synset-based structures. This prominence is a direct consequence of the way data is linked, as synset-driven models inherently rely on clusters of words that share the same conceptual meaning. Notably, five out of eight repositories examined exhibit comprehensive and high-quality coverage of semantic relations. However, other resources either only partially document such relations or fail to populate the relevant fields, even when such structures are theoretically accounted for in their underlying ontologies.

The strength of Linguistic Linked Open Data (LLOD) frequently lies in its synset-based structure, which groups synonymous words under a shared conceptual umbrella rather than organizing entries around individual lemmas, as seen in traditional lexicographic models. This approach, first introduced by WordNet, has been widely adopted by other datasets, reinforcing the efficiency of representing lexical meaning through semantic clusters.

A key theoretical foundation of synset-based models is the notion of synonymy, a concept that remains a debated topic in linguistics. While synsets often assume that words within them can be used interchangeably in all contexts, even the original architects of WordNet acknowledged that absolute synonymy is a rare phenomenon. This assumption is rooted in the principle of truth value (often attributed to Leibniz), which posits that two terms (x and y) are identical if, and only if, they share every property (Miller et al., 1990; Tarski, 1995). However, most linguistic research suggests that true

synonymy, where two words are entirely interchangeable in every context, are virtually nonexistent. Instead, languages exhibit degrees of synonymy, where words share significant overlap in meaning but differ in connotations, register, or domain-specific usage. This nuance is critical for computational models, as failing to account for partial synonymy can lead to errors in natural language processing tasks such as machine translation and information retrieval. To illustrate this, consider the following WordNet synsets:

Synset 1 (n): color, colour, coloring, colouring – a visual attribute of things that results from the light they emit or transmit or reflect.

Synset 2 (v): recognize, recognise, realize, realise, agnize, agnise – to be fully aware or cognizant of something.

At first glance, synsets appear to merge different phonological and morphological variations. In Synset 1, for example, the American English (*AmE*) spelling color and the British English (*BrE*) spelling colour are grouped together. Similarly, the *BrE* spelling realise and recognise correspond to the *AmE* forms realize and recognize. While these variations do not alter the conceptual meaning, they highlight the need for lexicographic resources to document regional distinctions explicitly.

Despite the structural efficiency of synset-based models, they introduce challenges when dealing with dialectal variation, lexical ambiguity, and interlinguistic differences. A notable example is the term *sidewalk* in American English, which corresponds to pavement in British English. However, in *AmE*, pavement typically refers to the hard surface of a road rather than a pedestrian walkway. This distinction is reflected in various lexicographic definitions:

WordNet: (n) *sidewalk*, pavement – a walk consisting of a paved area for pedestrians, usually beside a street or *roadway*.

Oxford English Dictionary (OED): (n) *sidewalk* (U.S.) – a path with a hard surface on one or both sides of a road that people walk on.

Merriam-Webster: (n) a usually paved walk for pedestrians at the side of a street.

BabelNet: (n) *sidewalk*, *pavement*, *footway* – a walk consisting of a paved area for pedestrians, usually beside a street or *roadway*.

This variation leads to inconsistencies in *BabelNet*, where both *sidewalk* and *pavement* are grouped under the same synset while still being marked as *differentFrom* each other. The conflation of these terms creates semantic ambiguity, as their meanings diverge across dialects. A similar issue arises

with the *AmE* term truck and its *BrE* equivalent lorry. In *BabelNet*, lorry is classified as a synonym for truck but is also listed separately as *differentFrom* large goods vehicle, defined as any lorry with a gross combination mass over 3,500 kg.

The improvements in resources such as the Open English Wordnet (OEW) demonstrate a growing awareness of the complexities associated with dialectal variation and semantic precision in synset-based systems. OEW has introduced refinements aimed at addressing lexical ambiguity and regional variation, making semantic relations more granular. In contrast to traditional synset-based systems might group regionally distinct words like sidewalk and pavement into the same synset, OEW incorporates additional usage domains and contextual markers. For example, it explicitly distinguishes between terms with different meanings based on regional or social factors by tagging words with specific usage notes. This approach helps mitigate ambiguity by clarifying regional preferences.

For instance, OEW treats fall (*AmE*) and autumn (*BrE*) as distinct concepts with clear usage notes indicating their regional preference. Similarly, apartment (*AmE*) and flat (*BrE*) are treated as separate lexical items within the synset, with their corresponding regional usage explicitly documented. This refinement in lexical representation makes OEW a more precise resource for applications that require a nuanced understanding of word meanings across different English varieties and, potentially, across multiple languages.

By addressing these challenges, the evolution of LLOD resources underscores the need for greater attention to dialectal variation, interlinguistic distinctions, and partial synonymy. Future developments in lexical databases should continue to refine these models to improve their applicability in natural language processing, machine translation, and other computational linguistics applications.

In addition to refining the mapping of regional differences, OEW has also enhanced the coverage of derivational relationships. As noted earlier, WordNet handles derived forms like color and colour separately, but OEW expands on this by linking these forms more explicitly, showing the morphological connections between words with different spellings. For instance, the derived form of colour in *BrE*, colourize, is now more clearly connected to its American counterpart, colorize. Similarly, realise (*BrE*) and realize (*AmE*) are more accurately represented with their respective derived forms, such as realization (*AmE*) and realisation (*BrE*). These refinements provide a more comprehensive representation of lexical variation, including morphological and orthographic aspects, which are crucial for understanding semantic shifts across different language varieties. Additionally, words like defence (*BrE*) and defense (*AmE*) now appear more clearly linked in the lexicon with their respective regional forms, reflecting the variation in spelling while maintaining semantic equivalence.

Further improvements in OEW involve its integration with larger ontological frameworks, enabling it to interface seamlessly with other semantic web technologies. This integration allows OEW to include data from a broader array of linguistic resources, making it more versatile in handling complex semantic phenomena. By incorporating data from resources such as WordNet, *BabelNet*, and *DBnary*, OEW enhances its ability to map inter-lexicon relationships more effectively, providing a richer and more nuanced semantic network. For example, the word *color* in OEW can now be linked to hue and shade, allowing for a more comprehensive understanding of color-related concepts across different languages and contexts. These enhancements extend to other lexical domains, such as emotions, where words like *happy*, *joyful*, and *ecstatic* are connected to various synonymous terms in different languages, enabling more accurate and context-aware translations.

These advancements contribute significantly to the ongoing development of the Linguistic Linked Open Data (LLOD) Cloud, allowing for greater interoperability between different lexical resources. Through such efforts, OEW has moved closer to offering a more robust framework for understanding and organizing lexical data, addressing the issues posed by regional variation and ambiguity in synset-based representations. For example, OEW's refinement in distinguishing between *biscuit* (*BrE*) and *cookie* (*AmE*) helps clarify the conceptual nuances between these terms, where both may refer to a similar food item but differ in preparation or usage based on regional context. Another example is the distinction made between *elevator* (*AmE*) and *lift* (*BrE*), where OEW links these terms with additional contextual notes, allowing for clearer semantic mapping.

The improvements seen in OEW underscore the importance of continuously refining synset-based models to better capture the full complexity of language use across different contexts. The inclusion of usage domains, regional markers, and morphological relationships allows for a more context-sensitive approach to lexical representation. This shift enhances the model's ability to support advanced applications in natural language processing (NLP), machine translation, and semantic web technologies. By improving the alignment of regional variations and dialectal distinctions within the synset framework, OEW moves closer to a more inclusive and representative model for lexical resources. These updates enhance the semantic clarity of resources like *BabelNet* and WordNet, making them more applicable for tasks that require cross-linguistic and cross-regional data integration. Moreover, OEW's ongoing improvements highlight the importance of developing dynamic lexicons that evolve in response to linguistic changes, technological needs, and the complexities of real-world language use. As the LLOD ecosystem continues to grow, we can expect further refinement of synset-based models, which will ultimately provide a more accurate, comprehensive, and interconnected representation of language in the digital age. Examples such as the treatment of *ferry* (*BrE*) and *boat* (*AmE*) in OEW demonstrate how the system has become more

nuanced, with added contextual distinctions making these terms clearer in cross-linguistic comparisons.

The benefits of formatting linguistic data in a structured and semantically rich way, such as through models like the Open English Wordnet (OEW), are numerous, particularly in the context of language learning and education. By organizing lexical resources in a way that explicitly reflects regional variations, synonyms, derivational forms, and semantic relationships, learners can achieve a deeper and more accurate understanding of language.

One significant advantage is that this structure aids in the development of vocabulary by exposing learners to a broader set of lexical variants and meanings. For example, having access to distinct entries for sidewalk (*AmE*) and pavement (*BrE*) not only helps students recognize the regional differences in English but also encourages them to understand the underlying concept (i.e., a path for pedestrians) shared by both terms. This kind of exposure is particularly valuable in an increasingly globalized world, where learners often interact with texts or speakers from different English-speaking regions. A learner can also explore other related concepts, like *sidewalk* in relation to *curb* and pavement in relation to *roadway*, fostering a more nuanced understanding of how words relate to each other within different contexts.

Furthermore, the ability to track derivational relationships is especially beneficial for learners seeking to expand their vocabulary and understand word formation. By linking words like color and *colour*, as well as *realize* and *realise*, with their derived forms (colorize, realization), OEW provides learners with tools to see patterns in word construction. This insight helps learners deduce the meanings of unfamiliar words by recognizing the roots and affixes. For example, knowing the connection between color and *colorize* allows learners to understand how new words can be created from base terms, strengthening their grasp of both vocabulary and grammar.

Additionally, learners benefit from the contextual markers and usage notes embedded within these resources. These markers help learners understand the appropriate usage of words in different contexts. For instance, a learner encountering the word *biscuit* would be made aware of the regional distinctions between British English (where it refers to a sweet baked good) and American English (where it refers to a type of bread roll often served with meals). By seeing this distinction explicitly noted, students gain not only a definition but also the contextual knowledge necessary for proper usage. In this way, learners are encouraged to pay attention to both meaning and context, enhancing their ability to use language flexibly and appropriately across different environments.

The semantic web structure also helps learners explore cross-linguistic comparisons. With more comprehensive data from diverse linguistic resources, learners can see how English words map onto concepts in other languages, such as how pavement might be linked to the equivalent term in Spanish

(acera) or French (trottoir). Such cross-linguistic associations deepen learners' understanding of both the language they are studying and the broader cultural and conceptual frameworks of the language. This global perspective enriches their learning experience by situating vocabulary within a broader context and helping them recognize similarities and differences across languages.

Moreover, synset-based approaches enhance learners' comprehension of synonyms and related terms by allowing them to see how different words that share similar meanings are connected. For instance, through OEW, students can explore how terms like happy, joyful, and ecstatic relate to each other within different contexts. This helps learners develop their vocabulary more effectively, allowing them to choose words with more precision based on the specific nuances they wish to convey. The connections between synonyms also promote a more dynamic understanding of meaning, encouraging learners to think critically about the subtle differences in word choices and how these differences can impact communication.

Finally, the structured approach to data in models like OEW allows learners to engage with language in a way that mimics real-world applications. By understanding how lexical items are organized within interconnected frameworks, students can develop more sophisticated skills in semantic analysis, which are critical for tasks such as translation, interpretation, and language synthesis. The ability to navigate complex webs of linguistic relationships, aided by the rich metadata and usage notes in these resources, enables learners to approach language as a living, evolving system rather than a static set of rules. This dynamic, interconnected approach is especially useful for learners who are working on advanced language skills, such as academic writing, research, or professional communication.

The structured approach to data in models like OEW allows learners to engage with language in a way that mimics real-world applications. By understanding how lexical items are organized within interconnected frameworks, students can develop more sophisticated skills in semantic analysis, which are critical for tasks such as translation, interpretation, and language synthesis. The ability to navigate complex webs of linguistic relationships, aided by the rich metadata and usage notes in these resources, enables learners to approach language as a living, evolving system rather than a static set of rules. This dynamic, interconnected approach is especially useful for learners who are working on advanced language skills, such as academic writing, research, or professional communication.

However, for these advantages to be realized, the data itself must be sensitive to the evolving nature of language and the diversity of learner contexts. By fostering awareness of the different linguistic and regional variations and the subtle nuances in meaning and usage, learners can become more attuned to the intricacies of language. A data structure that reflects this sensitivity not only enriches learners' vocabulary but also makes them more sensitive to the complexities of language, encouraging

critical thinking about word choice, meaning, and context. This heightened awareness can lead to a more profound understanding of language and its real-world applications.

3.3.6.2 Phonetic Transcriptions and Regional Variations

Esposito (2024) highlights that WordNet lacks phonetic transcriptions for its lexical units, whereas *BabelNet* includes them, often providing multiple International Phonetic Alphabet (IPA) transcriptions for a single entry. However, *BabelNet* frequently fails to explain these variations, leaving users without crucial information regarding pronunciation differences based on dialect, region, or register. This issue of insufficient phonetic and regional coverage is also prevalent in other synset-based repositories, limiting their effectiveness for linguistic research and language learning. In contrast, *Wiktionary* offers a more comprehensive and precise account of lemmas and their phonetic variants. For example, its entry for sidewalk includes distinct IPA transcriptions and explicitly marks the term as chiefly US, distinguishing it from pavement, which is more common in British English. *Wiktionary* also documents alternative pronunciations based on dialectal variation, such as differences between General American (GA) and Received Pronunciation (RP). This level of detail is particularly valuable for non-native speakers and comparative linguistic studies.

From a lexicographic perspective, *Wiktionary* follows a traditional descriptive approach, recording language as it is used, in contrast to WordNet's psycholinguistic methodology, which prioritizes cognitive representation and meaning relationships. Unlike WordNet, which organizes words into synsets (groups of synonymous terms), *Wiktionary* treats different formal representations of the same meaning as distinct entries. This structure allows for more granular documentation of phonetic, morphological, and syntactic variations.

Moreover, *Wiktionary*'s phonetic data is structured ontologically through *DBnary*, which builds on the OntoLex-lemon model. However, the *DBnary* framework currently omits phonetic information from its `DataProperty: PhoneticRepresentation`, resulting in a lack of phonetic details in its RDF triples (Esposito, 2024). This omission creates a gap between *Wiktionary*'s rich phonetic data and its machine-readable representation, limiting interoperability with other linguistic databases. Future improvements could integrate phonetic annotations into *DBnary*, enabling more robust cross-platform phonetic analysis.

Additionally, phonetic variation is a crucial aspect of sociolinguistics. The absence of explicit regional and dialectal distinctions in certain lexical databases raises concerns about linguistic bias, as some pronunciations may be overrepresented while others are marginalized. As global communication increases, the need for multilingual and multi-dialectal phonetic coverage grows, reinforcing the value of resources like *Wiktionary* that account for such diversity.

Another significant issue with synset-based knowledge graphs (KGs) is their handling of formality. Synsets often merge words with varying degrees of formality under the same concept without distinguishing between informal and formal variations. This approach disregards the pragmatic contexts that influence lexical choices at different levels of formality. In linguistic terms, context is crucial, and failing to account for it leads to inaccurate representations. For example, synset-based data might pair words with different formality levels together, treating them as equivalent representations of the same concept. This oversimplification can obscure important nuances in language use and fails to capture the true diversity of lexical expressions.

OEW: (v) skirt, hedge, put off, elude, dodge, duck, evade, parry, circumvent, fudge, sidestep avoid or try to avoid fulfilling, answering, or performing (duties, questions, or issues)

The set includes words from various levels of formality, including slang, colloquial, neutral, and formal language. However, the degrees of formality related to style and register variation are often unmarked, similar to how language variety is treated. While swear words, slang, and highly informal lexical units are typically labeled as highly informal, obscene terms, or obscenities, they are often semantically linked as hyponyms to more formal and non-vulgar counterparts. Synset-based approaches to digital lexicography consider the association of formal components with the cognitive system and aim to create a machine- and human-readable representation of the lexicon's semantic relations.

From this perspective, resources like *WordNet*, whose structure is primarily grounded in psycholinguistic principles (Miller, 1990), provide a robust foundation that has been used to model datasets such as *BabelNet*. Since synonymy is the central relation around which a synset is organized, other lexical-semantic relations between synsets are drawn based on the concepts encoded by the word units within the same set. However, this lack of detailed information at the top level of linguistic description can lead to inaccuracies that cascade through multiple nodes in the data.

WordNet's creators aimed to organize lexical information by word meanings rather than word forms, which is considered one of its most ambitious features. As a result, *WordNet* resembles a thesaurus more than a traditional dictionary (Miller et al., 1990), and resources modeled on its schema share this characteristic. Therefore, it is essential that resources based on this system's structure are supplemented with more accurate and detailed information, as proposed in more sophisticated ontologies.

The English language is widely recognized as a lingua franca on the Internet, where most English speakers use it alongside or in combination with other languages, often resulting in a hybrid form of online English (Seargeant & Tagg, 2011). While this hybrid nature is an undeniable reality of online communication, the domain of Linked Open Data for Linguistics (LLOD) which is specifically designed to describe language, requires a heightened sensitivity to such variation. LLOD's goal of publishing linguistic data through innovative approaches offers new perspectives on language description. This not only provides a well-organized framework for semantic relations but also introduces new methods for defining descriptive viewpoints of language.

The ontological models used in LLOD reflect the detailed approach applied to linguistic data description. For instance, the OntoLex-lemon specification, designed for lexicographic language description, allows for the representation of different realizations of the same linguistic unit. This is evident in the ontology's `DatatypeProperty: Written Representation` section. As OntoLex explains, for a lexical entry like *color*, there are two written representations: *colour* for British English and *color* for American English. Although these variants share the same pronunciation and meaning, they are treated as distinct lexicographic variants within the same lexical entry, denoted as `ontolex:written colour@en-GB` and `color@en-US`.

OntoLex-lemon demonstrates significant sensitivity to variation, especially through its core modules, such as the OntoLex-VarTrans module for Variation and Translation, which, though primarily focused on inter-language variation in translation, reflects this attention to detail. However, linguistic studies on language variety and speakers' perception of meaning indicate that certain lexical elements are not shared between British and American English. As shown by Scotto Di Carlo (2013), while some terms like *chemist/pharmacist* and *trousers/pants* are familiar to speakers of both varieties, others like *pavement/sidewalk* and *truck/lorry* can be entirely unfamiliar. In such cases, these are examples of polysemy rather than simple overlaps between forms associated with the same concept. Consequently, a synset might include various formal representations of the same concept without fully accounting for the variation within the English language.

We acknowledge that synset-based organization is an efficient method for grouping words that share the same conceptual reference. However, this approach can homogenize linguistic variation, resulting in unmarked descriptions that overlook important differences in language. Traditional dictionaries, like the Oxford English Dictionary (2023) and Merriam-Webster (2023), are typically organized according to specific language varieties, such as British English (*BrE*) and American English (*AmE*). These dictionaries often include terms from both varieties, explicitly indicating which variety a term belongs to. For example, while the word *lorry* is included in Merriam-Webster's definitions, it is excluded from the thesaurus section because it is not considered synonymous with American English

terms. This level of sensitivity to language variation is a hallmark of traditional linguistic resources but is often lacking in many Linked Open Data Lexical Resources.

BabelNet, however, shows greater sensitivity to language variation, allowing users to select dialects or older language varieties. For instance, in Italian, users can choose between dialects such as Lombard, Neapolitan, and Sicilian, while in English, one can even search for words in Middle English. Despite this, the General English descriptions in *BabelNet* do not adequately address the significant sociolinguistic variations in phonetics and lexicon, resulting in a mixed and unmarked representation that often fails to distinguish between American and British English varieties.

Non-LLOD resource	LLOD resource
<i>WordNet</i>	Open English Wordnet
<i>Wiktionary</i>	DBnary
Non synset and machine learning/embeddings dependent	
<i>ConceptNet</i>	

Table 6 LLOD and non LLOD resources taken into account for the observation

The representative of synset-based resources is Princeton *WordNet* as it represented the innovation and the bulwark upon which the largest part of current data modeling has been built. According to its authors, *WordNet* was compiled following psycho-linguistics principles and by posing the attention towards synchronic lexical-semantics to what is defined as psycholexicology. It represents a hybrid framework to organize units of meaning and model a lexical semantic network in a practical way. As defined by Millert et al., *WordNet* is a proposal for a more effective combination of traditional lexicographic information and modern high- speed computation. The *WordNet*’s synset was thought to be monolingual and aggregate synonyms in a single ‘node’. Miller defines that this organization makes words denotationally equivalent and can be substituted for one another in many, but not all, contexts. Formality relevance measures are absent also for lexical units in the synset-based resources such as *WordNet* and OpenEnglishWordNet (OEW).

Wiktionary is recognized as a collaboratively built resource intended to serve as a free, web-based alternative to traditional expert-built lexicons. From a lexicographic perspective, *Wiktionary* adopts a traditional descriptive approach, distinct from *WordNet*’s psycholinguistic model, allowing for a more nuanced description of each lexeme. This approach separates different formal representations of the same meaning, enabling greater specificity. As noted in the accompanying article, *Wiktionary* also embraces and explicitly reports linguistic labels related to domain, style, and register, thereby

providing pertinent information for each lexeme. In contrast, *DBnary*, the ontology-based representation of *Wiktionary* that adheres to the OntoLex model while extending it for specific purposes, does not include register or domain labels.

ConceptNet, defined as a semantic network (Speer et al., 2017), applies a one-lexeme, one-node structure similar to that of *Wiktionary*. It is designed to assist computers in understanding the meanings of words as they are used by people. Furthermore, *ConceptNet* is part of the Linked Open Data (LOD) ecosystem and benefits from an embedding-based structure that integrates distributional semantics with relational knowledge through *ConceptNet* Numberbatch (Speer et al., 2017), effectively bridging both approaches. However, like *Wiktionary* and *DBnary*, *ConceptNet*'s lexical descriptions lack measures of formality relevance.

When observing Linked Linguistic Open Data (LLOD) from a purely linguistic perspective, it becomes evident that no single resource can provide a comprehensive description to satisfy all linguistic needs. Typically, a thorough search across the entirety of lexicographic features necessitates the integration of information from multiple resources, each of which must be specifically understood for effective querying. Moreover, in the context of language, parameters such as formality, style, register, and domain cannot be overlooked. Despite this, LLOD represents an expanding movement that is invaluable for linguistic tasks, including supporting language learning and synonym retrieval. Adjustments are necessary to enhance the linguistic quality of data representation, increase completeness, and align digitally-based linguistic knowledge with traditional methodologies. This alignment will enable computer science to benefit from high-quality, linguistically validated resources while ensuring that linguists can rely on computer science contributions without encountering incongruities or inaccuracies (Esposito, 2024).

3.3.6.3 Formality and Lexicographic Resources

Formality in language is a complex and multifaceted phenomenon, encompassing lexical choice, syntactic structures, discourse conventions, and sociocultural context. Over the years, scholars have developed various approaches to assess and quantify formality, ranging from frequency-based statistical models to sophisticated machine learning algorithms. These methods are widely used in computational linguistics, psycholinguistics, and natural language processing (NLP) to analyze texts across different domains, including legal discourse, academic writing, social media communication, and conversational speech.

Linguists have traditionally linked formality to register variation, distinguishing between formal and informal language based on grammatical structure, vocabulary, and discourse markers. Halliday

(1978) described register as a function of field (what is being talked about), tenor (the relationship between interlocutors), and mode (spoken vs. written). In this framework, formal language typically exhibits:

Feature	Description	Linguistic Effect / Function
Nominalizations	Frequent use of nouns derived from verbs or adjectives (e.g., <i>development, analysis</i>).	Increases formality and abstraction by focusing on concepts rather than actions.
Passive Constructions	Higher occurrence of passive voice (e.g., <i>The data were analyzed</i>).	Shifts focus from the agent to the process or result, creating an impersonal tone.
Reduced Use of Personal Pronouns	Limited reliance on first- and second-person references (e.g., <i>I, we, you</i>).	Enhances objectivity and detachment, aligning with academic conventions.
Complex Sentences	Frequent use of subordinate and embedded clauses.	Allows for nuanced expression and logical connections between ideas.

Table 7 Key linguistic features characteristic of formal and academic language use.

These linguistic features form the basis of computational models for formality assessment. Beyond tracking semantic shifts over time, lexicographic resources must also account for the role of formality in lexical choices. Different contexts, such as academic writing, legal discourse, casual speech, and social media interactions, exhibit varying degrees of formality, influencing word selection and syntactic structures. Understanding and quantifying this variation is essential for applications in natural language processing (NLP), computational linguistics, and language pedagogy.

One of the most influential computational measures of formality was introduced by Heylighen and Dewaele (1999). Their formula quantifies formality (F) based on the relative frequencies of specific word categories:

$$F = \frac{\left(\begin{array}{l} \text{noun frequency} + \text{adjective frequency} + \text{preposition frequency} + \text{article frequency} - \\ \text{pronoun frequency} - \text{verb frequency} - \\ \text{adverb frequency} - \text{interjection frequency} + 100 \end{array} \right)}{2}$$

Formula 6 The formality score (F) as proposed by Heylighen and Dewaele (1999) quantifies linguistic formality by combining the relative frequencies of grammatical categories typically associated with formal versus informal registers. Higher values indicate greater formality in a text.

This metric ranges from 0% to 100%, with higher values indicating greater formality. The formula is based on the assumption that formal texts tend to use more nouns, adjectives, prepositions, and articles, while informal speech relies more on pronouns, verbs, adverbs, and interjections. Because it is based on word frequency distributions rather than sentence structure, this approach provides a flexible tool for comparing different genres and registers.

While Heylighen and Dewaele's model applies to entire texts, other scholars have explored formality at the lexical level. Brooke et al. (2010) proposed an alternative approach that assigns formality scores to individual words rather than analyzing an entire document's formality. This lexical approach enables finer distinctions, such as differentiating between near-synonyms like *commence* (formal) and *start* (neutral).

Building on this, Eder et al. (2021) developed a weighted lexicon for German that assigns an F-score to each lexical item. This measure is derived through a corpus-based methodology that integrates manual annotation, vector space modeling, and statistical text scoring. Unlike static dictionaries, corpus-based approaches provide dynamic insights into evolving linguistic trends, making them particularly valuable for tracking diachronic changes in formality levels.

Another significant application of formality metrics is in educational technology. Language learning platforms use formality assessments to guide learners in choosing appropriate vocabulary based on situational context. For example, an AI-driven writing assistant might suggest replacing *gonna* with *going to* in professional emails while preserving informal expressions in casual messages.

While computational models of formality have proven effective, they are not without limitations. One challenge is accounting for pragmatic factors, such as irony or cultural variation, which may not be fully captured by word frequency alone. Additionally, formality is often influenced by external social factors, including the relationship between speakers, power dynamics, and discourse conventions. Future research could enhance formality modeling by integrating contextual embeddings from large-scale language models like GPT or BERT, allowing for a more nuanced understanding of register variation.

As NLP and computational linguistics continue to evolve, future research on formality could explore deeper contextual analysis beyond lexical and syntactic features. One promising direction is the incorporation of transformer-based models, such as BERT, RoBERTa, and GPT, to develop context-aware formality detection systems. These models could analyze pragmatic cues, discourse coherence, and speaker intent, offering more precise distinctions between formal and informal registers.

Additionally, interdisciplinary research combining sociolinguistics and computational methods could provide richer insights into how formality operates in multilingual settings, cross-cultural communication, and online discourse. The increasing prevalence of code-switching in digital communication presents a challenge for current formality models, requiring innovative approaches that can dynamically assess mixed-language texts.

Moreover, as language evolves, formality itself may shift in response to changes in technology, social norms, and communicative preferences. This dynamic nature of linguistic formality underscores the need for adaptable models that can process emerging trends in real time. Further investigation is required to refine formality measurement techniques and integrate them into broader applications, ensuring their relevance in an ever-changing linguistic landscape.

In this context, ongoing advancements in AI-driven text analysis may bridge gaps between rigid formality scoring and the fluid nature of human communication. The integration of sentiment analysis, contextual embeddings, and real-time discourse adaptation could enable more sophisticated tools capable of adjusting language formality based on nuanced situational awareness. As research progresses, the intersection of human linguistic intuition and machine-driven assessment will likely redefine how we perceive and measure formality, opening new avenues for both theoretical exploration and practical applications.

Despite these advancements, our research and evaluation of Linked Open Data (LOD) resources for formality analysis remain unsatisfactory. Existing LOD datasets lack comprehensive coverage of formality variations across diverse registers and languages, limiting their applicability in computational models. Furthermore, inconsistencies in data annotation and the absence of standardized benchmarks hinder the reliability of LOD-driven approaches. Addressing these challenges requires the development of enriched, high-quality LOD resources that integrate contextual and pragmatic elements, ensuring a more robust foundation for future formality research.

3.3.7 LLOD Accessibility

In the context of Linked Linguistic Open Data (LLOD), accessibility via a SPARQL endpoint is primarily considered from the perspective of developers rather than end-users, such as learners. For developers, having access to a functional SPARQL endpoint is crucial for interacting with and leveraging the rich linguistic data available within the LLOD cloud. This access allows developers to explore, query, and utilize the data effectively, enabling them to design applications that can make the most of the vast linguistic resources hosted in the cloud. For our project, we specifically investigated

the accessibility of these resources, as it is vital for the development of tools that rely on LLOD data. While certain resources, such as those under the Corpora category, do not require a SPARQL endpoint, most other categories of LLOD resources should ideally be accessible via functional endpoints. However, as depicted in Figure 6, many LLOD resources with a SPARQL endpoint often fail to function properly, presenting significant challenges for developers. Furthermore, the amount of data accessible through a functional SPARQL endpoint is also a critical factor, with the volume typically measured in RDF triples. In our investigation, we compared the amount of data that should be accessible through SPARQL endpoints with the data that remains inaccessible due to non-functional endpoints, as shown in Table 1. This analysis highlights the ongoing need for better accessibility and functionality of SPARQL endpoints, as developers need reliable access to LLOD data to fully capitalize on the potential of the Semantic Web for linguistic applications. In the present project, we distinguish between resources that do not require a SPARQL endpoint, such as datasets in the Corpora category and ontologies listed under the *Other* category, and those that should be accessible via a functional SPARQL endpoint. For all categories except Corpora and Other, it is essential that resources are accessible through a SPARQL endpoint to maximize their utility. However, as shown in Figure 2, even when a SPARQL endpoint is provided, it often does not function properly in many cases, hindering the effective use of these resources. The amount of data in this context refers to the volume of information available in solutions published using Semantic Web technologies, typically measured in RDF triples. Figure 6 below compares the amount of data that could potentially be accessed through a functional SPARQL endpoint against the data that remains inaccessible due to non-functioning endpoints. This section presents the results of a qualitative analysis of 276 unique linguistic resources. Due to space limitations, we summarize the findings in terms of accessibility, the volume of data, and coverage of different languages, rather than presenting the raw results. Accessibility in terms of SPARQL endpoints, their status, the amount of data, and the language coverage of resources registered within the LLOD Cloud were automatically retrieved from the cloud. Resources described in peer-reviewed articles, but not registered in the LLOD Cloud, were manually assessed. Two independent reviewers searched for links to these resources within the articles or performed a web search using the resource names. If links were retrieved, their working status was verified to ensure the resources remained available, though the level of detail in the descriptions was not examined.

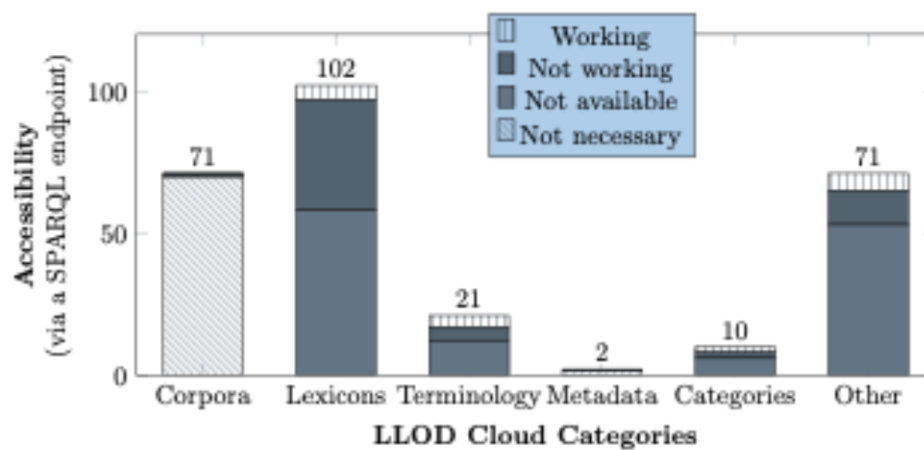


Figure 6 LLOD cloud resources accessibility for category

To verify the availability of SPARQL endpoints, the same procedure was applied: both the articles and the web were searched for the term *SPARQL* or *SPARQL endpoint*. When a potential endpoint was found, its accessibility was tested by attempting to open the link; however, no performance evaluation or query execution was performed. The language used to model each resource was manually extracted from the official descriptions in the LLOD Cloud, the articles, or the associated links. Regarding the amount of RDF triples, data was automatically retrieved for resources registered in the LLOD Cloud, while for others, the number of triples was extracted either from the articles, the resource links, or by querying the SPARQL endpoint, if available.

Linguistic resources were categorized according to the LLOD Cloud's classification system, which includes corpora, lexicons and dictionaries, terminologies, thesauri and knowledge bases, linguistic resource metadata, linguistic data categories, typological databases, and other types. This categorization and the comprehensive evaluation of accessibility are essential for understanding the extent to which LLOD resources are available for further development and utilization in various applications.

This section reports results of the qualitative analysis of 276 not duplicated linguistic resources. Due to space constraints, we opt for reporting summarized results in terms of accessibility, amount of data, and coverage of different languages rather than raw results. While accessibility in terms of SPARQL endpoint, its status, amount of data and language of resources registered to the LLOD Cloud are automatically retrieved from the cloud, we manually assessed those resources that are described in peer-reviewed articles without being registered to the cloud. As a result, resources in the LLOD Cloud are automatically assessed via Luzzu (Debattista et al., 2016). The manual assessment of the remaining resources has been performed by two independent reviewers that looked for links to the resource in the included articles or googling the name of the resource. If retrieved, links status has been verified and working status is reported if the resource is still available, without inspecting the

level of details used to describe the project or the resource. The same process has been followed to verify the availability of a SPARQL endpoint by both checking the included article and googling the name of resource attached to SPARQL or SPARQL endpoint. If the endpoint is successfully identified, its status has been verified in terms of reachability of the link, without testing its performance via a testing query. Concerning the retrieval of the language used to model each resource, it has been manually extracted by all the resources looking at the official description of the resource within the cloud, the article or the link describing the resource. About the number of triples contained in each resource, the evaluators automatically extracted it from the LLOD Cloud for those resources registered to it, while it has been extracted from the article, the link presenting the resource, or by querying the SPARQL endpoint, if available, for the others. Linguistic resources have been clustered as corpora, lexicons and dictionaries, terminologies, thesauri and knowledge bases, linguistic resource metadata, linguistic data categories, typological databases and other according to the LLOD cloud. All the resources published on the cloud are already identified as part of one of this cluster as a result of the publishing process, while a linguist hypothesizes the cluster for those resources that are described in the included articles but not registered in the cloud.

		Corpora	Lexicons and Dictionaries	Terminology, Thesauri, Knowledge Base	Linguistic Data Categories	Other
Total	resources	71	85	14	10	37
	resources with SPARQL endpoints	1	38	8	4	12
	potential triples from a SPARQL endpoint	2,000,000	3,524,654,263	32,369,655	11,688,500	10,031,008,889
Not-accessible	resources	1	34	5	2	10
	triples	2,000,000	3,109,247,394	21,310,206	3,000,000	9,917,629,432
Accessible	resources	0	4	3	2	2
	triples	0	415,406,869	11,059,449	8,688,500	113,379,457
	percentage	0	12	34	74	1

Figure 7 Amount of data in terms of RDF triples potentially accessible via a SPARQL endpoint

As reported in *The Linguistic Linked Open Data Cloud: Phenomenal Cosmic Powers... Itty Bitty Quality Space!* (Esposito et al., 2024) The distribution of Linguistic Linked Open Data (LLOD) resources is uneven across different categories. Most of the LLOD resources are found in corpora, lexicons, dictionaries, and similar datasets, with less than half associated with open licenses (OL). In some cases, outdated licenses like CC BY-NC 2.0 are still in use, rather than the updated CC BY-NC 4.0. This pattern is generally consistent across categories, although corpora show a relatively higher proportion of resources with OL (60%). In contrast, resources in linguistic metadata and typological

databases lack OL entirely. Furthermore, while 40% of resources offer a functional download mechanism, only 5% provide access via a working SPARQL endpoint. The situation deteriorates for openly accessible resources, dropping to 14% with any download capability, and to just four resources when SPARQL endpoints are required. For example, in the lexicons and dictionaries category, despite having 1 billion triples of data, only 322 million are downloadable, a mere 4 million are freely reusable, and none are accessible through SPARQL endpoints. This underscores the substantial untapped potential in LLOD due to limitations in accessibility and the rare use of open licenses.

For our purposes, ensuring both quality and availability of the LLOD resources is essential. While accessibility through SPARQL endpoints is necessary for interacting with the data, it is the combination of accessibility and the quality of the data that will truly enable the development of effective and impactful linguistic tools. The volume of data alone is not sufficient; the data must be accurate, well-structured, and reliable to be of real value to developers.

Quality data is essential for building meaningful linguistic applications, whether for natural language processing, linguistic analysis, or machine learning. Without high-quality, well-curated resources, even the most accessible endpoints may fall short in supporting sophisticated applications. For our project, the goal is to ensure that we not only have access to a large amount of data but that this data is structured, standardized, and have qualities that are indispensable when integrating multiple resources and building complex tools.

The challenges posed by non-functional SPARQL endpoints and inaccessible data underscore the need for improvements in both the accessibility and the integrity of LLOD resources. It is clear from our investigation that, while there is significant potential in the LLOD ecosystem, much of that potential remains untapped due to limitations in how the resources are made accessible. For developers to take full advantage of LLOD, the availability of reliable SPARQL endpoints is a minimum requirement; however, the quality of the data they provide must also meet a high standard to ensure that applications can operate efficiently and accurately.

As indicated by the present literature review and resource assessment, the volume of RDF triples available is often not the main issue; rather, it is the functionality of the endpoints and the overall reliability of the data that present significant challenges. In many cases, even when data is technically accessible, the lack of consistency and timely updates means that developers struggle to fully draw on the potential of LLOD data. Therefore, for the effective development of our tools, it is crucial that both data quality and accessibility be prioritized within the LLOD ecosystem.

Ultimately, improving the availability and functionality of SPARQL endpoints, alongside ensuring the quality of data, will unlock the true potential of LLOD resources, enabling the development of

sophisticated, reliable linguistic applications. For our project, the combination of these two factors is not just important, it is indispensable.

To further emphasize, without quality data, even the most accessible endpoints are rendered ineffective, leading to wasted efforts in development and hindering progress. In the context of our project, the value of LLOD resources lies not only in their accessibility but also in their ability to provide accurate, consistent, and comprehensive linguistic information. This is crucial for the development of applications that rely on high-fidelity data, such as language processing systems, corpus-based analysis tools, or even language learning platforms.

Moreover, the diversity of linguistic data sources in the LLOD cloud, ranging from corpora to ontologies, lexicons, and typological databases, requires a level of interoperability that can only be achieved when both the data's quality and the accessibility of the endpoints are optimized. The lack of effective integration between resources due to poor data quality or endpoint failures limits the ability to develop holistic linguistic applications. It also means that key linguistic resources, which could contribute to the advancement of tools in fields like machine translation or computational linguistics, remain underutilized.

The challenge, then, is twofold: ensuring that SPARQL endpoints are not only operational but that they are connected to high-quality data that accurately reflects the richness of linguistic resources. Our project seeks to address this by focusing on the identification of reliable resources and ensuring they meet the necessary standards for both accessibility and quality. As we continue to explore the landscape of LLOD, we must prioritize the identification of resources that offer both functional endpoints and high-quality, well-curated data to build effective linguistic tools.

In the broader context of LLOD development, the quality and availability of data cannot be seen as separate issues but must be tackled in tandem. Efforts to improve data quality, such as enhanced metadata, improved data curation, and ongoing resource maintenance, must go hand-in-hand with efforts to ensure that SPARQL endpoints remain operational and accessible. For developers, this dual focus will facilitate the creation of applications that can rely on stable, high-quality data, ultimately leading to more robust and innovative tools for linguistic analysis and language technology development.

3.3.8 Practical Application of LLOD

Linguistic Linked Open Data (LLOD) provides a structured, interoperable framework that can significantly enhance a wide range of natural language processing (NLP) tasks by integrating diverse linguistic resources. One primary area of application is word sense disambiguation (WSD), where LLOD facilitates the resolution of ambiguous terms by linking words to structured semantic

representations in knowledge graphs (KGs) and ontologies. Similarly, semantic similarity and word similarity tasks employ LLOD to compute relationships between linguistic units by utilizing linked lexical resources. Knowledge integration and ontology alignment benefit from LLOD's structured representation, allowing seamless merging of disparate knowledge bases for applications such as ontology extraction, knowledge graph (KG) completion, and ontology expansion (Jain et al., 2010; Sugimoto, 2022).

In machine translation and cross-lingual model transfer, LLOD supports language interoperability by providing multilingual linked resources, improving translation accuracy and adaptation to different linguistic contexts. Language detection and multilingual querying also benefit from structured linguistic data, enabling efficient cross-lingual search and retrieval. Text annotation and semantic mapping tasks use LLOD to enrich textual data with metadata, improving applications such as text enrichment, metadata reconciliation/harmonization, and text representation. In named entity recognition (NER) and entity linking, LLOD connects textual entities to structured knowledge bases, improving disambiguation and contextualization.

For question answering (QA) systems, LLOD enhances question interpretation, reasoning, and answer retrieval by aligning queries with structured knowledge, supporting fact-based QA, reasoning tasks, and dialogue generation. It also plays a role in text generation and data-to-text tasks, where structured data is transformed into natural language output. Corpus annotation and frame-based reasoning benefit from LLOD by enabling the enrichment of corpora with structured knowledge, facilitating NLP model training. In sentiment analysis, emotion recognition, and emotion analysis, linked data provides contextual sentiment lexicons and knowledge-based sentiment evaluation, enhancing affective computing.

Applications in lexical change detection and dictionary generation utilize LLOD to track linguistic evolution and automatically generate lexicons. Plagiarism detection and personal information extraction include structured linguistic resources to improve text analysis and entity recognition. LLOD also supports text-to-concept mapping, entity alignment, and relation extraction, which are critical for semantic search and KG construction. Knowledge extraction and KG navigation tasks utilize LLOD to retrieve and organize structured information efficiently. In domain-specific vocabulary creation and custom taxonomy development, LLOD helps build specialized linguistic resources tailored to specific fields.

Additional applications include OCR (Optical Character Recognition), where linked linguistic resources help correct and contextualize extracted text, KG materialization, which derives inferred knowledge from linked datasets, and cross-lingual querying, which facilitates multilingual information retrieval. LLOD also supports synonym lookup, interlinking, and encyclopedic

knowledge extraction, ensuring that NLP systems can effectively utilize structured knowledge. Narrativization of KGs converts structured knowledge into human-readable narratives, enhancing explainability. Finally, complexity assessment and semantic prediction tasks rely on LLOD to evaluate text complexity and predict semantic relationships.

Previous research underscores the significant impact of Knowledge Graphs (KGs) in Deep Learning models, particularly in Graph Neural Networks (GNNs), translational models, and explainable AI. Integrating computational measures with LLOD could enhance its effectiveness as a resource for linguistic research and optimize the use of KGs in training machine learning systems. In particular, the implementation of metrics-based labeled resources, such as those accounting for register variation, could improve the sensitivity of NLP tools that rely on KGs as training references. For instance, a resource labeled for formality based on speakers' perceptions could serve as a valuable benchmark, enabling AI systems to generate speech that aligns more closely with human linguistic expectations.

A natural intersection between linguistics, Natural Language Processing (NLP), the Semantic Web, and AI involves utilizing corpora as both qualitative and quantitative references to analyze and model the authentic reproduction of speech. Computational measures can be employed to extract weighted lexicons, integrating them into KG-based linguistic resources to capture the state of perception of lexical items. This approach moves beyond frequency analysis and word co-occurrence, allowing for a more nuanced representation of qualitative linguistic phenomena, such as those explored in the present study.

A key contribution to this area is found in *Comparative Probing of Lexical Semantics Theories for Cognitive Plausibility and Technological Usefulness* (Branco et al., 2020), which develops an experimental framework for evaluating different lexical semantics theories, including feature-based, inference-based, and co-occurrence-based approaches. One particularly influential lexical knowledge base is the Semantic Web of Words (SWOW), a crowdsourced resource that organizes approximately 12,000 words based on semantic relationships derived from word associations provided by 83,000 participants. Unlike other lexical resources that prioritize statistical co-occurrence or inference-based relationships, SWOW adopts a feature-based approach, representing lexical knowledge through semantic properties and associations. This methodology allows SWOW to offer valuable insights for both cognitive science and NLP, particularly in tasks involving semantic similarity and relatedness. The study demonstrates that feature-based lexicons, such as SWOW, outperform other models in these tasks and exhibit strong applicability in extrinsic NLP applications. Moreover, it highlights the discrepancy between intrinsic and extrinsic performance, showing that different embedding techniques are optimal for different tasks. The findings suggest that feature-based lexicons could

serve as core semantic knowledge bases and call for further exploration of cross-model conversions between various lexical resources (Branco et al., 2024).

An example of a versatile and expandable system for crowdsourcing language resources is presented in Rodosthenous et al. (2019). This architecture supports multiple user interfaces, enabling the collection and management of diverse linguistic resources and exercise types while incorporating mechanisms for evaluating and updating user responses. The initial prototype, a vocabulary trainer, has demonstrated promising results in preliminary tests using a Telegram chatbot and a Bootstrap web application. Future developments include formal evaluations and integration with the *Revita* platform (Katinskaia, 2018), along with the implementation of knowledge-based games, particularly those focusing on geography-related exercises.

In addition to this system, numerous advanced tools facilitate linguistic and knowledge management tasks. *Babelify* employs a graph-based approach to entity linking and word sense disambiguation, leveraging a densest subgraph heuristic to identify high-coherence semantic interpretations, and is accessible via a web interface or API. YAGO serves as a comprehensive knowledge base containing 50 million entities and 90 million facts, detailing relationships among real-world entities, and is available through a web interface and a SPARQL endpoint. Saffron enables multi-stage text analysis using modular NLP technologies, with accessibility through both web-based and command-line interfaces. PIKES extracts knowledge from text using a rule-based approach to semantic role labeling and is available via a web interface and a Java API. *Sematch* provides semantic similarity metrics within Knowledge Graphs (KGs), incorporating both taxonomic and statistical information, and is implemented in Python (Python Software Foundation, 2022). SEMAFOR delivers semantic analysis based on the Berkeley *FrameNet* framework for English sentences, available as standalone software. NAISC facilitates entity comparison and linking in graphs, implemented in Java. NASARI offers semantic vector representations for *BabelNet* and Wikipedia pages, aiding in semantic similarity and word sense disambiguation, and is provided as pre-trained vectors in W2Vec format. *SenseDef* is a multilingual corpus featuring disambiguated glosses from *BabelNet*, Wikipedia, and other sources, with extensive annotations across 256 languages. *LIdioms* models concepts as triples, incorporating multilingual translations and linking them to external resources. Lider focuses on building a Linguistic Linked Data cloud to enhance content analytics for multilingual and cross-media applications. Coh-Metrix analyzes texts based on linguistic and cognitive dimensions and is available as standalone software. Apertium is an open-source machine translation platform supporting various languages, with publicly available resources for community contributions. Lexvo.org provides structured data on languages and language-related entities within the Linked Data Web. FRED (Gangemi et al., 2017) automatically generates RDF/OWL ontologies and linked data from natural

language sentences, integrating Named Entity Resolution and Word Sense Disambiguation, and is accessible via an API and standalone code. COCO-EX (Becker et al., 2021) maps meaningful concepts from texts to *ConceptNet* nodes using relational information and is implemented in Python. Grapher (Melnik et al., 2022) constructs Knowledge Graphs by generating nodes from pre-trained language models and extracting features for edges, also implemented in Python.

The role of Linked Data in AI training is particularly significant for linguistic tasks, as it provides high-quality, interconnected data essential for processing diverse languages, dialects, and cultural variations. Linguistic Linked Open Data (LLOD), as a structured, machine-readable resource, enhances AI models by enabling them to learn from various linguistic features, including syntax, semantics, pragmatics, and sociolinguistic factors such as formality and regional variation. By integrating LLOD into AI workflows, language technologies can achieve greater accuracy, adaptability, and contextual awareness, ultimately improving applications across machine translation, natural language understanding, and information retrieval.

By integrating Linked Data, AI can go beyond simple word-level representations. Knowledge Graphs, a key component of Linked Data, connect language to broader ontologies, helping AI models grasp relationships between words, phrases, and concepts. Resources like *ConceptNet* and SWOW (Semantic Web of Words) provide rich semantic connections, enabling AI to infer meaning based on context and to model human-like understanding of language.

Moreover, Linked Data supports better multilingual and cross-cultural understanding. Resources like Lexvo.org, which provides information about languages and linguistic entities, allow AI to access detailed data about language variants, crucial for building models that can accurately process and generate text in different languages while respecting cultural nuances.

Incorporating LLOD into AI workflows enhances its ability to manage linguistic diversity in language learning and research. For example, leveraging metrics-based resources that account for linguistic features such as formality helps AI replicate human speech patterns more effectively. By using corpora and weighted lexicons, AI can extract deeper insights into speech authenticity and semantic similarity, reinforcing the value of LLOD in improving AI's understanding of linguistic nuances.

AI's dual role in language learning extends beyond technology application to recognizing and representing the unique grammatical structures, idioms, and cultural contexts of different languages. Languages with rich morphological or tonal systems require AI to not only analyze these subtleties but also generate appropriate responses. LLOD's structured, semantic data can be used to build more sensitive AI systems that address these linguistic variations, ensuring that AI tools can accurately reflect the complexity and richness of human languages.

AI's role in language learning transcends the mere application of technology; it demands a comprehensive understanding of the complexities inherent in linguistic diversity. Language is rich with various structures, idioms, and cultural contexts, and AI tools must be designed to accurately represent and process these unique linguistic features.

To address linguistic diversity effectively, it is imperative to recognize the unique characteristics of different languages. Each language exhibits distinct grammatical structures, vocabulary, and phonetic nuances that AI systems must capture (Helm et al., 2023). For example, languages with rich morphological systems require AI to analyze and generate forms that reflect various inflections or derivations. Similarly, languages that utilize tonal distinctions need AI algorithms capable of recognizing and producing these subtleties for effective communication.

Equitable access is another critical aspect of integrating AI into language learning. It is vital that AI tools support all learners, irrespective of their linguistic backgrounds. A failure to consider linguistic diversity in AI design can lead to biased outcomes, where certain languages or dialects are inadequately represented. This inequity may hinder the learning experiences of students who speak underrepresented languages or dialects, exacerbating existing disparities in educational contexts.

Additionally, understanding the cultural contexts that shape language use is essential. AI systems should be sensitive to regional variations and the cultural significance of language, ensuring that learners receive relevant and contextually appropriate feedback. This cultural awareness enhances the learning experience and fosters inclusiveness and respect for linguistic diversity. By developing AI tools adaptable to various languages and their unique characteristics, educators can create more inclusive and equitable learning environments that empower all learners to succeed.

As mentioned in 2.4, achieving this level of inclusivity requires AI systems in language learning to employ advanced computational approaches that deliver adaptive and personalized experiences. From a technical perspective, two key methodologies underpin this process: Computer-Assisted Language Learning (CALL) and Intelligent CALL (ICALL). While CALL focuses on integrating computational tools into educational contexts, ICALL extends this framework by incorporating Natural Language Processing (NLP) techniques to enhance language acquisition. Components such as tokenizers, morphological analyzers, and syntactic parsers are instrumental in processing linguistic structures, enabling systems to provide immediate, context-aware feedback that enriches the learner's experience.

However, despite the potential of these tools, bias within AI systems remains a significant challenge. When AI fails to represent certain languages or cultural contexts adequately, it can lead to unequal learning opportunities. Misalignment between a learner's linguistic background and AI-generated content can hinder comprehension, reduce engagement, and negatively impact confidence and

motivation, further highlighting the importance of developing AI tools that are sensitive to linguistic diversity.

Furthermore, educational inequities can be exacerbated by the presence of bias in AI tools. When assessments and instructional materials are not inclusive of various linguistic contexts, learners from underrepresented backgrounds may be unfairly evaluated. This misinterpretation of capabilities can lead to skewed perceptions of a learner's potential, adversely affecting their academic trajectory, self-esteem, and access to further educational opportunities.

Moreover, limited exposure to languages can occur when AI tools prioritize certain languages or dialects over others. This restricted access stifles learners' ability to engage with varied linguistic inputs, which is crucial for developing comprehensive language skills. When AI systems fail to resonate with the cultural and linguistic experiences of diverse learners, it can lead to reduced engagement, as learners may feel the content does not reflect their realities, resulting in diminished participation and investment in their learning journey (Helm et al., 2023).

To address AI bias in language learning, a multifaceted approach is essential. One foundational step is the development of inclusive data sets that reflect a diverse array of languages and dialects. This involves not only including widely spoken languages but also recognizing and incorporating underrepresented ones. Ensuring that data captures various linguistic features and cultural nuances allows AI systems to better serve learners from different backgrounds, promoting a more inclusive learning experience.

Regular audits of AI tools are crucial for identifying and rectifying biases. These evaluations should assess AI performance across different linguistic contexts and learner demographics. By analyzing how AI tools respond to varied inputs, educators and developers can uncover biases and make informed adjustments, fostering an adaptive learning environment. Community engagement also plays a significant role in this process. Actively involving linguistic communities in the design and development of AI tools ensures that diverse linguistic needs are met. Collaborating with educators, linguists, and learners from various backgrounds provides valuable insights into the unique challenges faced in language learning.

Bias awareness training for educators and developers is vital. Training programs can enhance understanding of AI bias, its implications, and its impact on learners. By increasing awareness, stakeholders can remain vigilant in their efforts to create inclusive AI systems. Establishing robust feedback mechanisms is equally important, allowing learners and educators to report biases or inaccuracies they encounter while using AI tools. Open channels for feedback provide AI developers with insights directly from users, enabling necessary adjustments to enhance system performance and user experience.

In today's technologically advanced educational landscape, AI literacy is becoming increasingly essential for both educators and learners. As AI technologies become more integrated into language learning, understanding their functionality and implications is crucial for navigating these tools effectively. Fostering AI literacy offers numerous benefits: it enhances engagement, encourages critical thinking, and equips learners to question and analyze AI-generated outputs. This empowerment enables users to make use of AI technologies more effectively in their learning and teaching processes (cf. Cinganotto, 2025; Cinganotto et al., 2024).

To cultivate AI literacy among learners and educators, integrating dedicated training on AI concepts into language curricula is vital. These modules should cover fundamental principles of AI, practical applications in language learning, and the importance of recognizing biases in AI systems. Professional development for educators is equally crucial. Implementing training programs that equip teachers with the skills necessary to effectively use AI tools in the classroom can significantly enhance teaching methodologies. Workshops focused on AI tools and pedagogical strategies can help educators create a balanced approach that combines traditional and AI-driven instruction.

When AI systems are trained on imprecise or incomplete data, the consequences can be significant, particularly in language learning. AI models rely on the quality and accuracy of the data they are exposed to, and if the data is biased, outdated, or lacks representation of diverse linguistic and cultural contexts, the output generated by these systems can be unreliable. Imprecise data leads to misinterpretations of language structures, meaning, and usage, resulting in inaccurate translations, poor grammar suggestions, or culturally irrelevant content.

In language learning, these inaccuracies can cause confusion for learners, especially when AI-generated feedback contradicts the linguistic rules or norms of their native language or target language. For instance, learners might receive feedback that fails to account for the complexities of their language, such as unique grammatical structures or regional dialects. This misalignment can hinder their progress, reduce confidence, and, in some cases, perpetuate harmful stereotypes or biases. Additionally, AI tools trained on imprecise data may fail to provide equitable learning opportunities for all, disproportionately affecting students who speak underrepresented or minority languages, further widening the educational gap.

In the long term, AI systems with imprecise training data can erode trust in these technologies, leading educators and learners to question the reliability of AI-assisted language learning tools. It underscores the need for rigorous data selection, regular audits, and a commitment to incorporating diverse and accurate linguistic inputs into AI models to ensure they provide meaningful, contextually appropriate, and fair learning experiences.

3.3.8.1 Word-sense disambiguation

For the purposes of the present study, we focused on leveraging Linked Open Data (LOD) as a core resource for Word Sense Disambiguation (WSD). WSD is a critical task in natural language processing, aiming to resolve ambiguities in word meanings based on context. Given the richness and interconnectivity of LOD, we found it to be an invaluable resource for this task. By incorporating structured semantic knowledge from widely used ontologies and knowledge graphs, we were able to enhance the precision and contextual understanding required for effective disambiguation. This section discusses the methodology behind utilizing LOD for WSD, detailing the integration of linguistic resources and the advantages of using linked data for resolving semantic ambiguities in our approach. We compare our proposed *WordNet*-based disambiguation method with several well-established knowledge-based systems that serve as benchmarks in word sense disambiguation (WSD). These systems employ diverse techniques, ranging from gloss overlap to graph-based models, offering valuable insights into the strengths and limitations of different disambiguation strategies.

One of the earliest and most widely recognized methods for Word Sense Disambiguation (WSD) is the Lesk algorithm (Basile et al., 2014a), which determines the sense of a word by measuring the overlap between its definition and the surrounding context. Specifically, it compares the words in the word's definition (or gloss) with those in the surrounding context of the word, selecting the sense with the highest overlap. While effective in certain straightforward cases, the Lesk algorithm has several limitations. For instance, it struggles when context words are synonyms or semantically related but do not match the exact lexical terms in the word's definition. This issue arises because the Lesk algorithm relies heavily on exact word matches, making it less effective in handling more subtle semantic relations such as synonymy, hypernymy, or hyponymy (Basile et al., 2014b).

To address these limitations, the *Leskext* algorithm was developed as an improvement over the original Lesk algorithm. *Leskext* extends the basic Lesk approach by incorporating related terms from a Knowledge Graph (KG), which significantly broadens the semantic scope beyond the immediate lexical overlap. By leveraging KGs, *Leskext* can account for a wider range of semantic relationships and includes terms that might not appear in the word's gloss but are contextually relevant. Additionally, *Leskext* uses Term Frequency-Inverse Document Frequency (TF-IDF) to enhance the computation of similarity between the expanded gloss and the surrounding context. This weighting scheme helps prioritize more informative words in the context, improving the overall disambiguation process by reducing the influence of common, less meaningful terms.

An even further refinement, *Leskext+emb*, integrates Latent Semantic Analysis (LSA) to compute semantic similarity in vector spaces, moving beyond the limitations of the bag-of-words model that

was the foundation of the earlier Lesk-based approaches. LSA enables the model to capture deeper semantic meanings by mapping words to high-dimensional spaces based on their usage in large corpora, making it more capable of handling synonyms, polysemy, and other subtle linguistic phenomena. Additionally, *Leskext+emb* introduces the concept of Inverse Gloss Frequency (IGF), which replaces the traditional Inverse Document Frequency (IDF) used in TF-IDF. By using IGF, the model incorporates distributional semantics into the disambiguation process, which allows it to more effectively identify synonyms and related words that may appear in different contexts but share underlying meanings.

Despite these advancements, Lesk-based methods, including *Leskext* and *Leskext+emb*, still rely on term overlaps as a fundamental mechanism for disambiguation. This reliance on term overlaps means that these methods can struggle in cases where synonymous terms do not explicitly appear in the context or gloss. Furthermore, these models fail to account for word order, a crucial aspect in many languages where the arrangement of words can significantly affect meaning. For example, in languages with rich syntactic structures, the placement of a word within a sentence can alter its interpretation, making it essential for WSD systems to consider contextual ordering.

In contrast to these overlap-based methods, the UKB system takes a more sophisticated, graph-based approach. Instead of relying on word definitions and context overlap, UKB operates over the entire *WordNet* graph, utilizing PageRank to determine the most appropriate sense. The UKB system constructs a subgraph that includes not only the target word's senses but also nodes representing the context, typically drawn from a sentence or a short passage. This graph-based approach enables the system to grasp the rich relational structure of *WordNet*, considering not only the direct connections between senses but also their broader contextual relationships. By applying personalized *PageRank*, UKB selects the sense that is most contextually appropriate based on its position and relative importance within the graph. This structure-aware method is an improvement over traditional overlap-based methods because it considers the global structure of word senses, which can provide more nuanced and accurate disambiguation in complex contexts.

Further extending the capabilities of UKB, the UKBgloss extension enhances the PageRank-based approach by converting the glosses of *WordNet* senses into disambiguated synsets and incorporating frequency information to better initialize the context words. This extension improves the precision of the disambiguation process by using contextual information to select the most relevant synsets for the given task. The latest version, UKBgloss18, optimizes the parameters used in the model, including a larger context window of over 20 words and 30 iterations of the personalized PageRank algorithm, allowing for deeper and more accurate contextual understanding. However, despite these advancements, UKB's reliance on an undirected graph remains a limitation. The lack of directionality

between nodes may weaken the effectiveness of the PageRank algorithm, particularly in cases where directional relationships between words and senses are essential for precise disambiguation. In cases where the meaning of a word depends heavily on its syntactic or semantic directionality, such as verb-object relationships or cause-effect chains, UKB's undirected graph may not fully capture the complexity of the relationships between words.

In future iterations of WSD systems, addressing the limitations of graph-based methods such as UKB could involve incorporating directed edges in the knowledge graph, allowing for more accurate representations of the relationships between words and senses. Additionally, integrating deep learning models, such as Transformers or BERT-based models, with graph-based methods could potentially overcome the issues of word order and context dependency. These models, with their ability to understand complex syntactic and semantic dependencies, could provide further enhancements to traditional graph-based methods by integrating word order and contextual nuances in a way that is not possible with current approaches (Popov et al., 2019).

Babelfy is another notable system that integrates entity linking and Word Sense Disambiguation (WSD) through a graph-based approach. Specifically, *Babelfy* employs a random walk with restart algorithm applied to *BabelNet*, a multilingual semantic network that combines *WordNet*, Wikipedia, and *Wiktionary* into a comprehensive resource. Unlike traditional methods that analyze words in isolation, *Babelfy* processes entire documents, capturing a broader and more robust contextual understanding. This holistic approach is particularly effective in multilingual scenarios, where the system can leverage *BabelNet*'s rich, cross-lingual semantic data. However, similar to PageRank-based methods, *Babelfy* faces challenges when applied to *WordNet* alone, primarily due to its reliance on the densest subgraph heuristic. This method can be problematic in cases of polysemy, where multiple senses of a word exist within the same context, as the heuristic may fail to disambiguate among competing meanings effectively.

In contrast to *Babelfy*'s document-level, graph-based strategy, the WSD-TM system takes a different approach by employing topic modeling (Chaplot & Salakhutdinov, 2018; Moro et al., 2014). This method, based on a variation of the Latent Dirichlet Allocation (LDA) algorithm, disambiguates words at the document level by representing text as distributions over synsets (sets of synonymous words). WSD-TM benefits from a global contextual perspective, analyzing the overall themes or topics within a document. While this broader view is advantageous in many cases, it may struggle when fine-grained, local context is needed, particularly in instances where multiple senses of a word appear within the same document or sentence. The method's reliance on document-level topics may overlook subtle contextual variations that could play a critical role in accurate disambiguation.

Building on the strengths of existing approaches, our proposed method introduces a context-aware WSD system that offers more precise disambiguation by integrating a wider range of semantic relations within *WordNet*. Unlike traditional methods such as Lesk and UKB, our approach goes beyond simple sense matching by considering various semantic relationships, including synonyms, hypernyms, hypernyms, meronyms, holonyms, and antonyms. This rich integration of semantic relations ensures that the system can capture not only the word’s direct meaning but also its broader conceptual network, offering a more nuanced understanding of context. For example, our disambiguation function first uses the Lesk algorithm (via the `simple_lesk` function) to select the most relevant sense of a word in a sentence. This initial sense selection is then enhanced by identifying related concepts that enrich the disambiguation process, allowing the system to resolve ambiguities more effectively. By incorporating these semantic relations, we ensure that our system is not only considering the immediate context but also the broader semantic landscape, making it more robust in handling complex linguistic phenomena like polysemy and context-dependence. Additionally, our method constructs a relation graph that visually represents the connections between the word, its senses, and related concepts. The graph includes:

Node Type	Description	Function in the Network
Main Node	Represents the target word or lexical item being analyzed.	Serves as the central entry point from which semantic relations are derived.
Definition Node	Contains the disambiguated sense of the target word, typically drawn from a dictionary or lexical database.	Provides semantic grounding by specifying the intended meaning of the word.
Related Concept Nodes	Represent semantically associated words such as synonyms, hypernyms, hyponyms, meronyms, holonyms, and antonyms.	Establish labeled semantic relations connecting each concept to the main node, enabling structured representation of lexical relationships.

Table 8 Node types and relational structure in a lexical-semantic network representation.

This relation graph not only aids in disambiguation but also provides a visual representation of the semantic network surrounding the target word. For example, if the word *bank* refers to a financial institution, the graph will display connections to concepts like *money*, *investment*, and *economy* as hypernyms, along with terms such as *ATM* or *credit* as related concepts. Here’s a snippet of the disambiguation process implemented in our system:

```
def disambiguate(sentence, word):
```

```

sense = simple_lesk(sentence, word)

if sense:
    synonyms = get_synonyms(sense)
    hypernyms = get_hypernyms(sense)
    hyponyms = get_hyponyms(sense)
    meronyms = get_meronyms(sense)
    holonyms = get_holonyms(sense)
    antonyms = get_antonyms(sense)

    words = [
        {id: f{word.replace(' ', '_')}_main, label: word, isCore: True},
        {id: fdefinition_0, label: sense.definition(), isCore: False}
    ]
    links = [
        {source: f{word.replace(' ', '_')}_main, target: fdefinition_0, relation:
definition}
    ]

    def add_relations(related_words, relation_type):
        relation_labels = {
            synonym: also known as,
            hypernym: is part of,
            hyponym: has a,
            meronym: is made of,
            holonym: contains,
            antonym: opposite
        }
        for i, related_word in enumerate(related_words):
            related_word_id = f{related_word.replace(' ', '_')}_related_{relation_type}_{i}
            words.append({id: related_word_id, label: related_word, isCore: False})
            relation_label = relation_labels.get(relation_type,
relation_type.capitalize())
            links.append({source: fdefinition_0, target: related_word_id, relation:
relation_label})

    add_relations(synonyms, synonym)
    add_relations(hypernyms, hypernym)
    add_relations(hyponyms, hyponym)
    add_relations(meronyms, meronym)
    add_relations(holonyms, holonym)
    add_relations(antonyms, antonym)

    return {
        'sense': sense.name(),

```

```

        'definition': sense.definition(),
        'examples': sense.examples(),
        'synonyms': synonyms,
        'hypernyms': hypernyms,
        'hyponyms': hyponyms,
        'meronyms': meronyms,
        'holonyms': holonyms,
        'antonyms': antonyms,
        'relation_graph': {
            'words': words,
            'links': links
        }
    }
else:
    return {'error': 'No sense found'}

```

Table 9 Example of semantic relations generated by the disambiguate() function. The table lists the target word, its identified sense (based on the Simple Lesk algorithm), and the corresponding semantic relations, synonyms, hypernyms, hyponyms, meronyms, holonyms, and antonyms, used to construct the relation graph

The Lesk Algorithm was one of the first successful algorithms for Word Sense Disambiguation (WSD) and remains a foundational method in the field. The core idea behind Lesk's approach is relatively simple: it attempts to select the most appropriate sense of an ambiguous word based on the overlap between the definition (or gloss) of each sense of the word and the surrounding context in which the word appears. The Lesk algorithm works by first identifying all the possible senses of the ambiguous word from a lexical resource such as *WordNet*, where each sense is associated with a gloss (a textual definition). For each sense, the algorithm then compares the gloss of that sense with the context in which the target word appears, measuring the overlap between the words in the gloss and those in the context. This overlap is often calculated by counting how many words are shared between the gloss and context. The sense with the maximum overlap is chosen as the correct sense of the word. For example, consider the word *bank*. The two senses in *WordNet* might be: 1) a financial institution and 2) the side of a river. The algorithm will compare the glosses of these two senses with the surrounding context (e.g., *He sat on the bank of the river*). If the gloss of the second sense has a higher overlap with the context (because it mentions the word *river*), this sense will be selected. While effective in simple cases, the Lesk algorithm has several limitations. First, it relies on exact word matches, meaning it can fail when synonyms or related terms appear in the context. It also does not consider the order of words or semantic relationships between them, which can lead to inaccuracies in more complex sentences. To address these issues, several extensions of the Lesk algorithm have

been proposed. *Leskext* extends Lesk by incorporating additional information from a Knowledge Graph (KG), such as *WordNet*, and uses Term Frequency-Inverse Document Frequency (TF-IDF) to compute the similarity between the gloss and context. *Leskext+emb* (Huang et al., 2020) further improves the algorithm by incorporating Latent Semantic Analysis (LSA), using semantic vector spaces instead of simple word overlap to calculate the similarity, which allows the algorithm to better handle synonyms and polysemy. Some adaptations also use cosine similarity between the vector representations of the gloss and context, overcoming the word-matching limitation of the original algorithm. The Lesk Algorithm remains a cornerstone in the WSD field due to its simplicity and intuitive approach, though its performance can be enhanced significantly by incorporating more advanced techniques like semantic analysis and graph-based reasoning.

In our approach, we recognize that the words selected by users for disambiguation often require improvement and enhancement to achieve accurate results. Words taken out of context can lead to misinterpretation and incorrect sense assignments. To address this, our method ensures that each word is disambiguated in relation to the specific sentence in which it appears, taking into account the full context surrounding the target word. By doing so, we can significantly enhance the quality and precision of word sense disambiguation (WSD) for each individual unit. This is critical because understanding the meaning of a word is deeply tied to its context within a sentence.

One of the key advantages of our approach is its ability to provide a more accurate sense assignment by leveraging the surrounding sentence context. This improves the reliability of the disambiguation process, as it eliminates the ambiguity that can arise from relying solely on individual words or static definitions. Additionally, our method enhances the system's robustness, ensuring that even when faced with complex or ambiguous sentences, it can effectively differentiate between multiple meanings of a word. This adaptability to varied linguistic contexts makes the system reliable across different domains and applications.

Furthermore, our approach does not rely purely on gloss overlap or graph traversal algorithms, which can be limited in certain scenarios. Instead, we integrate a variety of semantic relations within *WordNet*, such as synonyms, hypernyms, hyponyms, meronyms, holonyms, and antonyms, which enrich the disambiguation process. By considering these relationships, we can achieve a more nuanced understanding of word meanings and resolve ambiguities more effectively. The incorporation of these relations ensures that our system is not only more accurate but also more reliable, as it can handle a wider range of linguistic phenomena, including polysemy and synonymy. For instance, in our system, when a user provides a sentence and a target word for disambiguation, the system analyzes not only the word itself but also its position and relation to other words within the sentence. This contextualized approach allows us to disambiguate the word by comparing its

possible senses in *WordNet* to the surrounding linguistic cues, improving the accuracy of sense assignment. The following code snippet illustrates how our endpoint processes the user's request and performs disambiguation:

```
def disambiguate_endpoint():
    data = request.json
    sentence = data.get('sentence')
    word = data.get('word')

    if not sentence or not word:
        return jsonify({'error': 'Sentence and word are required'}), 400

    result = disambiguate(sentence, word)
    return jsonify(result)
```

Table 10 Structured representation of the semantic relations extracted through the `disambiguate_endpoint()` API. Each entry lists the target word, its definition, and the related lexical items as returned by the Word Sense Disambiguation process.

In this implementation, the endpoint accepts a POST request with the sentence and word, processes the input using the `disambiguate()` function, and returns the disambiguation result. This process ensures that each target word is assessed and disambiguated in the proper context, making our method robust and adaptable to various real-world scenarios where words may have multiple meanings depending on their surrounding text.

3.3.9 Final evaluation and state of the art

The present study has sought to evaluate the state of the art in linguistic data resources and their applicability to the development of adaptive, data-driven language-learning environments. The analysis, drawn primarily from Scopus and conducted through a qualitative lens, inevitably reflects certain limitations. Future work should extend this investigation both qualitatively and quantitatively, incorporating a wider range of databases, languages, and data quality dimensions. At present, the focus has been placed on English-language resources and the coverage of linguistic features, rather than on systematic comparisons of data quality across different knowledge graphs (KGs). As emphasized in the FAIR framework (Wilkinson et al., 2016), overcoming barriers of accessibility and quality will require the concerted effort of all stakeholders, researchers, repositories, and developers alike.

The primary objective of this review was to assess the potential of Linguistic Linked Open Data (LLOD) in supplying reliable lexical and grammatical input for language-learning applications. Previous research has shown that large-scale linguistic corpora, when combined with machine

learning and natural language processing (NLP) techniques, can yield valuable insights into learner behavior, error patterns, and developmental trajectories (Granger & Leech, 2014; Meurers & Dickinson, 2017). These computational approaches have not only strengthened theoretical frameworks in Second Language Acquisition (SLA) (Gass & Selinker, 2008) but also paved the way for adaptive, data-informed educational systems (Tharp, 2020). Within this broader context, the present evaluation highlights both the promise and the shortcomings of existing open linguistic resources.

While LLOD offers significant advantages, expert curation, open accessibility, and flexible licensing, our findings reveal substantial gaps in coverage, uniformity, and practical accessibility. Many datasets remain unavailable through functional download mechanisms, and only a small subset can be queried via SPARQL endpoints, leaving large volumes of valuable data effectively inaccessible. These constraints underscore the central paradox of LLOD: its conceptual richness and potential are considerable, yet its practical deployment is hindered by infrastructural and licensing barriers. To address these challenges, this project adopted Princeton *WordNet* as a primary lexical reference, complemented by Wikipedia as a dynamic, crowdsourced textual repository. Together, these resources provide both the structured lexical backbone and the contextual diversity necessary for meaningful language learning.

Despite the limitations of the current LLOD landscape, the analysis confirms the considerable pedagogical value of knowledge graphs (KGs) as frameworks for structuring, interlinking, and visualizing linguistic data. Their ability to model lexical networks, semantic hierarchies, and usage patterns makes them particularly well suited for supporting second language acquisition. Moreover, integrating semantic networks and network analysis techniques opens new perspectives for cognitive and didactic applications, enabling the visualization of lexical growth, the identification of semantic clusters, and the mapping of language development in learner-centered contexts.

To this end, the present study incorporated direct visualizations of synonym and semantic networks derived from KG-based resources. These representations encourage learners to explore linguistic structures actively, fostering deeper comprehension and metalinguistic awareness. As supported by studies such as Wolter (2001), synonym-based learning strengthens semantic connections and enhances vocabulary retention. Embedding such strategies within a digital environment positions language-learning systems not merely as repositories of information but as interactive spaces where learners can navigate and internalize the relational nature of language.

In conclusion, while the current LLOD ecosystem remains uneven and fragmented, it nonetheless provides a valuable foundation for the development of innovative educational technologies. The integration of structured lexical databases with dynamic textual repositories demonstrates both

feasibility and pedagogical potential, provided that issues of accessibility, licensing, and technical infrastructure are systematically addressed. More importantly, the exploration of KG-based approaches points toward the future of language learning: from static, isolated resources to adaptive, interconnected, and cognitively informed systems capable of reflecting the complexity and dynamism of real-world linguistic experience.

3.4 Conclusion

In this chapter, we have sought to delineate the principal methodological approaches that inform both didactic theory and computer-assisted language learning. Our objective has been to identify and critically assess the reference frameworks capable of supporting the coherent integration of pedagogical principles with technological innovation. By examining these foundational perspectives, we have established the conceptual and methodological groundwork required for the development of an effective adaptive environment for second language acquisition.

This analytical effort constitutes a preparatory phase aimed at equipping subsequent chapters with the theoretical and practical tools necessary for the design and implementation of a tangible, pedagogically informed, and technologically robust system for second language learning. Through this alignment of didactic methodology and computational design, the ensuing sections will translate theoretical insight into a practical framework capable of supporting authentic and sustainable linguistic development.

4 Appil: Bridging the Gap in Adaptive Language Learning

4.1 Introduction

Building on the theoretical and technological foundations established in Chapter 1, 2, 3, the present research turns to the development of Appil, a newly created system designed and supervised within the framework of this PhD project. Appil was conceived to address some of the gaps identified in existing CALL and ICALL applications (as discussed in Chapter 2). Unlike many current platforms that prioritize general-purpose solutions and rely on AI-driven interventions whose internal mechanisms often remain opaque, Appil proposes a more linguistically transparent approach. It seeks to integrate dynamic adaptation, lexical organization, and cognitive-load management through methodologies grounded in linguistic description and empirical research. In this respect, it draws on insights from Second Language Acquisition (SLA) theory to outline a potentially more pedagogically principled and learner-oriented model for technology-enhanced language learning.

The rationale for Appil arises from two persistent limitations observed in current CALL systems. First, personalization often remains superficial: while some platforms may adjust task difficulty or recycle items based on performance, they seldom model a learner's internal lexical organization, the structuring of vocabulary knowledge that supports communicative competence. Second, cognitive-load considerations are frequently overlooked, leading to input that may exceed learners' processing capacity or fail to provide sufficient scaffolding in line with Nation's (2014) recommendation of maintaining unknown lexical items below a 2% threshold. These limitations may hinder efficient vocabulary growth and the development of phraseological competence consistent with authentic language use.

Appil aims to respond to these issues by combining linguistic analysis with adaptive strategies informed by empirical findings. Rather than relying on opaque AI procedures, it employs transparent text-processing operations, tokenization, sentence splitting, part-of-speech tagging, syntactic parsing, and lexical-frequency profiling, to adapt content dynamically to the learner's evolving profile. In this way, feedback and materials may remain pedagogically grounded and cognitively sustainable. Furthermore, by embedding lexical and phraseological units as key learning targets, Appil seeks to support not only vocabulary expansion but also collocational and idiomatic awareness, aligning with the communicative and pragmatic dimensions of SLA.

Equally central to Appil's design is its emphasis on adaptive progression. The system is intended to evolve alongside the learner, updating its representation of lexical competence and adjusting input accordingly. This process may foster an environment in which learners encounter language at an appropriate level of challenge, novel enough to stimulate development, yet sufficiently accessible to prevent overload or disengagement. In this sense, Appil functions less as a fixed exercise generator and more as a responsive framework that may scaffold learners from basic lexical thresholds toward more advanced phraseological competence.

In positioning Appil within the wider CALL landscape, this chapter suggests that lexical organization and cognitive-load management constitute complementary dimensions of adaptive language learning. By adopting a lexical perspective and foregrounding phraseology in meaningful contexts, Appil aspires to mirror authentic communicative practices while encouraging greater learner autonomy. Grounded in SLA theory and lexicographic resources, it thus illustrates how linguistic transparency, pedagogical soundness, and adaptive technology may intersect within a unified design.

The following sections outline Appil's theoretical underpinnings, design rationale, system architecture, and pedagogical implications, indicating how the project seeks to address long-standing gaps in CALL/ICALL research while exploring a direction for adaptive, learner-centered language technology.

4.2 The Design Philosophy Behind Appil

The philosophy behind Appil rests on the conviction that language learning technology must be more than an attractive digital interface; it must embody the principles of Second Language Acquisition (SLA) while remaining usable, transparent, and responsive. Where earlier CALL systems often fell short, whether through static content, superficial personalization, or neglect of cognitive and linguistic realities, Appil aims to provide a coherent solution that balances theory and practice. Its design is guided by three overarching principles: dynamic adaptation, lexical centrality, and cognitive sustainability.

First, Appil is built to be dynamically adaptive, evolving continuously with the learner. Unlike static platforms that provide identical tasks regardless of progress, Appil utilizes fine-grained tracking of vocabulary knowledge, lexical diversity, and phraseological competence to calibrate materials in real time. By aligning content to the learner's current profile, it ensures that input reflects Krashen's principle of $i+1$ (Krashen, 1982) and Nation's lexical thresholds (Nation, 2001, 2006, 2017). Adaptivity in Appil is not a vague promise but a concrete mechanism: text selection, annotation, and task generation adjust as soon as new vocabulary is consolidated, or old items require recycling. This

guarantees a learning path that is both progressive and responsive, avoiding the twin pitfalls of redundancy and overload.

Second, Appil adopts a lexical-first orientation. Drawing on Nation's emphasis on vocabulary and collocations as the backbone of communicative competence (Nation, 2001, 2013), the system treats lexical organization as its primary unit of analysis. Words are not presented as isolated items but as part of larger phraseological and discourse structures, reflecting the view that formulaic sequences and collocations play a central role in fluent, natural communication (Wray, 2002; Schmitt, 2010). Learners encounter collocations, idiomatic expressions, and register-sensitive forms in context, enabling them to move beyond recognition toward authentic usage. This lexical focus also allows for more precise differentiation between learners: two students with similar grammatical knowledge may still differ markedly in their phraseological competence, an aspect of proficiency often overlooked by traditional assessment tools but increasingly recognized in cognitive and usage-based models of language learning (Ellis, 2001). Appil captures such distinctions through data-driven tracking of lexical and collocational development, aligning with contemporary perspectives on the role of vocabulary in second language acquisition.

Third, Appil prioritizes cognitive sustainability, ensuring that learning remains engaging without becoming overwhelming. Insights from Dual Process Theory (Evans, 2008) and cognitive neuroscience inform its pacing strategies: difficult items are recycled across contexts to reduce processing load, while linguistic annotations and feedback provide opportunities for explicit reflection without disrupting implicit acquisition (DeKeyser, 2007; Evans, 2008). By carefully managing the proportion of unfamiliar vocabulary and gradually layering syntactic complexity, Appil maintains what Sweller's Cognitive Load Theory identifies as an optimal balance between challenge and capacity (Sweller, 1988, 1994). In practical terms, this means that learners are consistently stretched but rarely frustrated, a balance that fosters persistence and long-term retention (Baddeley, 2003; Ellis, 2005).

Another key element of Appil's philosophy is transparency and pedagogical alignment. While some ICALL systems rely heavily on opaque algorithms, Appil's adaptivity is openly anchored in CEFR proficiency levels (Council of Europe, 2001), second language acquisition (SLA) research, and lexicographical data. This transparency benefits both learners and teachers: learners can situate their progress within an internationally recognized framework, while instructors can integrate Appil seamlessly into CEFR-based curricula. In this way, Appil bridges learner autonomy and institutional coherence, functioning equally well as a self-directed learning tool and as a classroom complement. Finally, Appil embodies a principle of minimal reliance on AI-driven black boxes. While many modern systems promise adaptivity through machine learning, their pedagogical logic is often hidden

and their feedback inconsistent. Appil instead employs linguistic analysis and rule-based text processing, tokenization, part-of-speech tagging, syntactic parsing, and lexical profiling, to create adaptive pathways that are interpretable and accountable. This ensures that learners understand why certain materials are presented, and teachers can trace the pedagogical rationale behind the system's design. Far from rejecting innovation, this approach represents a deliberate choice: by prioritizing interpretability and linguistic accuracy, Appil avoids the risks of opacity while still benefiting from advanced computational tools.

In essence, Appil's design philosophy is one of principled innovation. It acknowledges the advances of CALL and ICALL while refusing to replicate their blind spots. By rooting adaptivity in SLA theory, placing vocabulary and phraseology at the center, managing cognitive load carefully, and aligning progression with CEFR standards, Appil aspires to set a new standard for adaptive learning environments. It is not simply a tool for practicing language, but a structured and evolving environment that brings learners into sustained engagement with the language system, guiding them toward communicative competence in a way that is at once personalized, transparent, and pedagogically sound (Amiruddin et al., 2025).

4.3 Core Functionalities of the Appil Platform

The design philosophy outlined above establishes the principles that guide Appil's development, but these principles must be operationalized into concrete features that learners can interact with on a daily basis. Each functionality of the system is carefully designed to embody the theoretical commitments already discussed, ensuring that the learner experience is not only technologically enhanced but also pedagogically grounded. What distinguishes Appil from many existing CALL applications, such as Duolingo (von Ahn, 2013; Loewen et al., 2019), Babbel (Köhler, 2015), or Busuu (Gao, 2022), is that its features are not conceived as isolated add-ons but as interdependent components within an integrated framework of lexical growth, adaptive scaffolding, and cognitive sustainability.

4.3.1 Adaptive Text Processing

At the heart of Appil is its adaptive text processing engine, which tailors reading materials to each learner's profile. Using techniques such as tokenization, sentence splitting, part-of-speech tagging, syntactic parsing, and lexical profiling, the system evaluates texts for lexical coverage and complexity. This allows Appil to dynamically select and adapt passages that maintain Nation's recommended $\leq 2\%$ threshold of unknown words, ensuring that learners are consistently exposed to

comprehensible yet challenging input (Hsueh-chao & Nation, 2000; Nation, 2001, 2013). Over time, as the learner's vocabulary expands, the system recalibrates and introduces more advanced structures, aligning progress with CEFR bands.

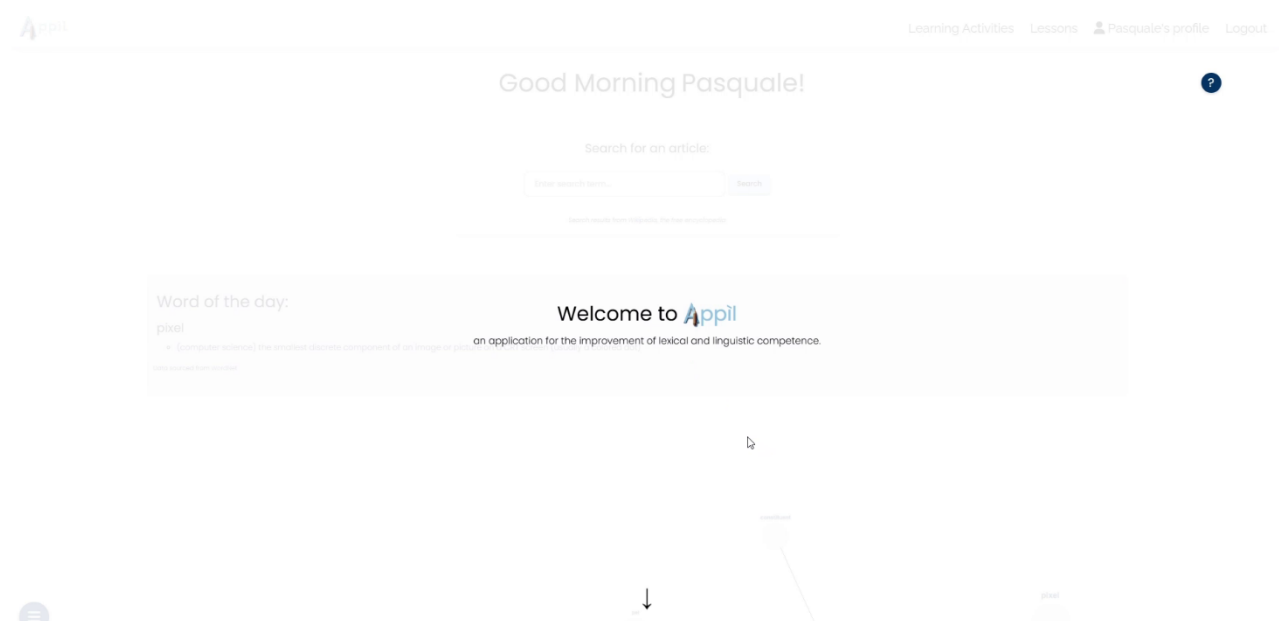


Figure 8 Appil Welcome Page

4.3.2 Lexical Organization and Phraseology

Vocabulary learning in Appil goes beyond the memorization of single words. The system emphasizes lexical organization and phraseological units, including collocations, idiomatic expressions, and discourse markers, which are widely recognized as central to communicative competence (Nation, 2001; Schmitt, 2010; Wray, 2002). These elements are highlighted within texts and reinforced through contextualized practice activities, supporting both noticing and form–meaning integration (Ellis, 2001). Learners are encouraged to observe how words combine in authentic patterns, bridging the gap between passive recognition and active use. This feature operationalizes Nation's (2001, 2013) lexical principles while also addressing Swain's (1985, 1995) argument that productive tasks foster deeper processing and promote the transformation of input into usable output.

4.3.3 Dynamic Personalization and Learner Profiles

Every learner's trajectory is unique, and Appil reflects this individuality through dynamic personalization that evolves continuously with user interaction. A custom learner profile monitors vocabulary breadth, phraseological competence, reading history, and task performance, enabling the system to fine-tune both input difficulty and activity type, whether receptive, productive, or reflective.

Beyond adaptive automation, Appil empowers learners with agency and choice. Users can freely select the texts, topics, or genres that resonate with their interests or academic goals, from literary works to scientific articles or multimedia content. This user-driven selection is then integrated into Appil's adaptive framework, ensuring that even self-chosen materials remain within an optimal linguistic range consistent with Krashen's (1985) $i + 1$ principle. By aligning personalization with autonomy, Appil promotes intrinsic motivation and authentic engagement, allowing learning to progress naturally from comprehension to mastery.

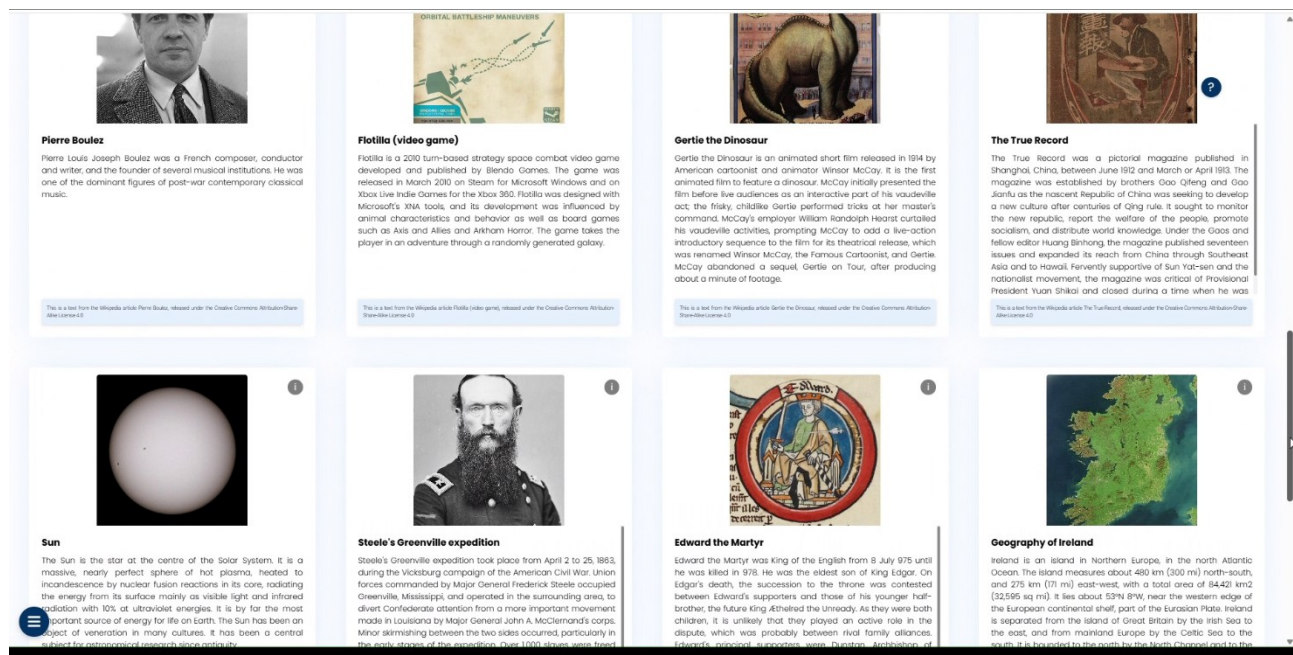


Figure 9 Appil's personalized content selection interface, presenting daily proposed texts alongside learner-selected materials sourced from Wikipedia and other CC BY-SA 4.0-compliant repositories. This design combines adaptive recommendation with open-access resources, supporting learner autonomy while maintaining pedagogical alignment and content transparency

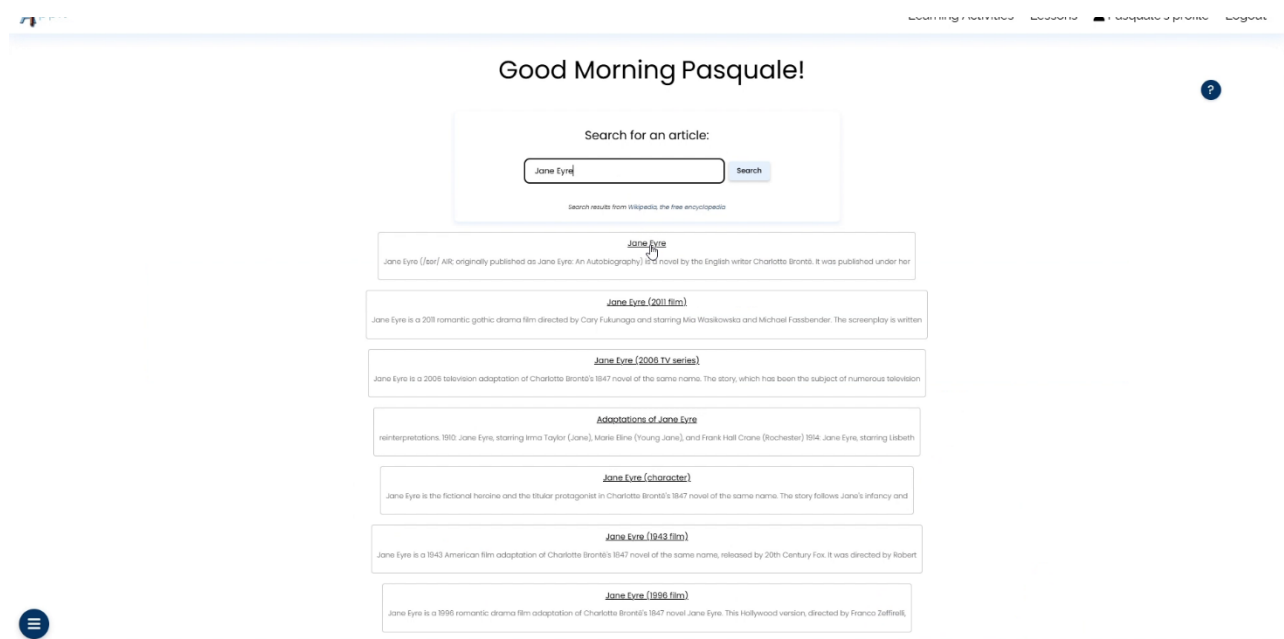


Figure 10 Appil's personalized content search and selection interface, integrating daily recommended texts with learner-driven exploration of Wikipedia and other CC BY-SA-compliant materials

illustrate this process: learners can browse and choose materials through an intuitive interface, while Appil dynamically analyzes lexical complexity and adapts feedback and practice modes accordingly. This synthesis of freedom and intelligent guidance ensures that personalization remains both meaningful and pedagogically grounded.

4.3.3.1 From Personalized Texts to Practice-Based Learning

Appil extends personalization beyond text selection, transforming individualized lexical input into practice-based learning activities. Once texts are analyzed and adapted to a learner's profile, the system automatically extracts relevant lexical and phraseological items to create interactive exercises such as flashcards, translation tasks, and retrieval games. These activities are dynamically generated from each learner's reading materials, ensuring that practice is contextually grounded and directly linked to previously encountered vocabulary.

This approach operationalizes principles of retrieval practice and spaced repetition, reinforcing memory consolidation through varied and meaningful engagement. By recontextualizing words in multiple formats: receptive, productive, and reflective, Appil supports durable lexical retention and productive fluency. The system thereby closes the adaptive loop: learners move from personalized input to personalized output, continuously refining their competence through interaction and feedback.

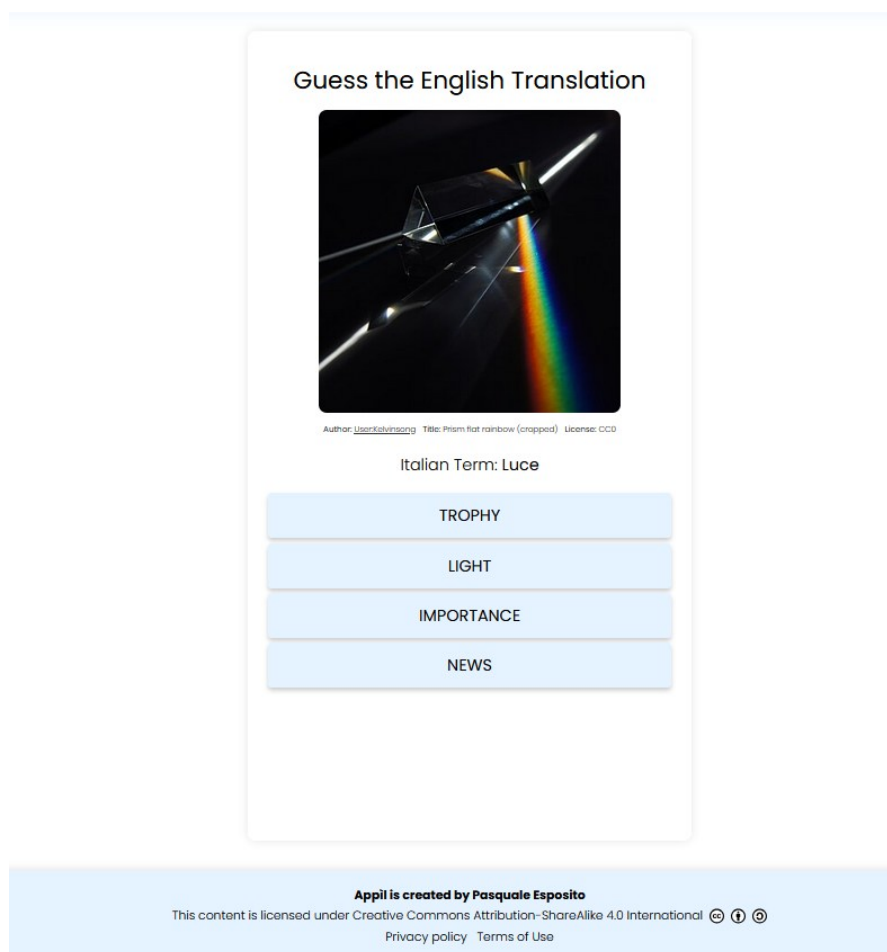


Figure 11 Appil's flashcard-based translation activity, illustrating how personalized lexical input is converted into adaptive, practice-oriented learning tasks

Figure 11 illustrates one such activity, where a lexical item drawn from the learner's reading corpus appears within an adaptive flashcard interface. The user is prompted to identify the English translation of the Italian term *luce*, integrating multimodal support through imagery and multiple-choice scaffolding. This design merges personalization with active recall, exemplifying Appil's goal of transforming individualized exposure into measurable language growth.

4.3.4 Lexical Reinforcement through Gamified Interaction

Appil extends its adaptive learning design to include gamified lexical reinforcement activities, such as dynamically generated crossword puzzles (Figure 12). These activities draw directly from the learner's personalized lexical database; recycling vocabulary items previously encountered in texts or practice modules. By engaging learners in clue-based retrieval, the crossword promotes deeper semantic processing, encouraging recall not only of isolated words but also of their collocational and definitional contexts.

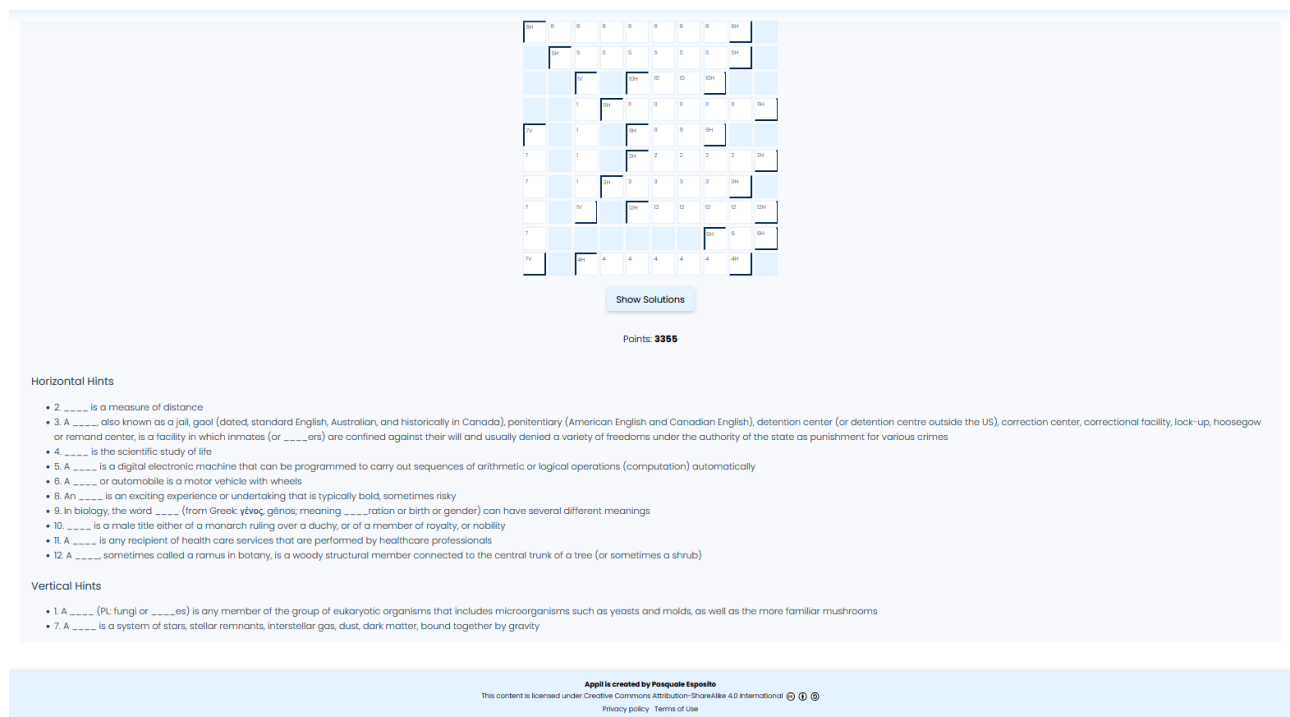


Figure 12 Appil's adaptive crossword activity, illustrating gamified retrieval practice for personalized lexical consolidation

This form of interactive practice aligns with principles of retrieval-based learning (Roediger & Butler, 2011), leveraging challenge and repetition to strengthen long-term memory traces. The integration of game elements adds intrinsic motivation, transforming review into an active, self-directed process. Within Appil's broader ecosystem, the crossword serves as both a diagnostic and consolidative tool, reinforcing lexical mastery while maintaining engagement and flow.

4.3.5 Feedback and Micro-Negotiations of Meaning

In Appil, feedback is conceived as an interactive learning process rather than a static act of correction. Grounded in Long's Interaction Hypothesis (Long, 1996) and informed by Sociocultural Theory (Lantolf, 2000; Lantolf & Thorne, 2009; Vygotsky, 1978), the system promotes micro-negotiations of meaning that foster engagement and cognitive depth. Rather than simply flagging an error, Appil initiates a dialogic sequence that guides learners toward self-repair through graduated prompts, lexical and collocational hints, and context-sensitive reformulations.

These adaptive feedback loops simulate authentic communicative exchanges, enabling learners to test and refine their interlanguage hypotheses in real time. When learners encounter difficulty, Appil provides dynamic scaffolding, for example, offering a collocational cue, highlighting a semantic contrast, or providing a reformulated version of the learner's utterance, to draw attention to subtle

distinctions in meaning and usage. Through this process, errors become opportunities for discovery, enhancing metalinguistic awareness and deepening language competence.

Figure 13 and *Figure 14* illustrate Appil’s feedback architecture and meaning-mapping interface. In these visualizations, learner input triggers a semantic network that reveals relationships among lexical items and conceptual fields, guiding reflection and hypothesis testing. The right-hand panels display dynamically generated *WordNet*-based relational graphs that exemplify Appil’s capacity to connect linguistic feedback with conceptual understanding.

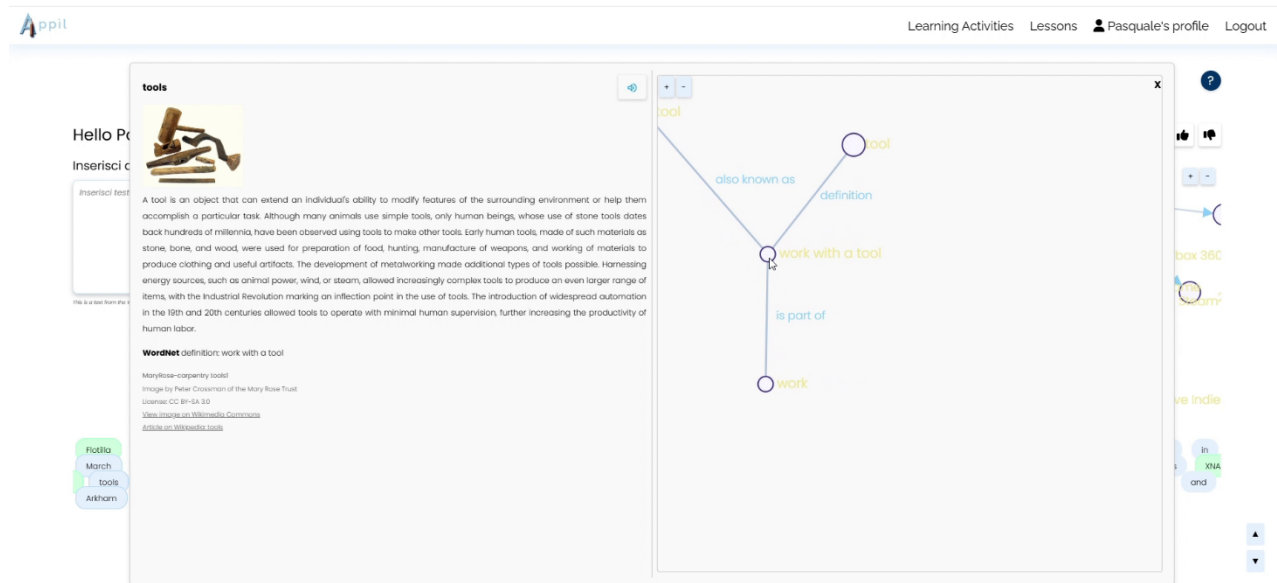


Figure 13 Appil's adaptive feedback and semantic mapping interface, demonstrating micro-negotiations of meaning through contextualized network visualization (word: tools)

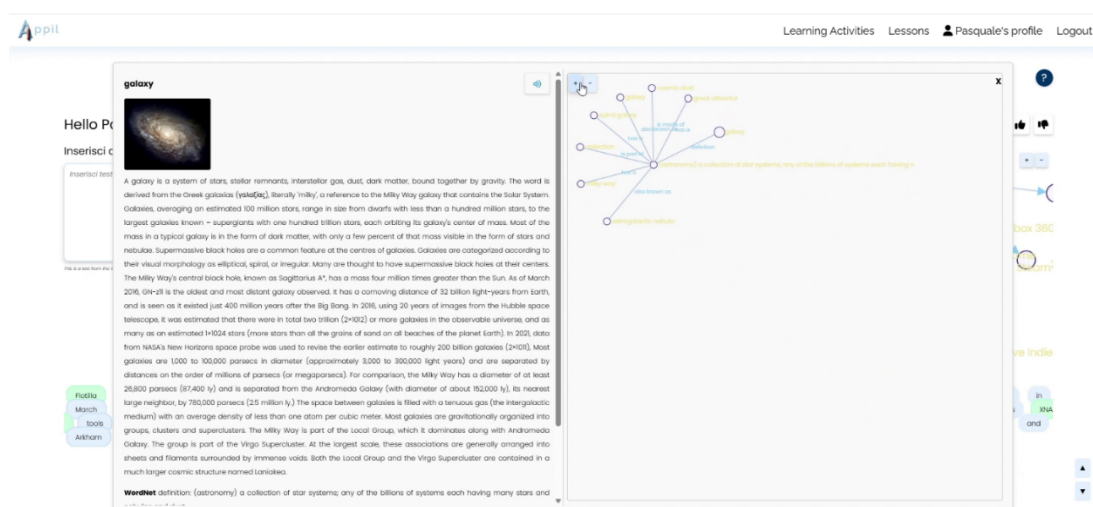


Figure 14 Appil’s adaptive feedback and semantic mapping interface, demonstrating micro-negotiations of meaning through contextualized network visualization (word: Galaxy)

4.3.6 Cognitive Load Management

To sustain engagement and prevent cognitive overload, Appil integrates cognitive load management strategies informed by Dual Process Theory (Evans, 2008; Kahneman, 2011) and contemporary psycholinguistic research. New lexical and syntactic elements are introduced incrementally and recycled across diverse contexts, allowing learners to consolidate understanding through spaced repetition and contextual variation. This gradual layering reduces processing demands while maintaining meaningful challenge.

Figure 15 Appil's adaptive interface visualizing text input, lexical mapping, and interactive task components as mechanisms for managing cognitive load and sustaining engagement

Figure 15 Appil's adaptive interface visualizing text input, lexical mapping, and interactive task components as mechanisms for managing cognitive load and sustaining engagement illustrates Appil's integrated interface, where textual input, lexical visualization, and interactive task modules operate in tandem to distribute cognitive effort across multiple representational modes. By dynamically balancing information flow and learner control, Appil exemplifies how adaptive system design can translate cognitive theory into effective, scalable pedagogical practice.

The screenshot displays the Appil adaptive interface. At the top, the Appil logo is on the left, and navigation links for 'Learning Activities', 'Lessons', 'Pasquale's profile', and 'Logout' are on the right. A user greeting 'Hello Pasquale' is shown. Below it, a text input area with the placeholder 'Inserisci qui il tuo testo:' and a 'Send' button is present. To the right of the input area is a complex lexical map. This map consists of a central node 'I' connected to several peripheral nodes: 'player', 'tokens', 'released on Steam for Microsoft Windows on Xbox Live Indie Games for Xbox 360', 'zero was based on strategy space combat video game released in March 2010 on Steam', 'released for Microsoft Windows on Xbox Live Indie Games', 'takes place through', 'wakes player in', 'galaxy', and 'released in March 2010 on Xbox Live Indie'. Below the lexical map is a horizontal bar containing a sequence of words: 'Flotilla', 'is', 'a', '2010', 'turn', 'based', 'strategy', 'space', 'combat', 'video', 'game', 'developed', 'and', 'published', 'by', 'Blendo', 'Games', 'The', 'game', 'was', 'released', 'in', 'March', '2010', 'on', 'Steam', 'for', 'Microsoft', 'Windows', 'on', 'Xbox', 'Live', 'Indie', 'Games', 'and', 'for', 'as', 'well', 'as', 'board', 'games', 'such', 'as', 'Axis', 'and', 'Allies', 'and', 'Arkhon', 'Horror', 'The', 'game', 'takes', 'the', 'player', 'in', 'an', 'adventure', 'through', 'a', 'Games', 'behavior', 'randomly', 'generated', 'galaxy'. At the bottom of the interface are four interactive buttons: 'Send words' with a thumbs up icon, 'Fill the gaps game' with a pencil icon, 'Play the memory game' with a memory card icon, and 'Read' with a speaker icon.

Figure 15 Appil's adaptive interface visualizing text input, lexical mapping, and interactive task components as mechanisms for managing cognitive load and sustaining engagement

Appil's instructional design balances System 1's intuitive, fluency-oriented processing with System 2's analytical, reflective engagement. Explicit annotations, metalinguistic explanations, and guided

tasks are interleaved with implicit exposure and comprehension-based activities, ensuring that attention alternates between automaticity and reflection. This calibrated interaction between cognitive systems supports sustained motivation, efficient encoding, and long-term retention, allowing learning to remain both demanding and cognitively sustainable.

4.3.7 CEFR-Aligned Progression

Appil explicitly aligns its adaptive learning architecture with the Common European Framework of Reference for Languages (CEFR), ensuring transparent and internationally recognized benchmarks for learner progress. The learning platform continuously monitors vocabulary size, syntactic sophistication, and phraseological competence, mapping these linguistic indicators to the appropriate CEFR proficiency bands (A1–C2). This systematic calibration allows learners to situate their development within a clear and measurable continuum of achievement.

For teachers and institutions, CEFR alignment facilitates seamless curricular integration, assessment, and certification, making Appil compatible with both formal and informal learning contexts. The system's analytics provide granular insights into learner progress, enabling data-informed instructional decisions and targeted pedagogical support.

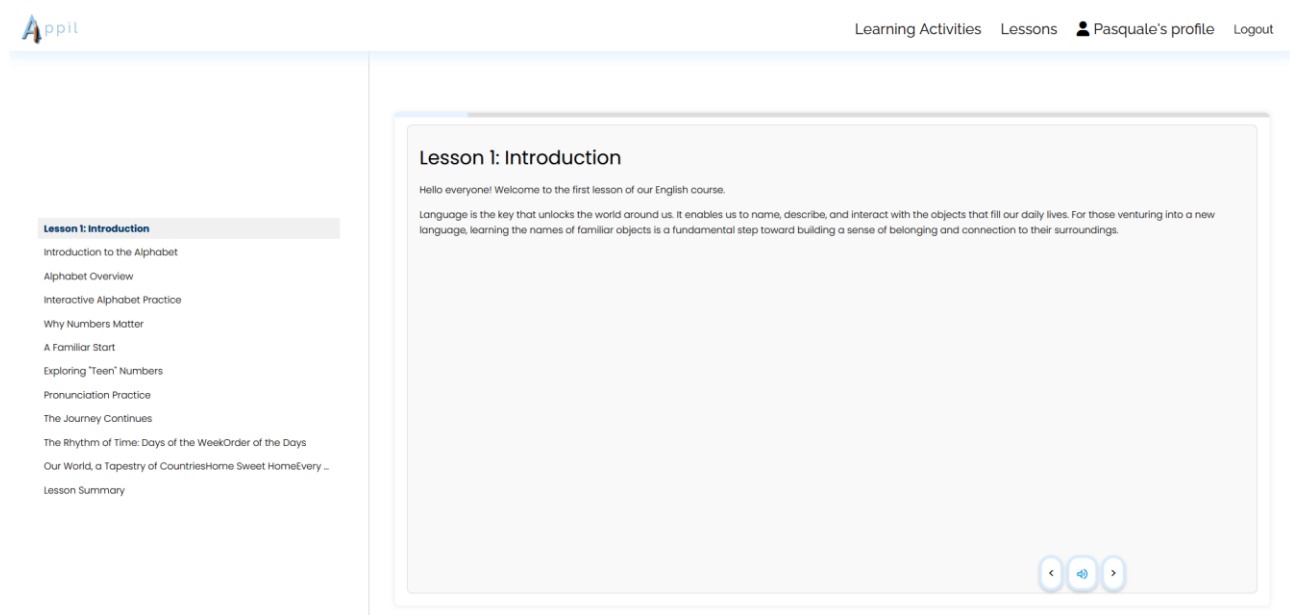


Figure 16 Appil's lesson progression interface demonstrating CEFR-aligned sequencing and cumulative skill development across modules

Figure 16 illustrates Appil's lesson progression interface, where each module builds on previously acquired competencies, reinforcing the incremental nature of CEFR-aligned development. This structured yet adaptive design bridges individualized learning trajectories with standardized

educational frameworks, ensuring that personalized advancement remains pedagogically coherent and globally interpretable.

4.3.8 Automated Generation of Lesson-Based Activities

Within Appil's adaptive learning framework, lesson content is not static but dynamically transformed into interactive exercises that mirror real-world language use. Drawing from the learner's profile and previously processed texts, the system automatically generates contextualized quizzes, classification tasks, and comprehension activities such as the Job Titles Quiz (Figure 17). These exercises are algorithmically composed using the system's lexical and syntactic database, ensuring that each activity remains pedagogically coherent and aligned with the learner's proficiency level.

Figure 17 Automatically generated Job Titles Quiz illustrating how Appil converts lesson content into adaptive, personalized exercises that reinforce contextual understanding

This automation serves a dual purpose: it personalizes practice while maintaining scalability, allowing each learner to experience a unique sequence of tasks derived from their evolving knowledge base. By converting input texts and lesson modules into self-adaptive exercises, Appil reinforces meaningful repetition and applied usage across multiple modalities. The resulting activities promote semantic precision, functional fluency, and active recall, turning structured lessons into a living, data-driven ecosystem of continuous learning.

Beyond generating adaptive practice, Appil also seeks to bridge the gap between abstract linguistic resources and concrete learner experience. The system integrates data from lexical databases, morphological analyzers, and semantic ontologies to transform decontextualized linguistic information into interactive, learner-centered activities. Exercises such as the Irregular Plurals Game

(Figure 18) exemplify this approach, translating abstract grammatical principles, like irregular inflectional patterns, into engaging, meaning-oriented practice.

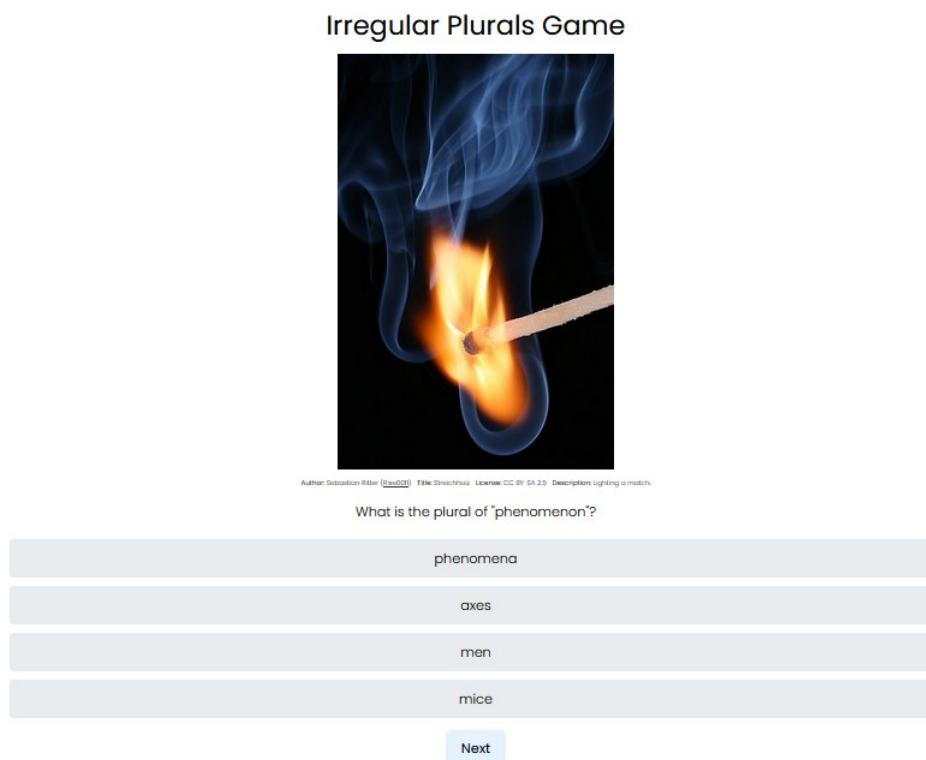


Figure 18 Appil's *Irregular Plurals Game* illustrating the integration of abstract linguistic resources with interactive, applied learning practice

Through this integration, Appil demonstrates how computational linguistic resources can be operationalized within pedagogically sound environments. Learners are not merely exposed to rules or lists, but actively apply linguistic knowledge in contextualized, task-based interactions that foster pattern recognition, hypothesis testing, and memory consolidation. By connecting formal linguistic representation with experiential learning design, Appil embodies a data-informed bridge between theoretical language models and practical language acquisition.

Among Appil's repertoire of adaptive activities, the Verb Matching Memory Game (Figure 19) exemplifies how gamified practice can serve as both an engaging and cognitively principled form of reinforcement. The exercise challenges learners to identify and match base forms with their corresponding inflected verbs, fostering morphological awareness, pattern recognition, and form–meaning mapping through repetition and retrieval. By embedding linguistic processing in an interactive format, Appil aligns playful engagement with evidence-based learning mechanisms such as deep encoding, retrieval practice, and spaced repetition.

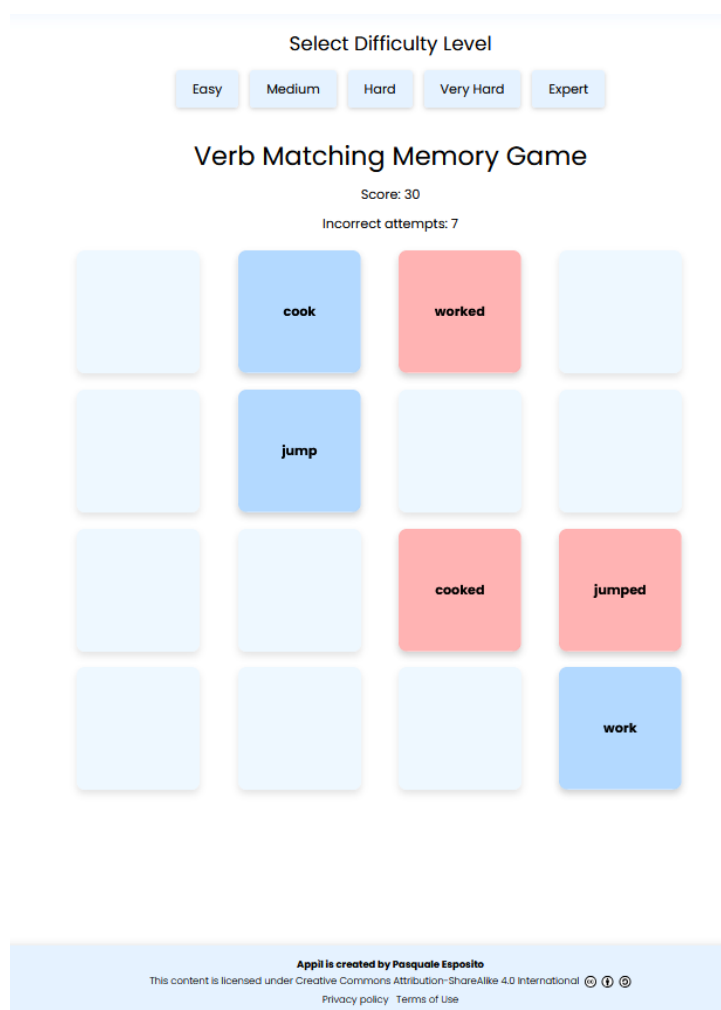


Figure 19 Appil's Verb Matching Memory Game, demonstrating adaptive difficulty scaling and gamified reinforcement of morphological competence

What distinguishes this activity is its competence-based scalability. Learners can choose from multiple difficulty levels, ranging from Easy to Expert, each calibrated according to their performance data and CEFR-aligned proficiency profile. The system dynamically recommends an appropriate level but allows users to self-select, promoting learner autonomy while preserving pedagogical coherence. As learners progress through lessons and expand their lexical–grammatical repertoire, the difficulty of the memory task adjusts automatically, introducing more complex verb forms, irregularities, and nuanced grammatical contrasts.

This flexible architecture not only supports differentiated learning but also enhances metacognitive control, enabling users to monitor and challenge their own progress. In doing so, Appil merges the motivational benefits of gamification with the precision of adaptive sequencing, ensuring that each learner encounters practice that is both attainable and incrementally demanding.

4.3.9 Contextualized Pronunciation and Interactive Reading

Another key feature of Appil’s adaptive design is its capacity to transform ordinary reading passages into interactive pronunciation and comprehension activities. As illustrated in Figure 17, learners can engage directly with texts where each lexical item, phrase, or grammatical unit becomes an active learning node. Words can be automatically highlighted, selected, and pronounced in context, allowing learners to perceive and internalize authentic prosodic patterns while simultaneously reinforcing lexical recognition.

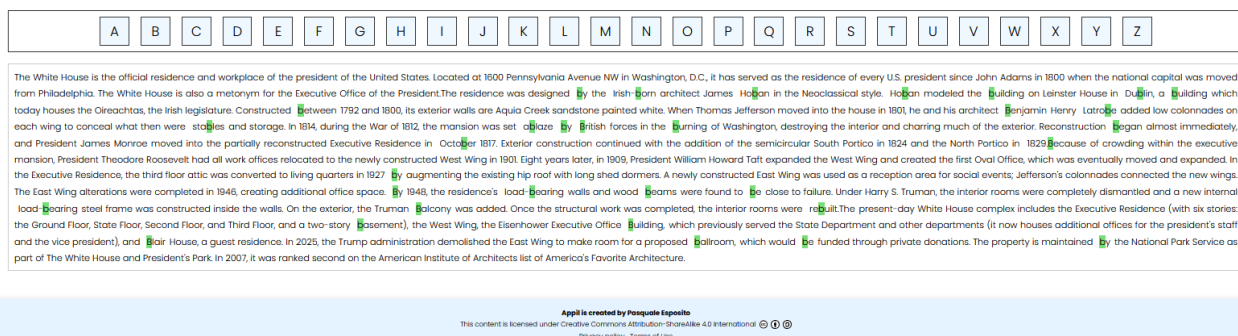


Figure 20 Appil’s interactive reading and pronunciation activity, demonstrating contextualized listening and user-driven lexical selection for adaptive multimodal learning

This activity draws on dual coding theory (Paivio, 1990) and multimodal input principles, emphasizing that learning deepens when auditory and visual channels are activated in parallel. Learners are thus exposed to real-time pronunciation models that situate phonological form within meaningful discourse, bridging the gap between textual comprehension and oral production. The system’s ability to automatically select content, whether drawn from Appil’s internal repository or user-provided material, ensures that listening practice remains personalized, relevant, and adaptive. Pedagogically, this approach supports phonological noticing and form–meaning integration, helping learners align segmental pronunciation with contextual cues such as stress, rhythm, and syntactic flow. The integration of interactive listening within reading tasks also fosters bottom-up decoding skills and top-down inferencing, enabling learners to consolidate comprehension while improving oral fluency.

4.3.10 Spatial Involvement and Multimodal Cognition

A further dimension of Appil’s adaptive design is its emphasis on spatial involvement as a means of enhancing cognitive and linguistic processing. Through interactive visual interfaces such as maps (Figure 21), learners engage with language in spatially grounded contexts, linking lexical and

conceptual knowledge to geographical, cultural, or situational representations. This feature draws on principles of embodied cognition and spatial learning theory, which suggest that knowledge retention is strengthened when abstract linguistic information is anchored in perceptual and spatial experience. In practice, activities such as identifying countries, cities, or landmarks combine linguistic recognition with spatial reasoning, allowing learners to integrate words with their real-world referents. The system's multimodal design, linking text, image, and location, encourages learners to move beyond purely symbolic processing and to form semantic associations grounded in sensory and spatial awareness. This not only enhances comprehension but also promotes contextual transfer, enabling learners to recall and apply vocabulary more effectively across communicative situations.

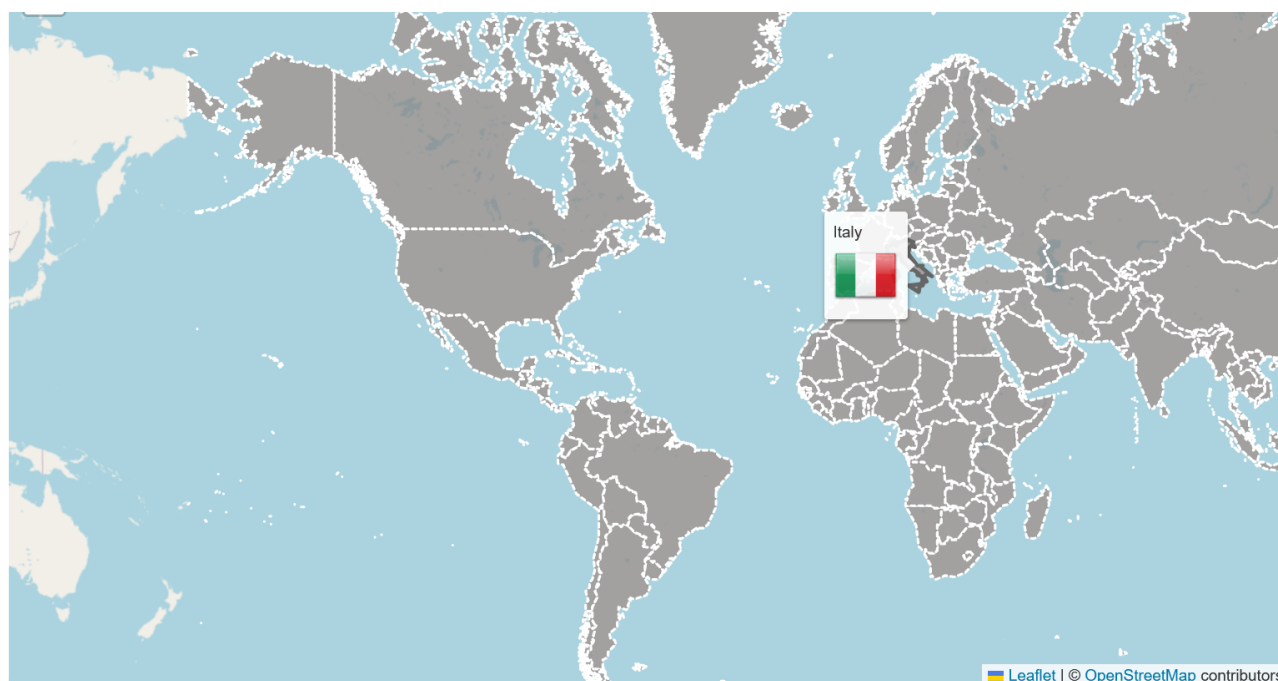


Figure 21 Interactive map activity in Appil, demonstrating spatial involvement and multimodal grounding of lexical and cultural knowledge

In this specific activity, learners explore a world map that dynamically associates countries with their corresponding nationalities, adjectives, and pronunciations. By clicking on a location, users can hear the country name pronounced in context, view the adjectival and demonymic forms (e.g., *Italy* → *Italian* → *Italian culture*), and observe how these lexical variants function across grammatical structures. This multimodal experience connects phonological, morphological, and semantic dimensions of language within a unified, exploratory task.

Pedagogically, such spatially anchored activities encourage learners to form integrated lexical networks that mirror real-world conceptual relationships. Instead of memorizing isolated vocabulary items, learners engage in contextualized pattern discovery, reinforcing both form–meaning mappings and cross-linguistic associations. The combination of audio, visual, and spatial input activates

multiple cognitive pathways, enhancing recall and supporting learners with diverse sensory preferences.

From a broader perspective, this feature exemplifies Appil's commitment to interactive, discovery-based learning, where users construct understanding through movement, perception, and multimodal feedback. By relating geography to language, through pronunciation, nationality, and adjectival use, Appil transforms abstract linguistic data into experiential knowledge, bridging linguistic form, cultural meaning, and spatial cognition in a cohesive and engaging way.

4.3.11 Augmented Reality and Immersive Lexical Exploration

Appil extends its adaptive and multimodal philosophy into the realm of augmented and virtual reality, providing users with immersive environments for lexical exploration and semantic mapping. Within this 3D lexical space (*Figure 22*), each node represents a word, concept, or thematic field, rendered as a navigable sphere in a multidimensional semantic universe. Learners can move freely through this environment, selecting items to reveal definitions, multilingual equivalents, images, and authentic usage examples. This spatialized lexicon transforms vocabulary learning from a static, list-based activity into an experiential journey, where semantic relationships are perceived through proximity, color, and spatial configuration.

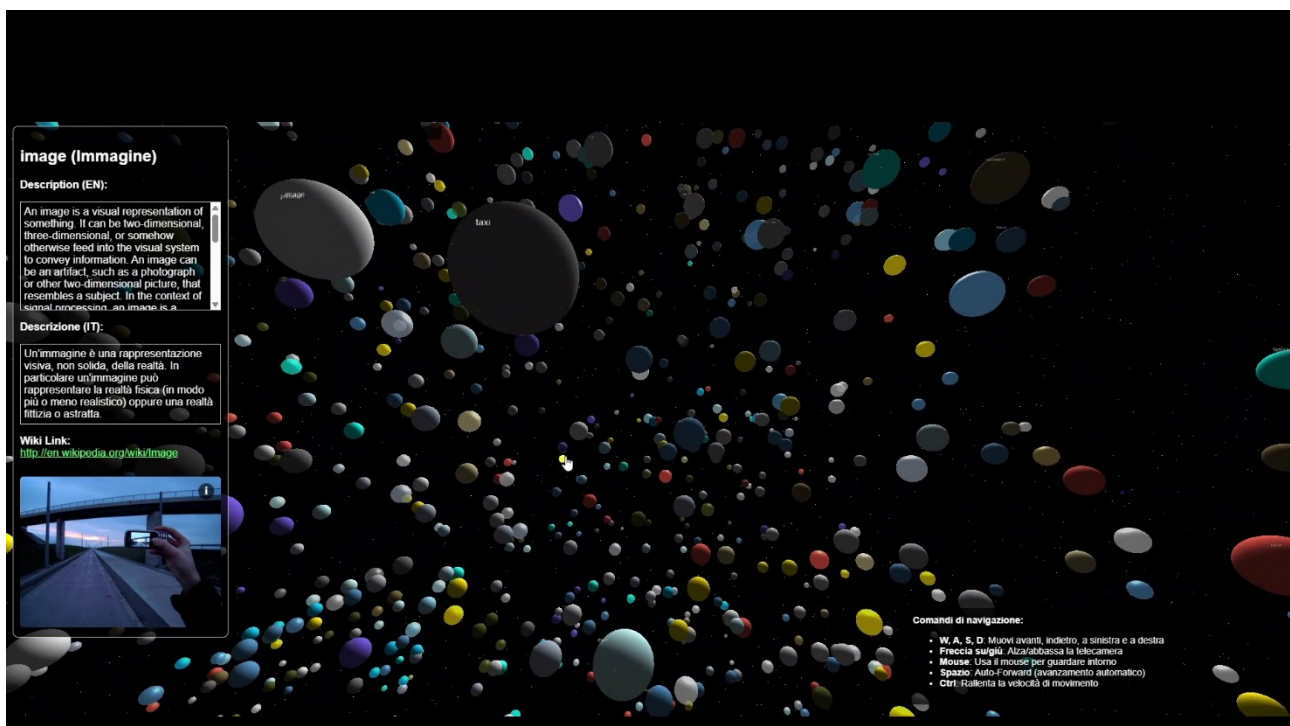


Figure 22 Augmented-reality lexical exploration in Appil, where learners navigate a three-dimensional semantic network enriched with bilingual descriptions and multimedia references

This approach draws on theories of embodied cognition and situated learning (Lave & Wenger, 1991), emphasizing that knowledge acquisition is most effective when learners interact with concepts in perceptually rich and meaningful contexts. By navigating through lexical constellations, users engage both spatial memory and conceptual reasoning, constructing dynamic mental maps that mirror the interconnected structure of language itself. The inclusion of bilingual annotations and visual cues, as shown above, where English and Italian entries coexist, also promotes cross-linguistic awareness and metalinguistic reflection, encouraging learners to notice form, meaning, and cultural nuance simultaneously.

Pedagogically, this augmented reality module serves not only as an exploratory tool but as a cognitive amplifier: learners experience language as a living network, not merely as an abstract system. The ability to zoom, rotate, and interact with lexical fields allows them to visualize how individual words cluster around broader semantic domains, such as movement, emotion, or art, thus reinforcing associative learning and conceptual integration. By blending linguistic data with immersive technology, Appil redefines vocabulary acquisition as a process of spatial discovery, bridging the gap between digital learning and embodied experience.

4.3.12 Network analysis and semantic connectivity

Appil employs network analysis to model and visualize the complex relationships among lexical, syntactic, and phraseological elements within its adaptive ecosystem. In this framework, words, phrases, and grammatical patterns are represented as interconnected nodes, while their co-occurrence frequencies and semantic associations form weighted links. This network-based representation enables the system to identify lexical clusters, central connectors, and emergent communities of meaning that reflect authentic language use. By mapping a learner's evolving vocabulary against this network, Appil can detect gaps, redundancies, and potential learning pathways that would otherwise remain implicit. The analysis also supports adaptive text selection and task generation by prioritizing items that strengthen underdeveloped connections or consolidate fragile ones. *Figure 23* illustrates this semantic network, where nodes represent lexical units and edges indicate associative strength, providing a transparent visualization of how Appil operationalizes data-driven language learning through graph-based modeling.

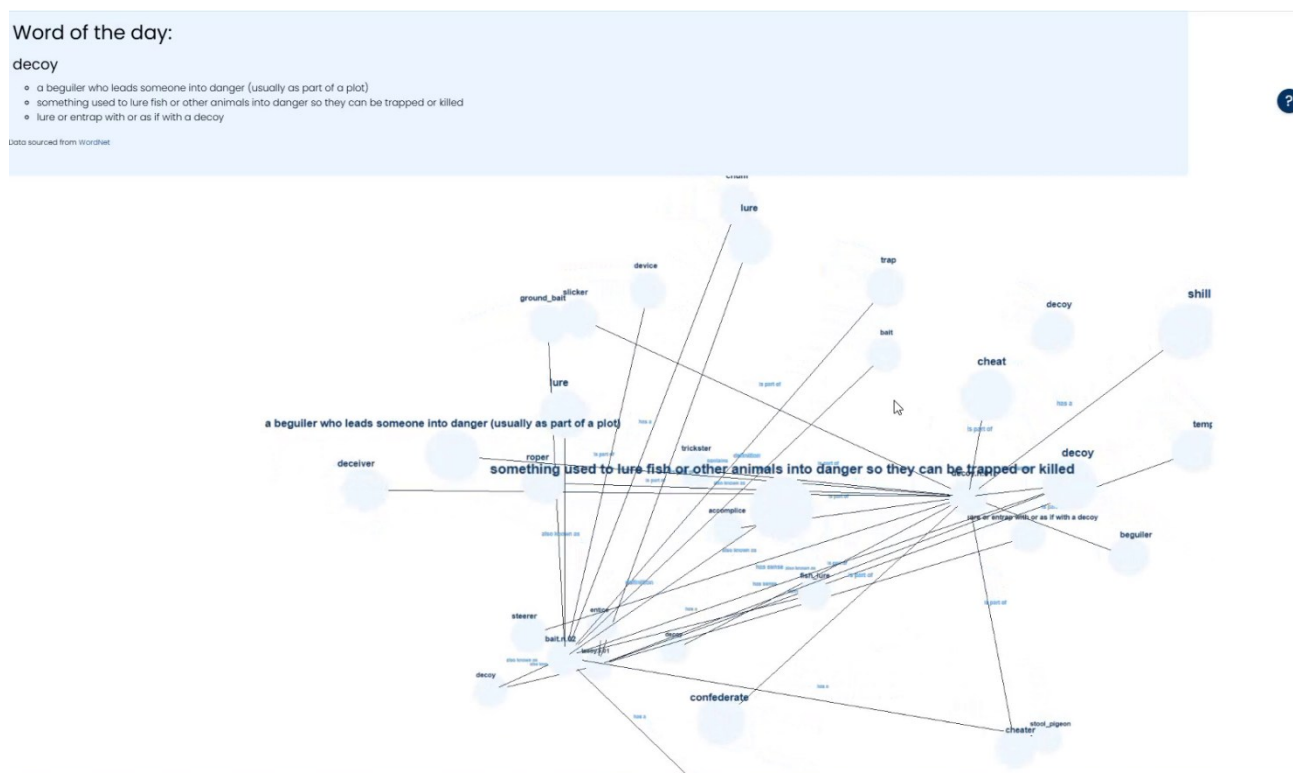


Figure 23 Network visualization of the lexical item decoy demonstrating semantic associations and relational structure within Appil's adaptive lexicon analysis framework

4.3.13 User-Centered Transparency and Interpretability

Unlike many Intelligent Computer-Assisted Language Learning (ICALL) systems that rely on opaque, data-driven algorithms, Appil is founded on the principles of user-centered transparency and pedagogical interpretability. Rather than functioning as a black box, the learning platform provides learners and instructors with clear insights into the rationale behind adaptive decisions. Each selection, whether a text, vocabulary set, or activity, is accompanied by accessible explanations detailing how it aligns with the learner's profile, lexical coverage, and current developmental stage. This design philosophy is rooted in the belief that informed awareness enhances engagement: when users understand why a task is presented, they are more likely to approach it with confidence, curiosity, and strategic intent.

For educators, transparency extends beyond surface feedback to encompass rich diagnostic analytics. Teachers can view comprehensive data visualizations that trace lexical and grammatical progression, phraseological development, error typologies, and task-specific performance trends. These insights enable educators to make data-informed pedagogical interventions, complementing the system’s adaptivity with human expertise and contextual understanding. In this way, Appil fosters a symbiotic relationship between automation and instruction, in which computational intelligence supports, but does not replace, pedagogical judgment.

From a cognitive perspective, transparency also contributes to metacognitive regulation and learner autonomy (Cotterall & Crabbe, 2012). When learners can see how their linguistic behavior influences system output, they develop an explicit awareness of their own learning trajectory. This awareness encourages reflection, goal-setting, and strategic behavior, key features of self-regulated learning. Moreover, the system's interpretability allows users to question, explore, and even challenge adaptive suggestions, turning the learning process into a collaborative negotiation between human agency and algorithmic guidance.

By making its adaptive logic visible, Appil distinguishes itself from traditional AI-based learning platforms that prioritize predictive efficiency over educational clarity. Designed and developed by the author within the framework of this doctoral research, Appil was purposefully conceived to embody principles of transparency, interpretability, and ethical AI in language education. Its architecture reflects a deliberate pedagogical design, ensuring that learners and teachers remain active participants in the interpretive loop rather than passive recipients of algorithmic output. Ultimately, this commitment to openness transforms Appil from a mere instructional tool into a shared pedagogical partner, one that empowers users through understanding, reflection, and agency.

4.3.14 Overt and Covert teaching processes

The course structure of Appil is intentionally organized into two main phases: the learning phase and the acquisition phase, inspired by Krashen's (1982) distinction between language learning and language acquisition. According to Krashen's theory, learning refers to the conscious process of developing explicit knowledge of linguistic rules, whereas acquisition denotes a subconscious process through which language is internalized naturally via meaningful exposure and interaction. Reflecting this distinction, the learning phase in Appil provides explicit instruction on grammar, vocabulary, and usage patterns, while the acquisition phase facilitates implicit language development through adaptive and communicative tasks that simulate authentic input and interaction.

Here, learners are exposed to dynamically modified content, which adapts to their individual progress and provides implicit opportunities for language acquisition. This content is not directly taught but encountered in real contexts, allowing learners to internalize the language naturally, through use, rather than through overt instruction. By integrating both phases, Appil ensures that users benefit from a holistic approach, where structured learning is complemented by immersive acquisition, fostering both conscious knowledge and subconscious fluency. By combining these two phases, Appil also ensures a comprehensive learning experience that balances explicit, overt teaching methods with covert, automated processes that promote natural language acquisition, enabling learners to absorb language more organically.

Both the inductive and deductive approaches, though differing in their methods of content delivery, align with the concept of overt teaching processes due to the clarity with which learning objectives are communicated and the active involvement of the teacher in guiding students' learning. Conversely, covert teaching processes involve more implicit methods, where learning is facilitated through indirect means. In these scenarios, instruction may not be immediately apparent to the learner; rather, it occurs through activities such as guided discovery, implicit feedback, or experiential learning, where students acquire knowledge or skills without direct explanation. This dichotomy between overt and covert teaching processes allows educators to employ a balanced approach, combining explicit instruction with opportunities for autonomous, self-directed learning.

In educational pedagogy, the distinction between overt and covert teaching processes is critical for understanding the diverse strategies employed to facilitate learning. Overt teaching processes refer to instructional methods that are explicit, clearly articulated, and easily observed by students. These include direct instruction, where teachers explicitly convey information, provide clear explanations, and model desired behaviors or skills, ensuring that learners are fully aware of the objectives and content being taught (Mathews & Cohen, 2022).

The distinction between inductive and deductive teaching methods further clarifies this concept. Both methods can be categorized as overt teaching processes, as they involve clear communication of learning goals and transparent guidance. Deductive teaching methods are overt in nature because they begin with a general principle or theory, which is then applied to specific examples. The teacher explicitly outlines the learning objectives, presents the overarching concept, and guides students through examples that illustrate how the concept applies. This structured, direct approach makes the learning process clear and observable to students.

In contrast, inductive teaching methods, while less direct in their approach, also maintain an overt structure. In this method, the teacher presents specific examples or data and encourages students to identify patterns or derive generalizations. Though the teacher's role is more facilitative, guiding students to make connections on their own, the process remains explicit as the teacher structures the learning experience to help students build their understanding through inquiry and discovery.

For the Appil's design, both overt and covert teaching methods are strategically integrated to support diverse learning needs. The application employs overt methods through deductive learning schemes, where explicit instruction is provided, allowing learners to directly engage with grammatical rules, vocabulary lists, and clear explanations of language structures. These features are designed to ensure that users have a solid foundation in the language, with well-defined learning objectives and measurable outcomes. Simultaneously, the application incorporates covert methods through inductive learning schemes, where learners are encouraged to discover linguistic patterns and rules implicitly

through contextualized activities, such as interactive exercises, reading comprehension, and problem-solving tasks. This balanced approach, combining overt deductive instruction with covert inductive learning, aims to foster both conscious understanding and intuitive language acquisition, catering to a wide range of learning styles and promoting deeper engagement with the target language.

By incorporating both overt and covert instructional methods, the application not only addresses the diverse cognitive preferences of learners but also aligns with contemporary theories of second language acquisition. Research in the field suggests that a combination of deductive and inductive approaches enhances language learning by providing a comprehensive learning environment that accommodates both analytic and experiential learners. The overt, deductive elements of the application allow learners to consciously process language rules and structures, facilitating accuracy and explicit knowledge. Meanwhile, the covert, inductive components encourage subconscious pattern recognition and language intuition, which are essential for fluency and the ability to use language in context. This dual approach is particularly effective in promoting self-regulated learning, as it enables users to engage with the material in multiple ways, fostering both deliberate practice and spontaneous language use. By balancing these methods, the application aims to create a holistic learning experience that supports sustained language development over time.

Within the application, the lesson section is intentionally designed to be primarily deductive, providing learners with explicit instruction on key language concepts. However, the deductive nature of these lessons is carefully balanced by limiting the amount of direct information presented. Rather than overwhelming learners with extensive explanations, the lessons offer concise, essential content that serves as a foundation for understanding. To reinforce this foundation and promote deeper comprehension, the application integrates inductive-based activities within and following the lessons. These activities encourage learners to apply the concepts introduced deductively, enabling them to discover language patterns and rules through exploration and practice. This approach ensures that while the lessons provide necessary guidance, learners are still actively engaged in the learning process, fostering a more intuitive grasp of the language.

While the transition from deductive lessons to inductive activities within the application might initially appear to present a wide gap for learners, this potential challenge is mitigated through carefully designed scaffolding. Both the didactic and technical aspects of the application are structured to provide ongoing feedback that supports learners as they bridge this gap. The activities are not only crafted to encourage discovery and application of language rules but also embedded with feedback mechanisms that guide learners in real-time. This feedback is strategically programmed to respond to user input, offering hints, explanations, or prompts that reinforce the concepts introduced in the lessons. By doing so, the application ensures that learners are supported throughout their

learning journey, allowing them to progress from understanding explicit rules to intuitively applying them in various contexts. This scaffolding is essential for maintaining learner engagement and promoting successful language acquisition, even as they navigate the shift from deductive to inductive learning tasks.

A key feature of the application is its ability to dynamically analyze each learner's input and progress to generate tailored learning experiences. Every piece of data, from the texts processed to the activities completed, is systematically analyzed by the software. This analysis informs the development of subsequent activities, which are enhanced to better align with the learner's current level and needs. The adaptive algorithm operates in real-time, ensuring that the content adjusts continuously, offering material that is slightly above the learner's current proficiency, an application of Krashen's 'input +1' theory. This constant calibration not only challenges learners but also supports gradual, manageable growth in their language skills. By continuously assessing learner performance and interaction with the content, the application can adjust the difficulty, focus, and feedback of activities, ensuring they remain both challenging and accessible. This adaptive approach not only personalizes the learning process but also helps maintain learner motivation and engagement by offering progressively more relevant and effective learning opportunities. Moreover, the system encourages learner autonomy, as the evolving content invites users to reflect on their progress and engage with increasingly complex tasks at their own pace. Through this real-time adjustment, the application fosters a responsive learning environment that evolves alongside the learner's language acquisition journey.

The design of our web application represents a fusion of semi-traditional course elements with an innovative learning interface. This hybrid approach combines a structured, semi-traditional curriculum with a dynamic, free-learning environment. The semi-traditional aspect ensures that learners benefit from a well-defined instructional sequence and clear educational goals, while the innovative interface offers a flexible, interactive platform for independent exploration and practice. Learners are given the freedom to navigate through a wide range of resources, drawing on both overt and covert learning processes. This allows them to experiment with the material, discover patterns, and implicitly acquire language skills while still benefiting from structured guidance when necessary. By integrating these two components, the application provides a comprehensive learning experience that supports both guided instruction and autonomous study. This design not only enhances the effectiveness of the learning process but also empowers learners to engage with the material in a way that best suits their individual needs and preferences.

While Appil prioritizes overt teaching processes, enabling students to engage directly with foreign language material, it also seamlessly integrates a multitude of background processes that enhance the learning experience. Learners are explicitly aware of the vocabulary, grammatical structures, and

language rules presented in a clear and structured manner. However, they may not fully recognize the sophisticated mechanisms at work behind the scenes, which refine and adapt content to meet individual learner's needs. One such component is the application's ability to deliver encyclopedic knowledge through dynamically processed texts. This ensures that learners not only receive well-organized information but also benefit from tailored adjustments that respond to their progress and preferences. By effectively blending overt instruction with these intricate background processes, Appil provides a comprehensive learning environment that fosters both explicit understanding and personalized language acquisition.

Moreover, the application's design embodies a dual focus on clarity and complexity, offering learners the comfort of explicit teaching while subtly engaging them in deeper cognitive processes. This approach encourages learners to explore language beyond surface-level comprehension, inviting them to connect concepts and recognize patterns that might not be immediately obvious. By promoting awareness of the material alongside the hidden intricacies of the learning system, Appil cultivates an environment where learners can confidently navigate their language journey, blending guided instruction with self-discovery.

The integration of overt teaching methods also plays a crucial role in demystifying the language learning process for students. As they engage with clearly articulated lessons, they develop a sense of agency and confidence in their abilities. This structured exposure not only lays a solid foundation for language acquisition but also empowers learners to take ownership of their progress. The interplay between explicit teaching and the sophisticated adaptive mechanisms operating in the background ensures that students receive a rich, nuanced educational experience, allowing them to thrive in their language studies while being supported by intelligent, responsive technology.

In essence, Appil's framework illustrates the potential of combining overt teaching strategies with complex underlying processes. This synergy creates a holistic educational experience that is both informative and transformative, enabling learners to appreciate the richness of the language while benefiting from a system tailored to their individual learning trajectories. By fostering an awareness of both explicit content and the nuanced mechanisms that enhance learning, Appil empowers students to engage deeply with the language and develop the skills necessary for effective communication in diverse contexts.

As learners progress through Appil, they encounter a carefully curated curriculum that emphasizes not only the mechanics of language but also its practical application in real-world contexts. This alignment with authentic language use ensures that students grasp the relevance of what they are learning, reinforcing their motivation to engage with the material. The overt teaching strategies employed within the application, such as explicit grammar instruction and vocabulary exercises, serve

as essential building blocks for learners. These foundational elements provide clarity and direction, enabling students to map their progress and understand the linguistic components at play in their communication efforts.

In addition, Appil embraces technology to enhance the learning experience through interactive and engaging exercises that complement the overt teaching processes. By incorporating gamification elements, collaborative activities, and real-time feedback, the application transforms language learning into an active, enjoyable process. These features encourage learners to take risks, make mistakes, and learn from them in a supportive environment. The balance of structured instruction with dynamic, interactive content fosters a sense of community among learners, as they share experiences and strategies while navigating their language journey together.

Furthermore, the ability of Appil to adapt content based on individual learner data exemplifies its commitment to personalized education. As learners interact with the application, their unique learning profiles are continuously assessed, allowing for the fine-tuning of instructional materials and activities. This responsiveness not only addresses diverse learning needs but also keeps learners engaged by presenting them with challenges that are appropriately calibrated to their current proficiency level. By ensuring that learners are consistently presented with content that is slightly above their comfort zone, consistent with Krashen's 'input +1' theory, Appil facilitates ongoing growth and development in language skills.

This dual approach not only enhances the learning experience but also aligns with contemporary theories of language acquisition that advocate for a balanced, multifaceted approach to teaching. As learners navigate through the layers of explicit instruction and implicit learning opportunities, they are equipped with the tools necessary for achieving fluency and confidence in their language use.

A significant aspect of Appil's innovative design lies in its strategic use of language models and advanced language models, which enhance the learning experience while operating largely in the background. Rather than serving as a generative tool that autonomously creates new content, AI in Appil functions as an analytical and adaptive mechanism. It facilitates nuanced language processing, refines personalized support, and optimizes learner engagement without overtly drawing attention to its presence. For instance, algorithms analyze learner interactions to identify patterns, preferences, and areas needing improvement. This covert integration ensures that content and activities seamlessly adapt to each student's unique learning trajectory without replacing structured instructional design.

To further enrich the user experience, Appil incorporates a semi-gamified approach to interaction, transforming language learning into an engaging and motivating journey. Game-like elements, such as points, levels, and achievements, encourage learners to explore the learning platform with curiosity and persistence. This gamification fosters a sense of accomplishment and reinforces commitment to

ongoing practice, making learning both effective and enjoyable. Additionally, the application provides predefined guidelines and structured pathways, reminiscent of traditional teaching methods. These serve as a supportive framework, offering clear objectives and milestones that mimic the role of a teacher, ensuring learners stay focused and guided throughout their journey.

While AI-driven systems dynamically tailor exercises and challenges to individual progress, Appil does not rely on generative AI to produce open-ended content. Instead, it utilizes adaptive algorithms to adjust and repurpose existing materials, ensuring that tasks remain pedagogically sound and contextually appropriate. This balance between guided interaction and learner-driven exploration encourages self-directed learning while maintaining the structure necessary for effective language acquisition.

In summary, Appil's integration of large language models, alongside semi-gamified interactions and adaptive learning strategies, demonstrates a commitment to providing a rich, multifaceted educational experience. By employing AI as a supportive rather than generative tool, Appil preserves the integrity of structured learning while fostering autonomy and engagement. This approach ensures that students develop language skills in a dynamic yet pedagogically grounded environment, equipping them to navigate an increasingly digital learning landscape.

In the development of our didactic application, a significant challenge emerged from the need to establish a flow of operations that is primarily conducted in real-time, precluding any substantial planning or execution in advance. The timing of interventions is intricately connected to both the selection and presentation of content, which requires ongoing, real-time adjustments and refinements throughout the user's learning process. This dynamic approach ensures that the learning experience remains responsive to the learner's needs but also imposes a complex layer of operational management. As a result, a considerable number of modules are essential for implementing these modifications and conducting continuous evaluations of the presented texts. These modules play a crucial role in monitoring learner progress, assessing content relevance, and adjusting instructional strategies accordingly.

This dynamic framework highlights the inherent complexity of maintaining a balance between flexibility and structure within Appil, ensuring that the application can adapt to diverse user requirements and the ongoing evolution of content. The necessity for real-time adjustments also reinforces the importance of continuous monitoring and iterative development to ensure that the learning platform remains effective, engaging, and academically sound.

In this context, the supervision of didactic operations is divided between two critical components: a content dispatcher module and a personalized learner space. Each learner is uniquely identified and tracked, allowing for the monitoring of their choices and progression throughout the learning process.

The content dispatcher module functions as the core mechanism of the application, managing the delivery of relevant educational materials based on real-time data. It plays a pivotal role in evaluating the learner's ongoing progress, preferences, and interactions with the content, dynamically selecting and presenting resources that best align with their evolving needs. By continuously assessing the learner's engagement and performance, the module ensures that the materials presented are tailored to the learner's current proficiency level and specific learning objectives. This adaptive selection process allows Appil to maintain an individualized and responsive learning experience, where learners are consistently challenged with appropriate content while reinforcing previously acquired skills.

Through this integration of real-time data and adaptive delivery, the content dispatcher module supports the creation of a personalized educational journey for each learner, providing timely access to resources that foster continuous development and sustained motivation.

Simultaneously, the personalized learner space serves a vital function in promoting individual engagement and accountability throughout the learning process. This space meticulously tracks each learner's progress, ensuring that a comprehensive record is maintained of the choices made, activities completed, and skills acquired. By consolidating this data, the application provides learners with timely, constructive feedback on their performance, as well as deeper insights into their learning patterns. These insights help learners identify their strengths, recognize areas in need of improvement, and better understand the trajectory of their language acquisition.

Moreover, by leveraging this data, Appil can suggest targeted resources and strategies designed to support each learner's ongoing development. These personalized recommendations not only enhance the effectiveness of the learning experience but also ensure that the content remains aligned with the learner's evolving needs and objectives. This personalized approach fosters a deeper, more meaningful engagement with the material, enhancing both the motivation and the outcomes of the learning process.

Furthermore, the integration of the content dispatcher module and the personalized learner space create a seamless interaction between content delivery and learner engagement. As the content dispatcher dynamically adjusts the learning material based on real-time assessments, the personalized space allows for continuous reflection and goal setting, empowering learners to take ownership of their progress. This synergy between adaptive content and personalized reflection fosters a comprehensive learning experience, where learners are not only exposed to tailored content but are also equipped with the tools to reflect on their own development.

Ultimately, this dynamic interplay between real-time content adaptation and personalized learner tracking enhances motivation, promotes accountability, and deepens the learner's understanding of

the subject matter. By providing a holistic view of the learner's journey, Appil facilitates a more robust and self-directed approach to language acquisition, ensuring that the learning process is both effective and engaging.

To conclude, the dual oversight of didactic operations through the content dispatcher module and personalized learner spaces exemplifies the complexity and dynamism required to create an effective didactic application. By ensuring that the learning experience is both adaptive and individualized, the application can respond to the diverse needs of learners, ultimately leading to more successful educational outcomes.

4.3.15 Assessing Text Complexity

As stated in 3.2.1.1, the aim is to identify appropriate resources based on a specific array of features to integrate into our learning environment. A fundamental aspect of this integration involves understanding text complexity, which is crucial in the context of language acquisition. The complexity of a text directly influences a learner's ability to engage with, process, and ultimately comprehend written materials, making it a key determinant in the effectiveness of language learning strategies. Text complexity encompasses a variety of linguistic features, including vocabulary richness, syntactic structures, and overall coherence, all of which can significantly impact a text's accessibility and appropriateness for learners at various proficiency levels.

Vocabulary richness refers to the diversity and depth of the lexical items used in a text. Texts with a more varied and extensive vocabulary may present challenges for beginners but can enhance learning opportunities for more advanced learners. Syntactic structures, which include sentence length, complexity, and the use of different grammatical constructions, also contribute to text complexity. Sentences that are long, contain subordinate clauses, or employ complex tenses can be difficult for learners to process and understand, particularly at lower proficiency levels. Moreover, the overall coherence and cohesion of a text, the logical flow of ideas and the use of discourse markers, play a critical role in how easily learners can follow and absorb the material. A well-structured text that presents information in a clear, organized manner is far more accessible than one with abrupt shifts or lack of clarity.

In the context of Appil, understanding and addressing text complexity is a pivotal element of the adaptive learning system. By carefully selecting texts that are linguistically appropriate for learners at different stages of their language development, we ensure that the learning experience is both challenging and achievable. This careful calibration of text complexity allows learners to gradually progress through levels of increasing difficulty, thereby fostering continuous engagement, improving comprehension, and enhancing overall language proficiency.

Recent advancements in computational linguistics and natural language processing (NLP) have significantly advanced the development of tools and methodologies for systematically assessing text complexity. These developments have enabled the creation of sophisticated algorithms that can analyze a wide range of linguistic features, such as vocabulary richness, syntactic structures, and overall text coherence, on a large scale. In the context of language acquisition, these tools are invaluable for evaluating texts with precision, ensuring that the complexity of the material aligns with the learner's current proficiency level. The ability to assess text complexity in this manner is particularly important for the development of language learning applications, as it allows for the careful selection of materials that are tailored to meet the diverse needs and interests of learners.

Efforts in this field are focused not only on identifying the most suitable repositories of texts but also on selecting appropriate frameworks that ensure the assessments align with learners' individual needs and personal interests. By leveraging resources such as lexical databases (e.g., *WordNet*), syntactic analysis tools, and corpus-based studies, researchers and educators can establish frameworks that support accurate evaluations of a text's difficulty. These evaluations go beyond simplistic metrics, incorporating a holistic view of the linguistic features that contribute to text complexity. This approach is crucial for providing learners with materials that are not only appropriately challenging but also engaging, as the selection of content based on interest is equally important for motivation and long-term retention.

This section will delve into the various methods employed to calculate text complexity, including word frequency analysis, readability indices, and syntactic parsing, and explore their contribution to the development of Appil's adaptive learning system. Word frequency analysis, for instance, is a key metric that assesses the prevalence of specific vocabulary items in a text, with more frequent words generally being easier for learners to understand. By determining which words appear most often in the target language, this method helps to evaluate how accessible the vocabulary is for learners at different stages of language acquisition. Readability indices, such as the Flesch-Kincaid index or the Gunning Fog Index (Flesch, 1948; Gunning, 1952; Kincaid et al., 1975), provide an overall estimate of a text's difficulty by analyzing factors like sentence length, word complexity, and syllabic structure. These indices offer a quantifiable approach to understanding how challenging a text may be, enabling the adaptive system to select materials that align with the learner's proficiency level. Syntactic parsing, which involves the analysis of sentence structure and grammatical relationships, further refines these evaluations by identifying the complexity of sentence constructions, such as the use of subordinate clauses or passive voice. By understanding the syntactic complexity, the system can better assess how challenging a specific text may be for learners in terms of sentence comprehension.

The integration of these methods into Appil's adaptive learning system ensures that content is dynamically adjusted to suit the learner's evolving proficiency. The system continuously evaluates the learner's progress and adapts the materials, accordingly, making adjustments in real-time to maintain a balance between challenge and comprehension. This dynamic adjustment of content ensures that learners are not only presented with texts that align with their language abilities but also with materials that keep them engaged and motivated.

By systematically analyzing text complexity, Appil can provide educational materials that are appropriately challenging, neither too easy to bore the learner nor too difficult to discourage progress. Ultimately, this systematic assessment of text complexity is crucial for tailoring the educational experience, leading to more effective learning outcomes in second language acquisition by ensuring that learners consistently interact with content that promotes language growth while also maintaining their interest and confidence.

To effectively calculate text complexity, it is essential to employ a multifaceted approach that integrates both quantitative and qualitative measures. Quantitative metrics, such as readability scores, are commonly used to assess elements like sentence length, word frequency, and syllabic complexity. These factors are key indicators of a text's linguistic accessibility, providing a clear, numerical evaluation of its difficulty level. Readability scores, such as the Flesch-Kincaid index or the Gunning Fog index (Gunning, 1952; Kincaid et al., 1975), offer valuable insights into how challenging a text might be based on structural components like sentence length and word complexity. These tools enable the adaptive learning system to select texts that align with the learner's current proficiency.

However, quantitative measures alone are insufficient for capturing the full scope of a text's complexity. Qualitative assessments are also necessary, as they evaluate aspects that are harder to quantify but equally important. These assessments focus on thematic depth, the relevance of the content, and the contextual richness of the material. For example, the incorporation of idiomatic expressions, cultural nuances, or abstract concepts may increase the cognitive load on learners, even if the overall sentence structure is simple. Evaluating these qualitative factors ensures that content remains relevant and intellectually stimulating, offering learners the opportunity to engage with the target language in a meaningful way.

Additionally, integrating feedback from learners about their perceptions of difficulty further enriches the complexity assessment process. Learners' self-reported experiences, such as their perceived comprehension and engagement with a given text, provide valuable qualitative data that helps refine the algorithm's evaluations. This learner-centered approach allows the system to adjust dynamically, fine-tuning the difficulty of materials based on real-time feedback and individual learner needs.

By combining both quantitative and qualitative assessments, along with iterative feedback loops from users, a more nuanced and accurate evaluation of text complexity can be achieved. This process ensures that the selected materials are not only linguistically appropriate but also pedagogically engaging. Ultimately, the integration of these multiple evaluation methods facilitates the creation of a learning environment where texts are tailored to the learner's evolving needs, fostering meaningful language learning experiences and promoting sustained motivation.

In our exploration of the identified repositories, we discovered a substantial number of dictionaries; however, we faced the challenge of establishing connections between these inventories, the obstacles encountered by second language learners, and the rich tapestry of language diversity. To address this complexity, we applied a series of computations focused on Italian-English lexical relationships. This approach allows us to create a more cohesive and effective learning environment that enhances language acquisition while accommodating the unique needs of learners from diverse linguistic backgrounds.

4.3.15.1 Related works in terms of text complexity assessment

To deepen our understanding of text complexity and its implications for second language acquisition, it is essential to integrate readability indexes with frequency analysis. Readability indexes, such as the Flesch-Kincaid Grade Level (Kincaid et al., 1975) and the Gunning Fog Index, offer quantitative measures of text accessibility by evaluating factors like sentence length, word length, and syllable count. These metrics complement Paul Nation's frequency levels by adding an extra layer of complexity assessment, helping determine how easily a text can be understood by learners at different proficiency levels.

By incorporating readability metrics into our computational framework, we can better align texts with the evolving needs of second language learners. This alignment ensures that texts are appropriately challenging, neither too simple nor too difficult, supporting gradual vocabulary development and fostering comprehension. The combination of frequency-based vocabulary analysis and readability formulas creates a dynamic, adaptable approach to selecting and adjusting learning materials. This approach offers a balanced learning experience that evolves with the learner's progress, enhancing both engagement and skill development.

In recent years, the evaluation of text complexity has increasingly gained prominence as a fundamental aspect of designing and refining effective tools for second language acquisition (SLA). The task of assessing text complexity is crucial not only for enhancing language learning resources but also for facilitating learners' engagement with texts that are appropriately challenging. Diverse methodologies have emerged to tackle this issue, ranging from traditional approaches grounded in

the analysis of lexical and syntactic features to more sophisticated, contemporary models that employ the capabilities of large language models (LLMs and machine learning (ML)). These approaches have found applications in a variety of academic disciplines, including linguistics, computational modeling, and applied linguistics, reflecting the growing interdisciplinary interest in text complexity. Given the relevance of this topic, it is essential to review and analyze key works that have significantly contributed to the development of the field, particularly those that focus on modern, widely adopted methods. In this section, we will provide an overview of these key contributions, emphasizing the contemporary approaches that have been both widely utilized and proposed in the extant literature on text complexity assessment.

A prominent approach to the assessment of text complexity is grounded in the analysis of linguistic features, including syntax, morphology, and semantics. Linguistic analysis has long been recognized as an essential component in the evaluation of text difficulty, as these features offer insights into the underlying structure and meaning of a text. Numerous studies (Crossley, 2025; Zhang & Lu, 2025) have explored and identified several key linguistic features that influence the perceived complexity of a text. Of particular interest are aspects such as sentence structure, lexical density, and overall readability, which have been found to be closely linked to the cognitive demands placed on readers. Sentence structure, for example, plays a pivotal role in determining how easily a text can be processed and understood, with more complex syntactic constructions generally posing greater challenges for readers. Similarly, lexical density, which refers to the concentration of content words in a text, has been shown to affect the ease with which readers can extract meaning from the text. Additionally, readability metrics, which often incorporate both lexical and syntactic factors, serve as widely recognized indicators of text complexity and have been frequently employed in the development of readability formulas and assessment tools.

Syntactic complexity, which pertains to the structure and length of sentences, has long been recognized as a crucial indicator of text difficulty. Notable research by Biber, Gray, and Poonpon (Biber et al., 2011; Biber & Gray, 2010) have highlighted that texts characterized by greater syntactic complexity generally result in increased cognitive load for readers, making comprehension more demanding. The relationship between syntactic structure and text difficulty is grounded in the notion that more complex sentence constructions require greater attention and processing resources to decode and interpret. To quantify syntactic complexity, various indices have been developed, with two of the most commonly employed being the Mean Length of T-unit (MLU) and the Subordination Index (SI). These indices measure the number and types of clauses present within a sentence, providing a standardized means of assessing the complexity of syntactic structures across different texts.

In addition to syntactic features, lexical characteristics also play a fundamental role in determining the complexity of a text. The frequency of words used within a text is particularly significant in this regard, as less frequent words, specialized terminology, and low-frequency word forms tend to contribute to greater complexity. For example, texts that incorporate highly technical or domain-specific vocabulary generally present increased challenges for readers, especially those who are still acquiring proficiency in the target language. In contrast, texts that predominantly feature high-frequency, familiar words, as often categorized in Nation's vocabulary bands, are typically easier for learners to process, making them more accessible and approachable. However, lexical complexity encompasses a broader range of factors beyond word frequency, including attributes such as word length, the presence of derivational affixes, and the use of collocations. These aspects can further enhance or diminish the overall complexity of a text. In recent years, computational methodologies have been increasingly applied to assess lexical complexity, with tools such as Lexical Frequency Profiles (LFP) and principles grounded in Zipf's Law being employed to analyze the distribution of word frequencies within a text. These computational approaches facilitate the prediction of how easily learners will be able to process and internalize new vocabulary, offering valuable insights into the cognitive demands of specific texts.

Although many contemporary approaches to text complexity assessment increasingly incorporate advanced techniques such as deep learning and machine learning models, traditional features continue to play a significant and indispensable role in evaluating the difficulty of a text. These traditional features, which are often rooted in basic yet effective linguistic metrics, offer valuable insights into the syntactic and lexical complexity of a text. Basic measures such as the average number of characters per word, syllables per word, and sentence length remain essential tools in the analysis of text complexity due to their ability to capture foundational aspects of a text's structure and accessibility. For instance, the average number of characters per word and the syllables per word provide quantifiable indicators of lexical difficulty, as longer and more complex words tend to increase the cognitive load on readers. Similarly, sentence length serves as a direct measure of syntactic complexity, as longer sentences typically involve more intricate constructions that may demand greater processing resources from readers. These metrics, while relatively simple in nature, remain highly relevant in the assessment of text complexity, particularly in contexts where computational resources or access to advanced models may be limited. In this regard, traditional features continue to complement more sophisticated approaches, offering a valuable balance of simplicity and precision in the evaluation process.

Traditional syntactic features remain a central component in readability assessments, with the Mean Length of Sentence (MLS) (Vajjala & Meurers, 2012; Zamanian & Heydari, 2012) being one of the

most widely employed measures. MLS refers to the average number of words per sentence and has been directly linked to text complexity. Studies have shown that texts with longer sentences tend to be more complex, as they require more cognitive effort to process and comprehend. The length of sentences is therefore a critical determinant of a text's difficulty, with shorter sentences generally being easier to understand, especially for language learners or readers with lower proficiency levels. Consequently, sentence length serves as a fundamental indicator of the cognitive demands placed on readers, impacting their ability to engage with the text effectively.

In addition to syntactic features, traditional lexical features are also essential in evaluating text complexity. Two such measures, the Average Number of Characters per Word (*NumChar*) and the Average Number of Syllables per Word (*NumSyll*), provide valuable insights into the lexical difficulty of a text (Vajjala & Meurers, 2012). Texts that contain longer words or words with more syllables are typically more challenging for readers, particularly for those who are still developing their language skills. Longer words and greater syllabic counts often correlate with more complex vocabulary and, as a result, increase the difficulty of comprehension. These traditional lexical metrics, when combined with other readability measures, offer a comprehensive picture of the text's overall complexity, helping to predict how accessible or challenging it may be for various audiences.

To assess text complexity in a systematic and standardized way, several traditional readability formulae have been developed over the years. Two of the most commonly used formulae in the field are the Flesch-Kincaid Grade Level and the Coleman-Liau Index (Coleman & Liau, 1975), both of which rely on the analysis of these basic linguistic features to estimate the difficulty of a text. The Flesch-Kincaid Grade Level (F-K) formula (Kincaid et al., 1975), for example, calculates the difficulty of a text by considering the average number of syllables per word and the average number of words per sentence. The resulting score indicates the grade level required to understand the text, making it a widely utilized tool in educational settings to evaluate reading materials for different age groups. Similarly, the Coleman-Liau Index is another readability formula that provides a grade level score, but it differs from the Flesch-Kincaid formula in that it focuses on the average number of characters per word and the average number of sentences per 100 words. While both formulae offer valuable estimates of readability, their reliance on different features, syllables versus characters, results in slight variations in the complexity scores they produce.

Both the Flesch-Kincaid Grade Level and the Coleman-Liau Index, often categorized as Traditional Formulae, have played a pivotal role in the evolution of text complexity assessments. These formulas are widely used in various educational contexts to provide a quick and accessible estimate of text difficulty based on fundamental linguistic features. By focusing on basic syntactic and lexical elements, such as sentence length and word characteristics, these traditional methods offer a

straightforward approach for evaluating readability. While they have proven effective in many cases, it is important to acknowledge that these traditional formulae, in their simplicity, may not fully capture the multifaceted nature of text complexity. Specifically, they may overlook critical elements such as discourse structure, thematic depth, or the broader cognitive demands that texts may place on readers. Nonetheless, these formulae remain invaluable tools for quickly assessing the basic structural features of a text, providing educators and researchers with useful, albeit simplified, insights into its accessibility. As a result, they continue to be widely implemented in educational settings to assess reading materials and tailor them to specific learner groups.

With the advent of Large Language Models (LLMs), artificial intelligence (AI) and machine learning (ML), there has been a notable shift toward more sophisticated and nuanced approaches to text complexity assessment. These modern models are able to go beyond the scope of traditional formulae by incorporating both linguistic and non-linguistic features of a text. For instance, AI and ML models can analyze a broader range of features, including rhetorical patterns, text coherence, and contextual factors, which contribute to the overall difficulty of a text. By leveraging large-scale corpora of language data, these models can predict text difficulty with a higher degree of precision, allowing for more individualized and context-specific assessments. This shift represents a significant advancement in the field, as AI and ML-based models provide a more comprehensive understanding of text complexity, taking into account a variety of variables that were previously difficult to quantify. As these technologies continue to evolve, their integration into text complexity assessments promises to enhance the accuracy and versatility of readability tools, offering a more refined approach to evaluating text difficulty for diverse audiences.

Recent research has increasingly explored the use of deep learning algorithms to predict text complexity by processing multiple layers of textual data. These models, such as Recurrent Neural Networks (RNNs) and Transformers, particularly pre-trained models like BERT (Bidirectional Encoder Representations from Transformers), have proven to be effective tools for assessing text difficulty. These advanced models function by analyzing patterns across vast datasets, learning how various linguistic and contextual features interact, and applying this acquired knowledge to evaluate the complexity of new, previously unseen texts (Devlin et al., 2019). Unlike traditional methods that primarily focus on syntactic and lexical features, deep learning models incorporate a wider range of factors. They assess not only sentence structure and word frequency but also contextual and semantic cues, which allows for a more nuanced understanding of a text's difficulty. By leveraging these capabilities, deep learning models offer a comprehensive analysis of text complexity, which extends beyond surface-level features to include the interpretation of meaning and context within the text.

Another prominent machine learning technique used in text complexity assessment is the Support Vector Machine (SVM). SVMs are powerful classifiers that have been applied to categorize texts based on their level of complexity. This method involves inputting a variety of linguistic features, such as sentence length, vocabulary richness, and syntactic variation, into the model. The machine learning algorithm then trains on a labeled corpus, which consists of texts that vary in complexity, allowing the model to learn how to classify new texts according to their difficulty. The application of SVMs has been particularly valuable in categorizing texts into distinct levels of complexity, aligning with established proficiency scales such as the Common European Framework of Reference for Languages (CEFR). Through this process, SVMs enable the accurate classification of texts, providing a reliable framework for evaluating texts across a range of proficiency levels.

While traditional methods of assessing text complexity have largely focused on syntactic and lexical features, recent advancements in the field have placed greater emphasis on semantic complexity. Semantics pertains to the meaning of words, phrases, and sentences within a text, and it plays a crucial role in determining how easily a reader can comprehend and interpret the content. As texts become more sophisticated, understanding the interaction between syntax, vocabulary, and meaning becomes essential for evaluating their overall complexity. By considering semantic complexity, researchers and educators can gain a deeper insight into the cognitive demands placed on readers, further enhancing the precision of text complexity assessments.

A notable method for assessing semantic complexity involves the integration of linguistic resources such as *WordNet*, a comprehensive lexical database of the English language. *WordNet* includes various relationships between words, such as synonyms, antonyms, hyponyms, and hypernyms. By mapping these semantic relationships, researchers can gain valuable insights into the richness and depth of vocabulary within a text. In the context of second language learning, texts that incorporate multiple layers of semantic relationships tend to be more challenging for learners, as they require a nuanced understanding of the subtle distinctions between words and their meanings. Moreover, integrating linked data from external sources has been explored to further enhance semantic assessments. The use of linked data techniques allows for a deeper understanding of the relationships between concepts within a text, providing additional context that aids in gauging its complexity for learners.

Recent research has also delved into the use of topic modeling techniques, such as Latent Dirichlet Allocation (LDA), to assess the thematic complexity of texts. Topic modeling algorithms analyze the distribution of words and phrases across different topics, enabling the identification of texts that engage with multiple or advanced themes. Thematic depth introduces an additional layer of complexity, as learners must not only comprehend individual words but also grasp the broader context

and subject matter being addressed. This deeper engagement with themes further challenges learners, as it requires them to interpret the underlying meaning and connections between topics.

In recent years, there has been a noticeable shift towards hybrid approaches that combine linguistic, machine learning, and semantic analyses. These hybrid methods aim to capitalize on the strengths of each approach, thereby enhancing the accuracy and adaptability of text complexity assessments. Some studies have integrated traditional syntactic and lexical features with more advanced machine learning and semantic models to create a more comprehensive framework for evaluating text difficulty. For example, Chang et al. (2017) developed a Deep Readability Prediction model that combines readability scores, syntactic features, and deep learning techniques to predict the difficulty level of a text. This hybrid approach holds significant potential for providing more precise and context-sensitive results, offering a more holistic understanding of the multifaceted nature of text complexity.

Hybrid models are particularly well-suited to adaptive learning environments. By incorporating real-time data and learner feedback, these models can dynamically assess text complexity, adjusting the material to align with the individual learner's proficiency level. For instance, platforms such as Appil use real-time data to modify the difficulty of texts according to the learner's progress, ensuring a more personalized and engaging experience. The integration of machine learning algorithms with semantic resources enables these platforms to gain a deeper understanding not only of a text's difficulty but also of its relevance and interest to each learner. This allows the learning platform to present content that is both appropriate in terms of complexity and engaging, enhancing the learning process by maintaining the right balance of challenge and comprehensibility.

In conclusion, the field of text complexity assessment has undergone substantial evolution, progressing from traditional readability indices to the incorporation of more advanced linguistic, machine learning, and semantic methods. While classic approaches, such as syntactic complexity analysis and vocabulary profiling, continue to form the foundation of text assessment, newer technologies like deep learning models, semantic networks, and hybrid systems offer powerful tools for refining text evaluations in the context of second language acquisition. These diverse methods enable more nuanced and dynamic assessments of text difficulty, which are critical for providing learners with engaging, comprehensible, and appropriately challenging materials tailored to their individual learning needs.

As computational linguistics and educational technology continue to evolve, these models are expected to play an increasingly central role in enhancing adaptive language learning platforms. In our case, we have been inspired by recent advancements in text complexity assessment but have developed a custom approach that integrates a multi-component index, which we present in the

following section. This tailored approach combines elements from various models, allowing for a more comprehensive evaluation of text difficulty. By ensuring that learners are exposed to content that is both suitably challenging and accessible, our method contributes to the creation of more effective and personalized educational experiences, better aligned with the diverse needs of learners.

4.3.15.2 Appil complexity assessment index

Our conceptualization of text complexity builds on the seminal work of Paul Nation, whose division of the English lexicon into ten frequency bands provides a foundational framework for vocabulary acquisition. This categorization highlights which lexical items are most essential at different stages of proficiency, thereby ensuring that the vocabulary selected for instructional materials mirrors authentic language use and strengthens both the relevance and credibility of the learning experience. Yet, while Nation's frequency levels offer a solid baseline, our approach acknowledges that text complexity in second language learning is inherently dynamic rather than fixed. It is shaped by a wide range of factors, including learner motivation and engagement, contextual cues, syntactic structure, and the interaction between an individual's proficiency level and the demands of the content. The central challenge, therefore, is to develop a framework that not only incorporates these interrelated dimensions but also adapts flexibly to learners' diverse needs and evolving competencies. Nation's work categorizes words into different frequency levels based on their occurrence in large corpora, providing valuable insights into how vocabulary usage influences text difficulty. By acknowledging these frequency levels, we establish a foundational understanding of text complexity, recognizing that higher-frequency words are generally easier for learners to process, while less frequent words present greater challenges. This starting point allows us to ground our approach in a solid linguistic framework, ensuring vocabulary selection aligns with real-world language use and learner proficiency.

Crossley and colleagues (Crossley et al., 2014) demonstrate that second language (L2) lexical acquisition is shaped by a constellation of interrelated factors that extend beyond simple word frequency in native speaker (NS) input. This challenges the traditional view that learners acquire vocabulary in direct proportion to its frequency of occurrence in the target language. Although it is often assumed that L2 learners build their lexicon according to the statistical distributions of the words they encounter, empirical evidence suggests that early L2 word production does not consistently follow the Zipfian frequency patterns observed in native usage. Instead, learners frequently produce words that are relatively rare in the target language.

This tendency may arise from several sources, including lexical transfer from the first language (L1), cognitive processing constraints, and individual variation in lexical learning strategies. In particular,

learners may rely on prior linguistic knowledge, selecting words that are salient or frequent in their L1, even if these items are infrequent in the L2, especially in contexts where the two languages share numerous cognates. Furthermore, deviations from NS frequency distributions may reflect the influence of factors beyond raw frequency, such as perceptual salience, recency of exposure, contextual distinctiveness, and word prototypicality, all of which contribute to lexical selection and retrieval in L2 acquisition. Additionally, the conditional frequency of words within specific grammatical constructions, such as verb-argument structures or tense-aspect patterns, may serve as a more accurate predictor of L2 lexical acquisition than absolute frequency alone, reinforcing the notion that learners are not merely passive receivers of linguistic input but active processors who employ complex cognitive mechanisms in vocabulary development. While these findings indicate that novice learners display marked discrepancies between the frequency patterns of their input and output, longitudinal research shows that, with increased exposure and experience in the target language, learners' lexical production gradually begins to approximate the Zipfian distributions typical of native speaker (NS) speech. This progressive convergence suggests that although frequency exerts an important influence on lexical acquisition, it operates in interaction with a broader set of cognitive and linguistic factors. A more nuanced perspective is therefore required to fully account for the mechanisms driving vocabulary development in L2 contexts.

Future research should examine the interplay between frequency and additional lexical properties, such as concreteness, imageability, and familiarity, as well as draw on larger learner corpora observed over extended periods to assess the extent to which L2 frequency distributions become increasingly native-like with proficiency. Moreover, studies exploring the relationship between lexical accuracy and production frequency, particularly in connection with the acquisition of morphological and syntactic constructions, could provide richer insights into developmental trajectories across proficiency levels. By integrating these dimensions, researchers will be able to construct a more comprehensive model of L2 lexical development, one that acknowledges both frequency effects and the broader cognitive and linguistic factors that shape vocabulary learning. Addressing text complexity is pivotal for calibrating learning materials and input types to meet the evolving needs of second language learners. By systematically analyzing and adjusting text complexity, we can tailor content to better match proficiency levels, supporting more effective language acquisition and creating a more engaging learning experience. This calibration allows educators and developers to fine-tune materials in real-time. As learners progress, input can be adjusted – through vocabulary, sentence structure, or contextual support – to provide consistent challenge without overwhelming them. This dynamic adaptation ensures learners encounter input optimally matching their developmental stage.

This calibration process enables the gradual incorporation of progressive enhancements that mirror both linguistic and cognitive development. For instance, once learners demonstrate control over basic sentence structures, they can be introduced to more complex grammatical patterns. Likewise, as their lexical repertoire expands, instruction can increasingly include nuanced and idiomatic expressions. Such a dynamic conceptualization of complexity allows educators to guide learners toward advanced content in a structured manner, while simultaneously reinforcing the stability of foundational knowledge.

By integrating these diverse dimensions, our approach seeks to provide a more comprehensive and individualized assessment of text complexity, thereby fostering a learning environment that is responsive to learners' developmental trajectories and optimally supports second language acquisition.

Given that our learning environment is designed for Italian learners of English, we used the *nuovo Vocabolario di Base*⁵ as an additional lexical reference. This resource, which includes approximately 7,000 words, is based on both the frequency of word usage in contemporary texts and the perceived availability of words by native speakers. The *Vocabolario di Base* categorizes words into those of high use, common in daily language, and those that, while less frequently used, are recognized as having a high level of availability or ease of comprehension. This classification, updated in its latest version by Tullio De Mauro and Isabella Chiari, serves as an invaluable resource for not only enhancing vocabulary acquisition but also providing a deeper understanding of word usage patterns in the Italian language. Thus, it can be an equally valuable tool for learners of Italian, as it captures both the most commonly used words and those that are crucial for effective communication but may be less frequent.

To further enhance our approach, we performed an extraction of the vocabulary from the provided PDF file, which presents the *Vocabolario di Base* in a structured tabular format. The vocabulary is divided into three categories: *fondamentale*, *alto uso e alta disponibilità* (De Mauro, 2016; De Mauro & Chiari, 2016). This clear categorization allows us to align vocabulary selection with the different levels of proficiency, ensuring that learning materials reflect both the most essential and the more nuanced words, tailored to the needs of learners at various stages.

Our intent was to create a cross-referenced database that highlights the gap in vocabulary between the *Vocabolario di Base* and Paul Nation's frequency-based lexicon (Nation, 2017). While Nation's database already includes around 25,000 more words, our cross-referencing approach allows us to

⁵ Available at: <https://www.internazionale.it/opinione/tullio-de-mauro/2016/12/23/il-nuovo-vocabolario-di-base-della-lingua-italiana> (last accessed on 26/02/2025)

identify the specific vocabulary differences, providing a more granular understanding of the lexical landscape for language learners.

To facilitate this process, we obtained a translation of the 7,500 words from the Vocabolario di Base using the MarianMT model (Junczys-Dowmunt, Grundkiewicz, Dwojak, Hoang, Heafield, Neckermann, Seide, Hermann, Aji, et al., 2018). By utilizing the Helsinki-NLP/opus-mt-it-en model for Italian-to-English translation and the Helsinki-NLP/opus-mt-en-it model for English-to-Italian translation (Tiedemann et al., 2024), we were able to ensure accurate and contextually appropriate translations for the words in both directions. This machine translation approach supports the development of learning materials that are both linguistically precise and aligned with the evolving needs of language learners.

After translating the VDB words, we cross-referenced the two lexicons to understand which meanings were excluded from Nation's frequency lists in comparison to the VDB. It's important to note that we started from the VDB as our pilot group consisted of Italian-speaking learners, making it more relevant to their needs. This dual-direction translation process, Italian-to-English and English-to-Italian allowed us to refine our database and ensure that the vocabulary used for language learning is consistent, accurate, and contextually appropriate in both languages. It further supports the creation of a comprehensive resource for learners, considering both Italian-to-English and English-to-Italian word relationships.

To refine the translation quality further, we implemented a validation step where the translated terms were compared against established bilingual dictionaries and manually reviewed in cases where multiple meanings were possible. This ensured that the most relevant translations were retained, particularly in instances where a word's meaning varies based on context. Additionally, we employed Named Entity Recognition (NER) and Part-of-Speech (POS) tagging to distinguish between different grammatical uses of words, helping to improve accuracy in the categorization of vocabulary items.

Further, we applied a morphological similarity analysis to assess the relationship between Italian and English words. Morphological similarity refers to the degree of resemblance between words based on their structural components, such as roots, prefixes, and suffixes. By employing stemming techniques, we extracted word roots and calculated their similarity using the *SequenceMatcher* algorithm. This allowed us to determine how closely related words were, even when their surface forms differed. Such an approach is particularly useful for identifying cognates and semantically related terms in translation tasks.

```
# Configure the stemmer for Italian and English
stemmer_it = SnowballStemmer("italian")
stemmer_en = SnowballStemmer("english")
```

Code 1 Python function to obtain the stem of the lemma with NLTK Snowball Stemmer Library

In this section, we initialize two instances of the `SnowballStemmer`⁶ class, one for Italian (`stemmer_it`) and the other for English (`stemmer_en`). `SnowballStemmer` is a widely used tool for morphological analysis, designed to reduce words to their root or base form, a process known as stemming. This step is critical in reducing the complexity of language processing tasks, as it allows the model to focus on the base forms of words rather than their various inflected forms.

```
# Function to calculate morphological similarity
def calculate_morphological_similarity(word1, word2):

    # Extract the stems of the words
    root1 = ''.join([stemmer_en.stem(word) for word in word1.split()])
    root2 = ''.join([stemmer_it.stem(word) for word in word2.split()])

    # Calculate similarity between stems using difflib
    similarity = SequenceMatcher(None, root1, root2).ratio()

    return similarity
```

Code 2 Implementation details of the morphological similarity algorithm. The function first stems each word using the appropriate Snowball stemmer for English and Italian, producing simplified root forms ('root1' and 'root2'). It then applies Python's 'SequenceMatcher' to compare the two stems and compute a similarity ratio based on shared character sequences. The resulting score ranges from 0 (no similarity) to 1 (identical morphology).

Here, the words are split into their individual components (if the input includes multiple words) and then passed through their respective stemmers. For `word1` (the English word), we apply the English stemmer (`stemmer_en`), and for `word2` (the Italian word), we apply the Italian stemmer (`stemmer_it`). This process transforms each word into its root form, which is crucial for assessing the true morphological similarity between the words, free from grammatical variations such as tense, number, or case.

After stemming both words, we use the `SequenceMatcher` class from the `difflib` library to compute the similarity ratio between the two stemmed words. The `SequenceMatcher` algorithm compares the two strings and calculates a ratio that reflects their similarity on a scale from 0 (completely different) to 1 (identical). This ratio serves as a measure of the morphological similarity between the two root

⁶ 2001, Dr Martin Porter; 2002, Richard Boulton used in Python with NLTK (`nltk.stem.snowball`) (Bird et al., 2009a)(<https://snowballstem.org/>)

forms of the words, which can be interpreted as the degree of overlap between their underlying structures.

Finally, the function returns the computed similarity score, which represents the degree of morphological similarity between the two input words, based on their root forms. This value can be further analyzed to assess the relationship between the words, such as identifying cognates or semantically related terms, which is particularly useful in tasks like machine translation or cross-linguistic lexicon alignment.

The similarity score calculated by the *SequenceMatcher* ranges between 0 and 1, where a value closer to 1 indicates that the two words are highly similar, while a value closer to 0 suggests that the words have little to no similarity. This ratio is a normalized measure of morphological alignment, reflecting the degree of overlap between the stemmed forms of the two words. In the context of linguistic analysis, a higher similarity score indicates a closer morphological relationship, which is particularly useful in identifying cognates, semantically related terms, or words with shared roots, even when their surface forms differ. As such, the similarity score can be interpreted as a quantitative indicator of how closely related two words are, both morphologically and semantically, with a higher value signaling a stronger degree of similarity.

The method described in this script is an efficient way to quantify the morphological similarity between words from different languages, leveraging the power of stemming and string comparison algorithms. This approach is particularly valuable in the context of multilingual natural language processing (NLP) tasks, where it is crucial to match words that may have different surface forms but share the same underlying meaning or root. The combination of stemming and the *SequenceMatcher* allows for the analysis of linguistic relationships at a more granular level, enhancing the accuracy of tasks such as bilingual lexicon construction, machine translation, and cross-lingual word alignment. Additionally, we computed the Levenshtein distance between Italian and English words to quantify their phonetic and orthographic differences. The Levenshtein distance measures the minimum number of edit operations, insertions, deletions, or substitutions, required to transform one string into another. This metric is widely used in various fields, including natural language processing and computational linguistics, to assess text similarity, detect misspellings, and enhance search algorithms. By integrating both morphological similarity analysis and Levenshtein distance calculations, we refined word pairings, ensuring that lexical alignment was both meaningful and precise.

The Levenshtein distance is a well-established metric used in computational linguistics to assess the minimum number of edit operations, insertions, deletions, or substitutions, required to transform one string into another. This approach is particularly useful for comparing words in different languages, as it accounts for various types of modifications that may occur during the translation process.

The algorithm operates by constructing a matrix that tracks the minimum number of operations required to transform each substring of the first word (s_1) into each corresponding substring of the second word (s_2). Initially, the matrix is populated with boundary values, where the transformation of an empty string to a non-empty string (or vice versa) requires several operations equal to the length of the string. The core of the algorithm iteratively fills the matrix by considering whether the characters at each position of the two strings are identical. If they are, no operation is needed; otherwise, the algorithm computes the cost of each possible operation (insertion, deletion, or substitution) and selects the minimum.

```
def levenshtein_steps(s1, s2):
    # Create a matrix to calculate the distance
    len_s1 = len(s1)
    len_s2 = len(s2)
    dp = [[0] * (len_s2 + 1) for _ in range(len_s1 + 1)]

    # Initialize the matrix
    for i in range(len_s1 + 1):
        dp[i][0] = i
    for j in range(len_s2 + 1):
        dp[0][j] = j

    # Populate the matrix with Levenshtein distances
    for i in range(1, len_s1 + 1):
        for j in range(1, len_s2 + 1):
            if s1[i - 1] == s2[j - 1]:
                dp[i][j] = dp[i - 1][j - 1] # No modification needed
            else:
                dp[i][j] = min(dp[i - 1][j] + 1, # Deletion
                               dp[i][j - 1] + 1, # Insertion
                               dp[i - 1][j - 1] + 1) # Substitution

    # Retrieve the sequence of steps to transform s1 into s2
    i, j = len_s1, len_s2
    steps = []

    while i > 0 and j > 0:
        if s1[i - 1] == s2[j - 1]:
            i -= 1
            j -= 1
        else:
            if dp[i][j] == dp[i - 1][j] + 1:
                steps.append(f"Remove '{s1[i - 1]}' from position {i}")
                i -= 1
```

```

elif dp[i][j] == dp[i][j - 1] + 1:
    steps.append(f"Add '{s2[j - 1]}' at position {i + 1}")
    j -= 1
else:
    steps.append(f"Replace '{s1[i - 1]}' with '{s2[j - 1]}' at position {i}")
    i -= 1
    j -= 1

# Handle remaining characters in either string
while i > 0:
    steps.append(f"Remove '{s1[i - 1]}' from position {i}")
    i -= 1
while j > 0:
    steps.append(f"Add '{s2[j - 1]}' at position {i + 1}")
    j -= 1

return steps[::-1]

```

Code 3 Implementation of the Levenshtein distance algorithm with step tracking. This function, 'levenshtein_steps()', computes the minimum number of edit operations (insertions, deletions, substitutions) required to transform one string (s1) into another (s2). Beyond returning the numerical Levenshtein distance, it reconstructs the specific sequence of operations needed for the transformation. Each operation is described as a human-readable step, such as 'Remove', 'Add', or 'Replace', making the algorithm useful for visualizing string transformations or explaining text similarity computations in educational contexts

Once the matrix has been populated, the algorithm proceeds by tracing back from the bottom-right corner to the top-left corner, recording the specific operations (insertions, deletions, substitutions) that were required to transform s1 into s2. This step-by-step trace provides a detailed sequence of operations, which can be useful in identifying precise differences between the two words.

For example, when comparing the Italian word *giocatore* (meaning *player*) and the English word *player*, the Levenshtein distance algorithm might reveal the following steps:

These steps clearly illustrate the required operations to convert one word into another and quantify the degree of orthographic similarity between them.

To further refine the comparison between words, particularly in the context of translation tasks, we considered integrating stem distance with the Levenshtein distance. The concept of stem distance refers to the comparison of word stems, or roots, which can help capture semantic relationships between words, even when their surface forms differ significantly due to morphological changes. For instance, the words *giocatore* (player) and *giocare* (to play) share the same root, *gioc-*, suggesting a closer semantic relationship, despite differences in form and grammatical function.

```

# Example of usage
s1 = "giocatore"

```

```
s2 = "player"
steps = levenshtein_steps(s1, s2)

# Display the steps
for step in steps:
    print(step)
```

Output:

```
Delete 'g' at position 0
Delete 'i' at position 0
Delete 'a' at position 0
Delete 'c' at position 0
Delete 'o' at position 0
Substitute 't' with 'p' at position 0
Substitute 'o' with 'l' at position 1
Substitute 'r' with 'a' at position 2
Insert 'y' at position 3
Insert 'e' at position 4
Insert 'r' at position 5
```

Code 4 Example of Levenshtein transformation steps between two words. This example demonstrates how the `levenshtein_steps()` function identifies and lists the precise edit operations required to transform one string ("giocator") into another ("player"). Each step specifies the type of operation: replacement, insertion, or deletion, along with the affected character and its position in the string. This output provides an interpretable view of the morphological and orthographic differences between two lexical items.

The combination of Levenshtein distance and stem distance provides a robust framework for analyzing linguistic relationships and refining translation accuracy. By leveraging these metrics, along with additional validation steps such as Named Entity Recognition (NER), Part-of-Speech (POS) tagging, and semantic transparency analysis, we can enhance the precision and relevance of word pairings in bilingual lexicons. This methodology not only contributes to the refinement of machine translation systems but also facilitates more effective language learning tools that are grounded in both morphological and semantic analysis.

Following the translation and validation process, we integrated the results into a structured database that facilitates dynamic querying and filtering. This database serves as a key resource for educators and learners, allowing them to retrieve vocabulary based on specific linguistic criteria such as frequency, category (fundamental, high use, high availability), and grammatical function.

	id	ita	eng	vdb_level	level_nation	stem_ita	stem_eng	stem_similarity	lev
▶	1	giocare	play	1	1	gioc	play	0.2	4
	2	giocatore	player	1	NULL	giocator	player	0.285714	6
	3	comprare	buy	1	1	compr	buy	0	5
	4	comprendere	understand	1	1	comprend	understand	0.4	7
	5	lezione	lesson	1	2	lezion	lesson	0.666667	2
	6	gioco	game	1	1	gioc	game	0.25	3
	7	li	there	1	1	li	there	0	5
	8	gioia	joy	1	2	gioi	joy	0.25	3
	9	giornale	newspaper	1	2	giornal	newspap	0.125	6
	10	giornalista	journalist	1	3	giornal	journalist	0.761905	6
	11	giornata	day	1	1	giorn	day	0.181818	5
	12	giorno	day	1	1	giorn	day	0	5
	13	giovane	young	1	1	giovann	young	0.181818	5
	14	girare	turn	1	1	gir	turn	0.222222	3
	15	giro	round	1	1	gir	round	0.444444	5
	16	giù	down	1	1	giù	down	0	4

Figure 24 Snapshot of the database with computed metrics

Moreover, this database acts as a foundational reference for the development of our complexity index, which quantifies the linguistic complexity of words based on various parameters, including morphological features and semantic relationships. By linking this resource with existing lexical databases, we enriched the dataset, thus supporting more robust vocabulary learning strategies. This enriched dataset, in conjunction with the complexity index, provides a comprehensive framework for understanding and acquiring vocabulary in a more targeted and efficient manner.

```
# Text preprocessing function
def preprocess_text(text):
    # Tokenize the text into words
    words = text.lower().split()

    # Remove stopwords
    stop_words = set(stopwords.words("english"))
    filtered_words = [word for word in words if word not in stop_words]

    # Initialize the stemmer
    stemmer = PorterStemmer()

    # Apply stemming and store both the original word and its stem
    words_and_stems = [(word, stemmer.stem(word)) for word in filtered_words]

    return words_and_stems
```


Code 5 Text preprocessing function for lexical simplification and normalization. The `preprocess_text()` function performs a series of text-cleaning operations to prepare raw input for linguistic analysis. It first converts text to lowercase and tokenizes it into individual words. Stopwords, common words that carry little semantic weight, are removed to reduce noise. The function then applies the Porter stemming algorithm, generating a list of tuples containing each original word and its corresponding stem. This preprocessing step helps standardize lexical forms and supports downstream tasks such as similarity analysis or vocabulary profiling

In the context of our application's development, the complexity index is designed to process freely chosen texts and establish a complexity threshold. This threshold allows us to calibrate enhancements accordingly, ensuring a more tailored and effective approach to vocabulary learning. By linking this resource with existing lexical databases, we have enriched the dataset, further supporting robust vocabulary acquisition strategies.

The `preprocess_text` function is designed to process a given input text by performing several important tasks commonly employed in Natural Language Processing (NLP). First, it tokenizes the text by converting it into lowercase and splitting it into individual words. The stopwords, which are common words like *and*, *the*, and *is*, are removed from the tokenized words using a pre-existing list from the NLTK (Bird et al., 2009a) library. This step helps to reduce noise and focus on the content-bearing words. Subsequently, a Porter stemmer is used to reduce each word to its root form (or stem). This process helps to group different inflected forms of a word under a common base form. The function returns to a list of tuples, where each tuple contains the original word and its corresponding stem.

```

# Function to retrieve semantic similarity from the MySQL database
def get_stem_similarity(stem, word):
    # First query: search for similarity using the full word
    cursor.execute("SELECT stem_similarity FROM metrics WHERE eng = %s",
(word,))
    results = cursor.fetchall()

    if results:
        # If results are found for the full word, calculate the average
similarity value
        print(f"Word found: {word}")
        print(f"Results: {results}")
        average_similarity = sum(float(result[0]) for result in results) /
len(results)
        return average_similarity
    else:
        # If no results for the full word, search using the stem
        cursor.execute("SELECT stem_similarity FROM metrics WHERE stem_eng =
%s", (stem,))
        results = cursor.fetchall()

        if results:
            # If results are found for the stem, calculate the average similarity
            average_similarity = sum(float(result[0]) for result in results) /
len(results)
            return average_similarity
        else:
            # If no results are found for either the word or the stem, return
0.0
            return 0.0

```

Code 6 Python function for retrieving semantic similarity from a MySQL database. This function, `get_stem_similarity()`, queries a MySQL database to obtain semantic similarity values for a given English word or its stem. The procedure first searches for direct matches of the full word in the metrics table. If no records are found, it performs a secondary lookup using the word stem. When multiple similarity values are returned, the function computes their arithmetic mean, providing a single representative similarity score. If no entries exist for either form, the function returns a default value of 0.0, ensuring robust handling of missing data.

The `get_stem_similarity` function calculates the semantic similarity of a given word and its corresponding stem by querying a MySQL database. Initially, the function attempts to retrieve similarity scores for the full word from the metrics table. If results are found, the function computes the average similarity score across all records. If no results are found for the word, it then checks for

similarity scores associated with the word's stem. If no data is found for either the full word or its stem, the function returns to a default similarity score of 0.0. This approach allows for flexible querying and handling of both full word and stem-based similarity metrics.

```
# Function to calculate the semantic similarity of a text
def calculate_semantic_similarity(text):
    words_and_stems = preprocess_text(text)

    # Remove duplicates (each word/stem is counted only once)
    unique_words_and_stems = set(words_and_stems)

    # Calculate the total similarity
    total_similarity = 0.0 # Use a float for the total
    for word, stem in unique_words_and_stems:
        # Pass both the word and its stem to obtain the similarity value
        similarity = get_stem_similarity(stem, word)
        total_similarity += similarity

    # Normalize similarity by the number of stems
    if unique_words_and_stems:
        normalized_similarity = total_similarity /
len(unique_words_and_stems)
    else:
        normalized_similarity = 0.0 # If there are no stems, similarity
is 0

    # Limit similarity to a maximum of 1
    return min(normalized_similarity, 1.0)
```

Code 7 Python function for calculating text-level semantic similarity. This function, `calculate_semantic_similarity()`, computes the overall semantic similarity of a given text by integrating word-level similarity scores retrieved from a database. The process begins with text preprocessing and stemming to extract unique lexical items. Each word–stem pair is passed to the `get_stem_similarity()` function, which returns a similarity value based on stored semantic metrics. The cumulative similarity is then normalized by the number of unique stems, producing a score between 0 and 1. This approach provides a scalable metric for assessing the semantic cohesion and lexical relatedness of a text sample.

The `calculate_semantic_similarity` function includes the previously defined `preprocess_text` and `get_stem_similarity` functions to calculate the semantic similarity of an entire input text. Initially, it processes the text to obtain the list of words and stems. To ensure that

each word and stem pair is counted only once, duplicates are removed by converting the list into a set. The function then iterates over the unique word-stem pairs and calculates the similarity for each pair using the `get_stem_similarity` function. The total similarity is accumulated, and the average is computed by normalizing the total similarity by the number of unique stems. Finally, the function ensures that the similarity score is capped at a maximum value of 1.0, providing a final semantic similarity score between 0.0 and 1.0.

```
# Function to calculate and display semantic similarity
def on_calculate_similarity():
    text = text_entry.get("1.0", "end-1c") # Get the entered text
    if text:
        similarity = calculate_semantic_similarity(text)
        messagebox.showinfo("Similarity Result", f"Normalized total
semantic similarity: {similarity:.4f}")
    else:
        messagebox.showwarning("Error", "Please enter some text before
calculating similarity.")
```

Code 8 Function for calculating and displaying semantic similarity results in a GUI. The `on_calculate_similarity()` function serves as a user interface handler for the semantic similarity module. It retrieves user-input text from a graphical entry widget, processes it through the `calculate_semantic_similarity()` function, and displays the resulting normalized similarity score in a message box. If no text is provided, the function generates a warning prompt. This integration of computational processing with a graphical user interface enhances usability, enabling users to perform semantic evaluations interactively and view immediate feedback.

The `on_calculate_similarity` function is an event handler designed for a graphical user interface (GUI). It retrieves the input text from the text entry widget in the interface, calculates the semantic similarity using the `calculate_semantic_similarity` function, and then displays the resulting similarity score in a pop-up message box. If no text is entered, the function raises a warning to prompt the user to input text before performing the similarity calculation.

The logic includes a mechanism to handle cases where multiple words share the same stem in the database. For each word, the system first retrieves its similarity score from the database. If multiple entries exist for a word (e.g., if it shares the same stem with other words), the program computes the average of the retrieved similarity values. This averaging process ensures that the final similarity score for a word reflects the overall similarity of all related forms, contributing to a more nuanced and accurate similarity measure.

For example, consider the word *moment*. The program retrieves two similarity scores: 0.923077 and 0.333333. By averaging these values, the program generates a more balanced similarity score,

enhancing the semantic similarity calculation for the entire text. Similarly, for words like *pleasant* and *music*, which appear multiple times in the database with different similarity values, the program computes the average across all these values to ensure that their contribution to the overall similarity score is appropriately represented.

This approach improves the robustness of the semantic similarity measure, as it accounts for variations in meaning that can arise from different word forms sharing the same root. By calculating and averaging the similarity scores, the program provides a more reliable and accurate assessment of semantic similarity, especially when working with diverse lexical forms and variations.

Given the output you provided, where each word appears with associated similarity values (e.g., *moment* having two similarity scores), the program will compute the average similarity for words that share the same stem in the database. Here's an example of how this process works for a word like *moment*. For *moment*, the retrieved similarity scores are (0.923077, 0.333333). The program will compute the average similarity as:

$$\text{Average similarity for } \textit{moment} = \frac{0.923077 + 0.333333}{2} = 0.628205$$

Formula 7 Example of average similarity score computation

This average similarity score will then be used in the calculation of the total semantic similarity for the entire text, ensuring a more consistent and accurate similarity measure. By automatically calculating the average similarity between words that share the same stem, the program better captures the nuances of semantic relationships between different word forms. This methodology provides a more comprehensive assessment of text similarity, particularly in cases where multiple related words contribute to the overall meaning. Building upon this foundation, an index capable of calculating text complexity based on interlanguage similarity and cognate transparency offers several advantages, particularly in the fields of language acquisition, translation studies, and computational linguistics. By incorporating interlanguage similarity, the index can assess the extent to which a given text aligns with the linguistic structures and lexical resources of different languages, thereby enhancing its accessibility for multilingual users. This approach is particularly beneficial for language learners, as it allows for the identification of texts that are more comprehensible based on their existing linguistic knowledge. Furthermore, the integration of cognate transparency into the complexity assessment provides an additional layer of refinement by distinguishing words that share etymological and semantic similarities across languages. Cognates, which are often intuitively understood by speakers of related languages, can significantly reduce cognitive load and facilitate vocabulary acquisition. By leveraging these factors, the complexity index can dynamically adjust text

difficulty, ensuring that learners receive content tailored to their linguistic proficiency. This adaptability supports more effective language learning strategies, enhances the efficiency of text simplification processes, and contributes to the development of advanced computational models for automated text adaptation. Moreover, the index can serve as a valuable tool in cross-linguistic research by providing quantitative insights into lexical and structural similarities across languages, thereby fostering a deeper understanding of language relationships and acquisition processes.

The primary objective of our index is to establish a robust comparative framework that captures and quantifies form-meaning relationships across the two languages. The initial step in constructing the bilingual index involved the careful selection of high-frequency and pedagogically relevant vocabulary for both English and Italian. This was essential not only for ensuring the comparability of lexical items across the two languages, but also for enabling a meaningful analysis of semantic transparency grounded in language learners' experience. Central to this objective is the concept of semantic and lexical transparency, which guided the selection of vocabulary and the alignment of corresponding terms. By integrating computational techniques such as Levenshtein distance and word stem analysis, the methodology aims to yield a more precise indicator of interlinguistic complexity, grounded in measurable linguistic features. The linguistic resources used, Nation's word lists for English (Nation, 2006, 2017, 2001), Coxhead's New Academic Word List (AWL) (Coxhead, 2000), and the *Nuovo vocabolario di base della lingua italiana* by De Mauro for Italian, served as the foundation for vocabulary selection and analysis, providing weighted lexical references essential for assessing semantic transparency. These resources were chosen not only for their empirical validation in corpus linguistics but also for their pedagogical alignment with school-level language instruction. By grounding the similarity index in these curated lexical sets, the model aligns with expected lexical competence in formal education, offering a tool that can meaningfully reflect the linguistic demands placed on learners.

For English language analysis, Nation's BNC/COCA (BNC Consortium, 2007) Frequency Lists (Nation, 2001, 2006, 2017), hereafter referred as NBC-FL(s) for conciseness, were adopted due to their well-established utility in language teaching, lexical profiling, and curriculum design. The NBC-FL are designed primarily for learners of English as a foreign language and divide the lexicon into ten distinct levels based on frequency and range, allowing for a fine-grained stratification of vocabulary. Each level reflects how commonly a word appears across diverse texts, which is particularly useful for modeling vocabulary accessibility and potential semantic transparency from a learner-oriented perspective. The NBC-FLs consist of 28 core word family lists based on frequency and range data, supplemented by five additional lists covering proper names, marginal words (such as swear words, exclamations, and letters of the alphabet), transparent compounds, acronyms, and

foreign words. Prioritizing words that support opportunities for use (e.g., foreign travel, academic study, internet communication) and learning (e.g., graded readers and coursebooks), the NBC-FLs ensure that essential survival vocabulary appears early. Consequently, a substantial portion of the Academic Word List is located within the first five bands. Although most of the AWL word families are already embedded within the early levels of the NBC-FL, we chose to treat the AWL separately. Nation (2006) noted that in frequency lists designed to reflect native-speaker vocabulary size growth, most of the AWL would typically emerge around the fourth or fifth thousand-word families. However, because the AWL focuses specifically on high-frequency academic vocabulary that is less common in general discourse, maintaining it as a separate entity allowed for more precise modeling of academic language proficiency. Coxhead's AWL (Coxhead, 2000) offers a curated selection of 570-word families frequent across a broad range of academic texts. By treating the AWL separately and integrating it alongside the NBC-FL, we aimed to capture not only the breadth of general English usage but also the specialized vocabulary demands that characterize academic communication.

For Italian language, we adopted the *Nuovo vocabolario di base della lingua italiana* (De Mauro, 2016b), hereafter also referred as the NVdB, developed by Tullio De Mauro, was selected as the primary lexical resource for Italian. This inventory classifies approximately 7,000 lemmas into three core categories: *vocabolario fondamentale* (fundamental vocabulary), *vocabolario di alto uso* (high-usage vocabulary), and *vocabolario di alta disponibilità* (high-availability vocabulary). Together, these levels offer a pedagogically oriented and psycholinguistically informed framework for identifying essential vocabulary in contemporary Italian. The classification is based on a combination of frequency and contextual availability. In particular, the concept of *uso* (usage) is calculated as the product of a word's absolute frequency and its distribution across different categories of text. This dual criterion corrects for distortions caused by high frequency in narrowly defined domains and ensures that selected words are both common and broadly representative of general language use. The NVdB is grounded in a manually verified, electronically processed corpus consisting of 18,843,459 lemmatized word occurrences, drawn from six equally weighted categories: print media (e.g., newspapers and magazines), academic writing (e.g., essays and textbooks), literary texts (fiction and poetry), performance texts (film and theatre scripts), computer-mediated communication (e.g., chat messages), and spoken language recordings. This structure reflects the notion of a balanced corpus, which aims to represent a language's usage across different registers and communicative contexts. Proper nouns, numerals, symbols, ideograms, and non-lexical elements were excluded, and all word forms were lemmatized and linked to lexical entries in the GRADIT dictionary (De Mauro, 1999). The final set of 7,078 lemmas in the NVdB thus represents a core vocabulary selected not only for its frequency but also for its communicative salience.

Similarly, Nation's frequency lists for English are grounded in data from two major balanced corpora: the Corpus of Contemporary American English (COCA) (Davies, 2008) and the British National Corpus (BNC) (BNC Consortium, 2007). Both corpora are constructed to represent language use across a range of communicative situations and registers. COCA includes over 560 million words drawn from spoken language, fiction, magazines, newspapers, and academic texts (Davies, 2008), while the BNC comprises 100 million words, with contributions from written sources such as fiction, newspapers, and academic prose, as well as spoken data from everyday conversations and formal speech. The use of such balanced corpora ensures that lexical frequency data reflect real-world usage and support meaningful cross-linguistic comparison in the bilingual index.

The emphasis on high-frequency lexical items in constructing the bilingual index is intentional and supported by a substantial body of psycholinguistic and applied linguistic research, which identifies frequency as a critical factor in second language acquisition (SLA). Over several decades, empirical studies have consistently demonstrated that frequency of exposure significantly influences the acquisition of vocabulary and grammatical morphemes. Learners tend to acquire high-frequency items earlier and with greater accuracy (N. C. Ellis & Schmidt, 1997; Freeman, 1975; Goldschneider & DeKeyser, 2001). Notably, even minimal repeated exposure, sometimes as few as six encounters, can lead to significant vocabulary gains (Rott, 1999). Furthermore, the inclusion of frequent lexical bundles in instructional materials fosters learner engagement with authentic language use (Koprowski, 2005). Frequent constructions also enhance processing efficiency, contributing to greater automaticity in adult learners (Arnon & Snider, 2010). Although some complex syntactic structures, such as English ditransitives, may be less responsive to frequency effects (Year & Gordon, 2009), the overall evidence underscores frequency as a central mechanism in SLA at both lexical and morphological levels. This empirically grounded and psycholinguistically informed approach underpins the value of the NVdB as a counterpart to Nation's frequency-based wordlists. The integration of Nation's stratified frequency bands with De Mauro's multidimensional lexical categorization facilitated the development of a bilingual, weighted vocabulary framework. This approach ensured the inclusion of items that are not only frequent and pedagogically relevant but also suitable for a comparative analysis of semantic transparency, accounting for usage frequency, contextual diversity, and lexical accessibility. These lexical resources provided the foundation for the subsequent computational analysis of semantic transparency.

A crucial methodological step in the construction of the bilingual index involved the alignment of lexical levels of the two reference vocabularies. While both Nation's wordlists and the *Nuovo vocabolario di base della lingua italiana* offer frequency-based, pedagogically relevant lexical inventories, they differ significantly in their structure, level of granularity, and overall lexical

coverage. This required not only selecting and aligning the vocabulary but also ensuring that the data were accessible in a machine-readable format suitable for further computational analysis.

For the Italian component, NVdB as developed by Tullio De Mauro, was initially available in a non-structured format. To enable its integration into our computational framework, a process of data extraction was performed to convert the NVdB into a machine-readable format. This was achieved by manually extracting the lexicon from the online source provided by *Internazionale* (De Mauro, 2016a), where the vocabulary was initially available in a digitally printed PDF format. To enable its integration into our computational framework, the PDF was carefully transformed and cleaned into a series of structured tables. This process involved manual extraction and reorganization of the lexemes and their corresponding categories (*vocabolario fondamentale*, *vocabolario di alto uso*, *vocabolario di alta disponibilità*) into machine-readable tables, ensuring consistency and clarity for further processing.

Nation's model consists of ten bands, each containing approximately 1,000-word families, ranked according to frequency and dispersion. This stratification supports detailed distinctions in vocabulary difficulty, commonly used in second language acquisition and reading comprehension research. The first 1,000–3,000-word families typically constitute the core lexicon essential for comprehension and productive use, while the remaining levels represent increasingly specialized or less frequent vocabulary. In contrast, the NVdB contains a more compact inventory of approximately 7,000 lemmas, categorized into three principal levels: *Vocabolario fondamentale*: the most frequent 1,991 lemmas, based on the *Lessico Italiano di Frequenza* (LIF); *Vocabolario di alto uso*: the next 2,750 high-frequency lemmas from the LIF; *Vocabolario di alta disponibilità*: 2,337 lemmas identified through various means, including frequency, but also dictionary analysis and availability in everyday communication.

To harmonize Nation's ten-level frequency framework with the Italian *Vocabolario di Base*, a proportional alignment strategy was implemented by consolidating Nation's original bands into three standardized categories: *basic*, *high_use*, and *availability*. This restructuring facilitates both cross-linguistic comparison and pedagogical alignment, addressing structural differences between the frameworks, namely, Nation's focus on word families, as compiled in his frequency lists, versus the NVdB's organization around individual lemmas, as developed by De Mauro (De Mauro & Chiari, 2016). Standardization at the level of lexical concepts, rather than surface forms, enables more meaningful cross-linguistic mapping. The three tiers correspond to major Common European Framework of Reference (CEFR) (Council of Europe, 2020) proficiency levels and expected vocabulary size: BASIC (Nation levels 1–2) aligns with CEFR A1–A2 and *foundational vocabulary* (~1,000 word families), *high use* (levels 3–5) with B1–B2 learners (~2,000–4,000 word families),

and *availability* (levels 6–10) with advanced C1–C2 proficiency, extending up to ~9,000 word families (Nation, 2001; Nation & Crabbe, 1991). This mapping closely mirrors the NVdB’s own tripartite division: *vocabolario fondamentale*, *vocabolario di alto uso*, and *vocabolario di alta disponibilità*. Specifically, the *vocabolario fondamentale* was aligned with Nation’s first 1,000-word families, reflecting the highest-frequency core vocabulary.

Furthermore, Nation’s lists are based on word families, and De Mauro’s NVdB offers a lemma-based vocabulary; it was therefore necessary to standardize the lexical units for comparison. To this end, word forms in both Italian and English were reduced to their root representations using the Natural Language Toolkit (NLTK) and the Porter stemming algorithm (Bird et al., 2009b; Porter, 2006; Sparck Jones & Willett, 1997; van Rijsbergen et al., 1980). This approach enabled a more accurate alignment based on underlying lexical units, rather than inflected or derived variants. In addition, the inclusion of the Academic Word List required a similar normalization process. The AWL, which consists of word families commonly found across academic texts, was also subjected to stemming to ensure compatibility with the other lexical resources. A relational database was created to support the bilingual index, consisting of a primary table that housed the core lexical information. This alignment strategy preserved the internal coherence of all three lexical systems, Nation’s frequency lists, De Mauro’s NVdB, and the AWL, while facilitating the creation of a harmonized, frequency-informed bilingual index. This approach enabled the development of weighted vocabulary lists that could be meaningfully compared across languages, thereby supporting the broader objective of examining semantic transparency in relation to lexical accessibility and text complexity.

To support the analysis of lexical and semantic properties across the selected corpus, a structured database was designed to store and process key linguistic and statistical features of the word forms. The database comprises two tables named *metrics* and *awl* which capture various dimensions of the words and their corresponding stems, incorporating both general and domain-specific semantic metrics. This structure allows for detailed comparisons between the original forms of the words and their normalized or stemmed counterparts, facilitating the evaluation of their semantic similarity and lexical distribution. Table 11 reflects tables’ fields such as unique identifiers (id), original word forms in Italian and English (ita, eng), and their respective stemmed forms (stem_ita, stem_eng). The table also contains several semantic similarity scores, including the raw Levenshtein distance (levenshtein_distance), its normalized counterpart (levenshtein_normalized), and a combined semantic similarity score (combined_score) that integrates multiple measures. Additionally, flags indicating the presence of words in both the internal vocabulary database (NVdB_present) and the Nation’s vocabulary list (nation_present) are included to provide context for each word’s lexical properties.

Column Name	Data Type	Description
id	bigint	Unique identifier for the entry in the dataset.
ita	text	The original Italian word or multi-word expression being analyzed.
eng	text	The corresponding English translation or equivalent.
NVdB_level	int	Level of match or alignment strength within the vocabulary database, based on predefined linguistic or frequency tiers.
NBC-FL_level	int	Level of presence or alignment in a national corpus or lexicon.
NBC-FL_level_normalized	int	Normalized form of NBC-FL_level, typically rescaled to a common range for comparative analysis.
stem_ita	text	Stemmed version(s) of the Italian term, reduced to root forms using stemming algorithms.
stem_eng	text	Stemmed version(s) of the English translation.
stem_similarity	float	Semantic or morphological similarity between the Italian and English stemmed forms, often calculated via cosine similarity or string similarity metrics.
levenshtein_distance	int	Levenshtein distance between the Italian and English terms, representing edit distance.
levenshtein_normalized	float	Levenshtein distance normalized by term length, yielding a proportion for comparison across pairs.
combined_score	float	Composite score integrating multiple similarity measures (e.g., stem similarity, normalized Levenshtein), intended to reflect overall lexical correspondence.
NVdB_present	tinyint (0/1)	Binary indicator (1 = present, 0 = absent) denoting whether the term pair exists in the vocabulary database.
nation_present	tinyint (0/1)	Binary indicator denoting whether the term appears in a national linguistic resource or lexicon.

Table 11 Multilingual lexical alignment table capturing morphological, orthographic, and corpus-based correspondences between Italian terms and their English equivalents, enriched with similarity metrics and resource presence indicators

To assess surface-level similarity, the Levenshtein distance was calculated between English-Italian cognates. The Levenshtein distance is a metric used to quantify the difference between two strings, representing the minimum number of edit operations, insertions, deletions, or substitutions, required to transform one string into another.

$$\begin{aligned} & \text{Levenshtein}(A[1..m], B[1..n]) \\ &= \min \begin{cases} \text{Levenshtein}(A[1..m-1], B) + 1 & \text{(deletion)} \\ \text{Levenshtein}(A, B[1..n-1]) + 1 & \text{(insertion)} \\ \text{Levenshtein}(A[1..m-1], B[1..n-1]) + \text{cost}(A[m], B[n]) & \text{(substitution)} \end{cases} \\ & \text{cost}(A[m], B[n]) = \begin{cases} 0 & \text{if } A[m] = B[n] \\ 1 & \text{if } A[m] \neq B[n] \end{cases} \end{aligned}$$

Formula 8 Levenshtein distance measures the minimum number of edits, insertions, deletions, or substitutions, needed to transform one string into another

This algorithm calculates the minimum number of edits (insertions, deletions, substitutions) required to transform one word into another. In a cross-linguistic context, this metric captures orthographic and phonological similarity, which often correlates with semantic proximity. As supported by Konstantinos et al. (2018), lower Levenshtein scores tend to indicate closer semantic relationships, making it a suitable proxy for evaluating transparency between cognates or closely aligned lexical items.

In the context of string comparison, particularly when dealing with short strings, the Levenshtein distance (Levenshtein, 1966) is often the most suitable metric due to its clear and intuitive representation of the number of modifications required to transform one string into another. Unlike other string similarity metrics, such as Jaro-Winkler (Jaro, 1989; Winkler, 1990), Jaccard, or Cosine similarity, which focus on set-based or vector-based representations, Levenshtein distance directly quantifies the number of edit operations, insertions, deletions, and substitutions, necessary for transforming one string into another. While Jaro-Winkler, like Levenshtein, measures the edits required for string matching, it prioritizes prefix similarity, making it particularly useful for scenarios with minor typographical errors, such as name matching or deduplication. However, Jaro-Winkler

might not capture the full spectrum of transformations as effectively as Levenshtein, especially when significant non-prefix changes occur. Furthermore, while metrics like Jaccard and Cosine similarity offer valuable insights into the overall similarity of sets or vectorized text representations, they lack the precision to detail the specific differences between strings. Consequently, for tasks requiring an understanding of the degree of transformation or the specific modifications needed for string identity, especially in contexts where fine-grained edits are important or small typographical variations are common, Levenshtein distance is the preferred choice.

To provide a more consistent measure of similarity between strings of different lengths, we calculated the Normalized Levenshtein Distance (NLD). This metric adjusts the raw Levenshtein distance by the length of the longer string, yielding a value between 0 and 1, as shown in Formula 9:

$$NLD = \left(1 - \frac{d_{lev}(a, b)}{\max(|a|, |b|)} \right)$$

Formula 9 Normalizing edit distance to enable consistent similarity scoring between strings of different lengths

where $d_{lev}(a, b)$ is the Levenshtein distance between strings a and b , and $|a|$, $|b|$ are the lengths of the respective strings. An NLD value of 1 indicates perfect similarity (i.e., the strings are identical), while a value approaching 0 suggests significant dissimilarity. By normalizing the edit distance, NLD allows for meaningful comparison across inputs of varying lengths, which is particularly useful in applications involving text similarity, error detection, and natural language processing.

To complement the surface-level metric, a deeper morphological comparison was conducted through word stem analysis. Using Python programming language (Python Software Foundation, 2022 – version 3.11) *Sequence Matcher* (Bird et al., 2009b), which identifies the longest common subsequences between word pairs, it was possible to isolate shared roots and detect underlying semantic relationships.

This approach accounts for variations introduced by inflectional and derivational morphology, offering insight into core lexical equivalence even when superficial forms differ significantly. By combining this technique with Levenshtein distance (Levenshtein, 1966), the analysis captured both surface and structural similarities, providing a multidimensional view of transparency. The *SequenceMatcher* algorithm, based on the work of Ratcliff and Obershelp in the 1980s (Ratcliff & Metzner, 1988), is an effective method for comparing sequences such as strings or lists. The algorithm aims to find the longest contiguous matching subsequence (LMS) between two sequences and recursively applies the same process to the left and right segments of the sequences. This approach does not directly yield minimal edit sequences but tends to generate matches that align well

with human intuition, making it valuable for applications such as text comparison and file differencing (Black, 2021). *SequenceMatcher* pattern recognition technique offers a pragmatic and effective approach to sequence comparison. Its recursive matching strategy and emphasis on producing human-readable results make it particularly well-suited for use cases where clarity and interpretability are valued over formal optimality. The algorithm begins by identifying the longest contiguous matching subsequence between two input sequences. Once this match is located, the algorithm partitions each sequence into three components: the matching subsequence itself, along with the preceding and succeeding segments. It then recursively applies the same matching process to these non-matching segments. This recursive division continues until no further matching subsequences can be found.

To quantify the similarity between two strings based on their internal character sequences, we have used the *diffli* module implemented in Python, which employs a normalized ratio derived from the concept of matching blocks. The formula is defined as follows, where M is the number of matching elements and T is the total number of elements in both sequences.

$$\text{ratio}_{\text{seq}}(a_{\text{italian stem}}, b_{\text{english stem}}) = \frac{2 \cdot M}{|a| + |b|}$$

Formula 10 Sequence Matcher Formula: measures the similarity between two-word stems by comparing their longest matching subsequences. Here, M represents the number of matching characters in the optimal alignment, and $|a|$ and $|b|$ are the lengths of the Italian and English stems, respectively. This normalized ratio provides a value between 0 and 1, capturing partial matches and structural similarities, and is particularly useful for identifying morphologically related words across languages

The strengths of *SequenceMatcher* lie in its conceptual simplicity, computational efficiency, and its tendency to generate results that are more interpretable than those produced by more algorithmically rigorous methods, such as dynamic programming-based approaches (e.g., Levenshtein distance). While Levenshtein distance is effective for quantifying exact transformation costs, it treats all operations equally and often fails to capture the structural or semantic alignment between sequences which is meaningful to human observers. In contrast, *SequenceMatcher* emphasizes contiguous matching subsequences, which often results in alignments that better reflect the actual similarity of natural language text or structured data. Moreover, *SequenceMatcher* is more resilient to small, localized changes, such as word reordering or phrase insertions, that may result in disproportionately high edit distances under Levenshtein-based metrics. Its recursive strategy enables it to identify larger

structural similarities even in the presence of scattered differences, making it particularly suitable for use cases involving human-authored text, code diffing, or document revision analysis.

To support the process of term pairing, especially in identifying and validating cognate relationships, the *MarianMT* (Junczys-Dowmunt, Grundkiewicz, Dwojak, Hoang, Heafield, Neckermann, Seide, Hermann, Fikri Aji, et al., 2018) neural machine translation model was employed. This model facilitated the initial generation of Italian equivalents for each English term, offering a consistent and context-sensitive basis for candidate selection. However, given the limitations inherent in automated translation systems, particularly in handling domain-specific vocabulary and subtle semantic distinctions, each entry produced by the model underwent subsequent manual review. This human supervision ensured the accuracy, appropriateness, and semantic alignment of the final bilingual pairings, thereby enhancing the overall reliability of the index.

4.3.15.3 Data compiling

The analysis of the distribution of these indicators reveals a nuanced interaction between the two lexical sources. Of the 13,808 entries in the database, 4,568 lemmas (approximately 33%) were found in both the NVdB and Nation's frequency list, indicating considerable overlap between pedagogically and communicatively relevant vocabulary. An additional 2,660 lemmas (19%) were exclusive to the NVdB, reflecting its focus on domain-specific or educationally prioritized terms not captured by general frequency data. Conversely, 6,580 lemmas (48%) appeared only in Nation's list, highlighting its broader lexical coverage of high-frequency terms used in general discourse. This distribution underscores the complementary nature of the two resources: while the NVdB provides a curated pedagogical framework, the English frequency list expands the index's lexical range.

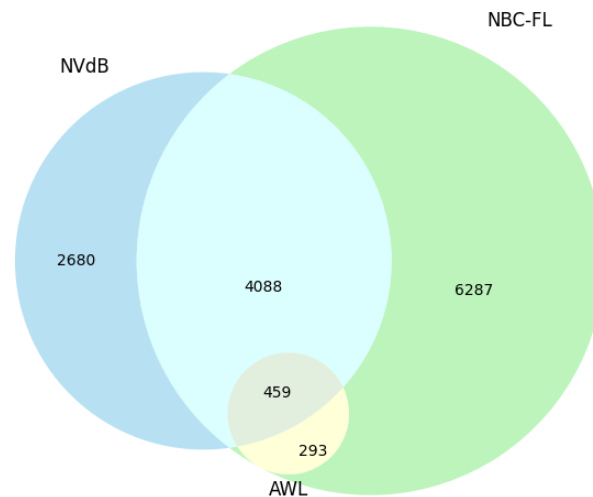


Figure 25 Distribution of lexical items across three distinct vocabularies: NVdB, NBC-FL, and AWL. Each circle represents one vocabulary, with the overlapping areas indicating the intersection of words present in multiple vocabularies

The Venn diagram (Figure 25) illustrates the distribution of lexical items across three distinct vocabularies: NVdB, NBC-FL, and AWL. Each circle in the diagram represents one of these vocabularies, with the overlapping areas highlighting words that are present in multiple vocabularies. The numerical values within each region correspond to the number of words that are either unique to a specific vocabulary or shared between different vocabularies. Specifically, the NVdB circle (blue) represents words exclusive to the NVdB vocabulary, while the NBC-FL circle (green) denotes words unique to the NBC-FL vocabulary. The AWL circle (coral) contains words found only in the AWL vocabulary. The overlapping regions indicate the words that are common to two or more vocabularies, illustrating the intersections between these lexical resources.

The distribution of presence indicators provides insight into the occurrence of specific linguistic features across the dataset.

For NVdB presence, the variable reflects the inclusion of words in a verb database, with 7,227 instances (52.3%) present and 6,580 instances (47.7%) absent. This near-equal distribution indicates that the verb database includes a significant portion of the words, though almost half of the words are not found in this database, potentially suggesting a focus on a smaller set of verbs. In terms of AWL presence, 752 instances (5.4%) exhibit presence from the AWL, while the remaining 13,055 instances (94.6%) do not. This imbalance suggests that the majority of words in the dataset are not part of the AWL, reflecting the predominance of more general vocabulary. Regarding NBC-FL's presence, the variable indicates whether a word appears in the NBC-FL's lexicon. A substantial majority, 11,127 instances (80.5%), are marked as present in the NBC-FL's lexicon, while 2,680 instances (19.5%) are

absent. This suggests that a large proportion of the words are recognized in the NBC-FL's lexicon, while a smaller subset is not, possibly corresponding to more specialized or less frequently used terms.

To compute a composite measure of lexical similarity between two words from different languages, we propose a normalized weighted formula that integrates the Levenshtein distance, and the similarity score provided by Python's *SequenceMatcher*. The Levenshtein distance quantifies the minimum number of single-character edits required to transform one word into another. To ensure comparability across strings of varying lengths, this distance is normalized by dividing it by the maximum length of the two strings, yielding a value between 0 (identical strings) and 1 (maximally different). The *SequenceMatcher* ratio, which provides a similarity score between 0 and 1 based on the longest contiguous matching subsequence, is converted into a distance metric by subtracting it from 1. The final *Weighted Similarity Score* (WSS) is computed as a weighted sum of these two normalized components:

$$\text{similarityScore}(a_{\text{italian stem}}, b_{\text{english stem}}) = \alpha \cdot (1 - \text{NLD}) + \beta \cdot (1 - \text{ratio}_{\text{seq}}(a, b))$$

Formula 11 This formula calculates the similarity score between two stemmed words, one in Italian and one in English, using a combination of Levenshtein distance (Formula 9) and sequence alignment ratio (Formula 10). The Levenshtein distance measures orthographic similarity, while the sequence alignment ratio captures structural similarity. The two components are weighted by parameters α and β , allowing flexibility in emphasizing one measure over the other. This score provides a comprehensive assessment of word similarity based on both character-level and sequence-level comparisons

This composite score provides a balanced and interpretable measure of string dissimilarity, leveraging both character-level edit distance and sequence-based alignment. In our implementation, both weights were set equally ($\alpha = \beta = 0.5$) to give balanced importance to both orthographic and structural similarity components. The resulting values are then combined using the weights α and β , which satisfy the condition $\alpha + \beta = 1$. These weights offer flexibility in modulating the influence of orthographic (edit-based) versus structural (sequence-based) similarity, depending on the linguistic or computational goals of the task. Unlike traditional edit distance metrics, which focus on transformations, the *SequenceMatcher* ratio emphasizes common structure and alignment, making it particularly suited for tasks such as approximate string matching, fuzzy search, and translation alignment in natural language processing.

In the process of merging the NVdB data with the English lexicon, specific attention was given to the handling of duplicated entries. Given the cross-linguistic nature of the database, the occurrence of

apparent duplicates was not uncommon. For example, in English, a single term like event can correspond to multiple Italian words, such as *evento*, *avvenimento*, and *manifestazione*, each with subtle differences in usage and connotation. In such cases, removing duplicates could result in the loss of valuable distinctions between these terms, which, while sharing the same English translation, are contextually different. This decision was informed by the understanding that lexical variations often reflect important semantic and pragmatic differences in language use. By retaining all entries, the database was better suited to support nuanced computational analyses, including semantic field studies, frequency analysis, and cross-linguistic comparison. It also ensured that the pedagogical purpose of the NVdB lexicon was preserved, allowing learners to access and understand the full range of meanings and contexts in which a word might be used, without the risk of oversimplification. For instance, consider the following entries in the database:

ID	578	934	4489
ITA	even to	manifestazione	avvenimento
ENG	even t	event	event
NVDB_LEVEL	1	1	3
LEVEL_NBC-FL	2	2	2
LEVEL_NBC-FL_NORMALIZED	1	1	1
STEM_ITA	even t	manifest	avven
STEM_ENG	even t	event	event
STEM_SIMILARITY	1	0.308	0.6
LEVENSHTEIN	0	6	3
LEVENSHTEIN_NORMALIZED	0	0.75	0.6
COMBINED_SCORE	1	0.2788	0.5
NVDB_PRESENT	1	1	1
NBC-FL_PRESENT	1	1	1
AWL_PRESENT	0	0	0

Table 12 This table illustrates a multilingual alignment dataset consisting of Italian terms and their English equivalents, along with computed similarity measures and metadata flags. The variables include both lexical similarity metrics (e.g., levenshtein_normalized, stem_similarity) and database-specific attributes (e.g., presence

in vocabulary databases and academic word lists). This combination of morphological, orthographic, and lexical information supports the evaluation of semantic alignment in bilingual terminology extraction and machine translation contexts

In Table 12, while all three Italian terms are mapped to the English word *event*, each Italian lemma represents a distinct nuance or usage in context. *Evento* is the most common, *avvenimento* may be used in more formal or historical contexts, and *manifestazione* might refer to a specific type of event, such as a demonstration or manifestation. Therefore, retaining these entries as separate records allows for a more detailed and accurate representation of the Italian lexicon, even though they share the same English translation. This lexical diversity underscores the importance of retaining these terms as distinct entries within the dataset. Despite sharing a common English equivalent, their unique contextual usages and degrees of formality justify separate records. This approach supports more nuanced cross-linguistic analyses and helps preserve the semantic richness of the source language, which would otherwise be flattened in a one-to-one translation model.

The table was defined with key attributes, including a unique identifier for each lemma (*id*) to ensure distinct referencing, the Italian form of the lemma (*ita*) as the target language, and the English equivalent (*eng*) to provide a cross-linguistic comparison. Additionally, each lemma was categorized by its NVdB level (*NVdB_level*), denoting whether it belonged to the *vocabolario fondamentale*, *vocabolario di alto uso*, or *vocabolario di alta disponibilità*. The corresponding frequency band from NBC-FL's model (*level_NBC-FL*) was also recorded, reflecting the lemma's placement within one of English ten frequency bands. Furthermore, each lemma was represented by its root form in both languages (*stem_ita* and *stem_eng*), allowing for cross-linguistic morphological analysis. These stems were obtained using the Natural Language Toolkit (NLTK) (Bird et al., 2009b) in Python, specifically employing the Porter stemming algorithm (Sparck Jones & Willett, 1997) for English (Bird et al., 2009). To quantify the morphological similarity between the Italian and English stems, a calculated value (*stem_similarity*) was included. The table also incorporated a numerical difficulty level (*levenshtein_distance*) to represent each word's pedagogical relevance, based on its frequency and complexity. This difficulty level was derived using the Levenshtein distance, which quantifies the minimal number of edits required to transform one string into another, providing a measure of lexical difficulty. Additionally, a composite metric (*combined_score*) was introduced, integrating factors such as frequency, relevance, and ease of acquisition, to rank lemmas according to their overall accessibility. Finally, binary indicators (*NVdB_present* and *NBC-FL_present*) were included to specify the presence of each lemma in the NVdB lexicon and NBC-, respectively. The precise nature and calculation of these scores are further presented in the article.

From this point onward, all lexical computations and comparisons presented in this article are based on stemmed forms of words rather than their surface forms. This approach ensures consistency in handling morphological variation across both languages and aligns with NBC-FL's word-family framework. By focusing on stems, we aim to capture underlying lexical concepts while minimizing the influence of inflectional and derivational variations. This is particularly crucial in cross-linguistic analyses, where equivalent lexical items may manifest in divergent surface forms due to typological differences. Consequently, all metrics, such as frequency alignment, morphological similarity, and pedagogical relevance, are derived from the stemmed representations, as outlined in the methods section.

With the NVdB data structured and accessible, a range of computational analyses became feasible. Lexical difficulty analysis was performed using the *levenshtein_distance* and *combined_score* fields, enabling the assessment of relative word difficulty for second language learners, which is essential for designing targeted instructional materials. Additionally, cross-linguistic studies were facilitated by the *stem_similarity* attribute, which allowed for the examination of morphologically related words between Italian and English, furthering research on language transfer and cognates.

In addition to preserving lexical duplicates, the database was deliberately designed to retain technical, domain-specific, and regionally marked Italian terms alongside more general vocabulary. This decision was grounded in the goal of maintaining the full semantic and lexical coverage represented across both source resources. Excluding such terms would have risked narrowing the scope of the index, thereby compromising its applicability in diverse communicative contexts. Given that the bilingual index is intended to serve as a comprehensive reference tool, ranging from general everyday usage to academic and specialized discourse, it was essential to include terminology reflective of varying registers and subject areas. By avoiding the filtering of entries based solely on frequency or generality, the resulting lexicon captures a broader spectrum of language use, ensuring its relevance for both pedagogical and domain-specific applications.

4.3.15.4 Index assessment

Our initial investigation focused on understanding how lexical similarity, as quantified by the Combined Score (a composite metric integrating Stem Similarity and Normalized Levenshtein Distance), influences the inclusion of terms in multilingual lexical databases. By disaggregating the analysis across three distinct groups, terms exclusive to the NBC-FL, terms found in both the NBC-FL and the NVdB, and terms exclusive to the NVdB, we uncovered a series of insightful patterns about structural alignment and editorial selection criteria.

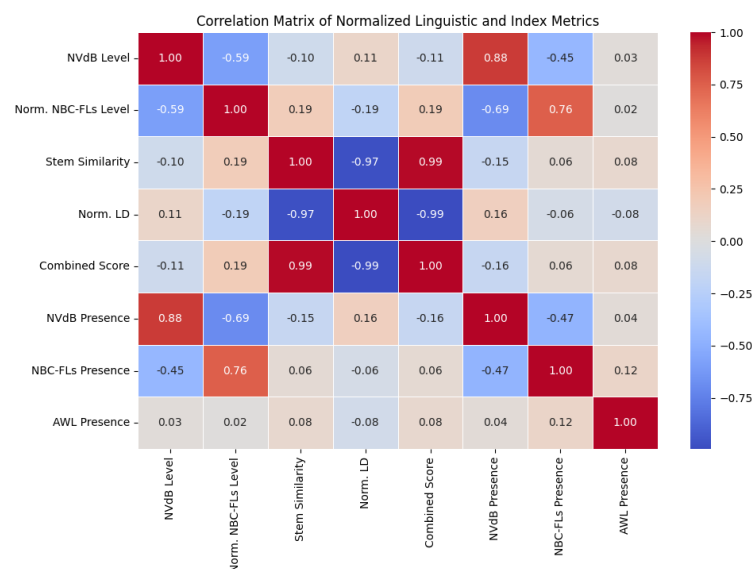


Figure 26 The correlation matrix illustrates the relationships between key numerical features in the dataset. Each cell represents the Pearson correlation coefficient between two variables. Values close to 1 indicate a strong positive relationship, values near -1 indicate a strong negative relationship, and values around 0 suggest little to no linear association.

The abbreviations used in the matrix are as follows: NVdB Level refers to the lexical level assigned within the NVdB; Norm. NBC-FL's Level denotes the normalized lexical level from the NBC-FL; Stem Similarity indicates morphological similarity between word stems; Norm. LD refers to the Normalized Levenshtein Distance, a measure of orthographic similarity; Combined Score is a composite similarity metric integrating both stem and edit-distance-based features; NVdB Presence is a binary indicator of inclusion in the NVdB; NBC-FL Presence represents inclusion in the NBC-FL; and AWL Presence denotes presence in the Academic Word List. The matrix provides insight into how lexical similarity metrics and resource inclusion indicators co-vary, highlighting patterns that contribute to the composition and differentiation of lexical resources

The correlation matrix (Figure 26) illustrates relationships among key numerical features in the dataset. Each cell in the matrix represents the correlation between two variables, with values approaching 1 indicating strong positive association, values near -1 denoting strong negative association, and values close to 0 suggesting minimal or no linear relationship. The matrix includes the following metrics: NVdB Level, Normalized NBC-FL Level, Stem Similarity, Normalized Levenshtein Distance (Norm. LD), Combined Score, and the binary presence indicators for the NVdB, NBC-FL and the AWL. Together, these features provide a multifaceted view of both linguistic structure and lexical inclusion.

As previously discussed, terms exclusive to the NBC-FL exhibited the highest average similarity ($\mu = 0.590, \sigma = 0.347$). This suggests that these lists, often compiled with communicative efficiency and linguistic standardization in mind, may preferentially include terms that exhibit high morphological

or orthographic coherence. Their content likely reflects widespread and regularly patterned lexical items that support general linguistic competency and accessibility.

In contrast, terms found in both the NVdB and NBC-FL displayed a lower average similarity ($\mu = 0.474$, $\sigma = 0.338$), potentially reflecting the compromises involved in reconciling different inclusion criteria across heterogeneous lexical resources. These overlapping entries may prioritize semantic breadth and contextual flexibility over surface-form regularity, indicating a more conceptual alignment than a purely structural one. Terms exclusive to the NVdB demonstrated an intermediate average similarity ($\mu = 0.489$, $\sigma = 0.333$), suggesting a balanced emphasis on both morphological consistency and domain-specific relevance. Given the NVdB role as a foundational lexicon for Italian, often used in pedagogical and language planning contexts, this balance likely reflects a carefully curated selection process that accounts for both frequency and conceptual representativeness.

Importantly, the strong negative correlation between Stem Similarity and Normalized Levenshtein Distance ($r = -0.974$) confirms the expected inverse relationship between morphological proximity and orthographic difference. While stem similarity captures deeper structural parallels, editing distance emphasizes surface variation, validating their complementary integration within the Combined Score. The Combined Score itself shows near-perfect positive correlation with Stem Similarity ($r = 0.991$) and a nearly equivalent negative correlation with Norm. LD ($r = -0.994$), affirming its robustness as a synthetic measure of lexical similarity. However, its explanatory power with respect to resource inclusion reveals more nuanced patterns.

Specifically, presence in the NVdB correlates negatively with the Combined Score ($r = -0.159$), implying that formal lexical similarity alone does not govern inclusion in this foundational vocabulary. This may reflect the NVdB deeper emphasis on conceptual salience, pragmatic value, or didactic necessity, dimensions that extend beyond string-level resemblance. In contrast, NBC-FL presence exhibits a weak positive correlation with the Combined Score ($r = 0.061$), suggesting a slight editorial preference for structurally coherent forms, consistent with the goal of promoting accessible and functionally central vocabulary within national language policy frameworks.

Finally, the moderately negative correlation between NVdB Presence and NBC-FL Presence ($r = -0.468$) underscores the partially divergent editorial philosophies of these two resources. While overlap exists, the NVdB emphasis on core pedagogical vocabulary and inclusion of specialized or culturally salient terms contrasts with the broader, frequency-driven approach of NBC-FL-specific lists. This divergence may be strategic: the NVdB seeks to offer a linguistically and semantically rich foundation for language users, whereas NBC-FL prioritize communicative generality and statistical centrality. The correlation patterns thus suggest a complementary but independently curated relationship between these resources, each tailored to serve distinct yet intersecting linguistic functions.

The analysis of the Combined Score provides valuable insights into the overall similarity of terms within the dataset. The mean value of the Combined Score ($M = 0.5322$) suggests that, on average, the terms exhibit a moderate degree of similarity. The relatively high standard deviation ($SD = 0.3456$) indicates substantial variability in the dataset, with some terms showing high similarity while others exhibit lower similarity. The minimum score of 0.0 and maximum of 1.0 further illustrate the wide range of term similarities, from negligible to perfect.

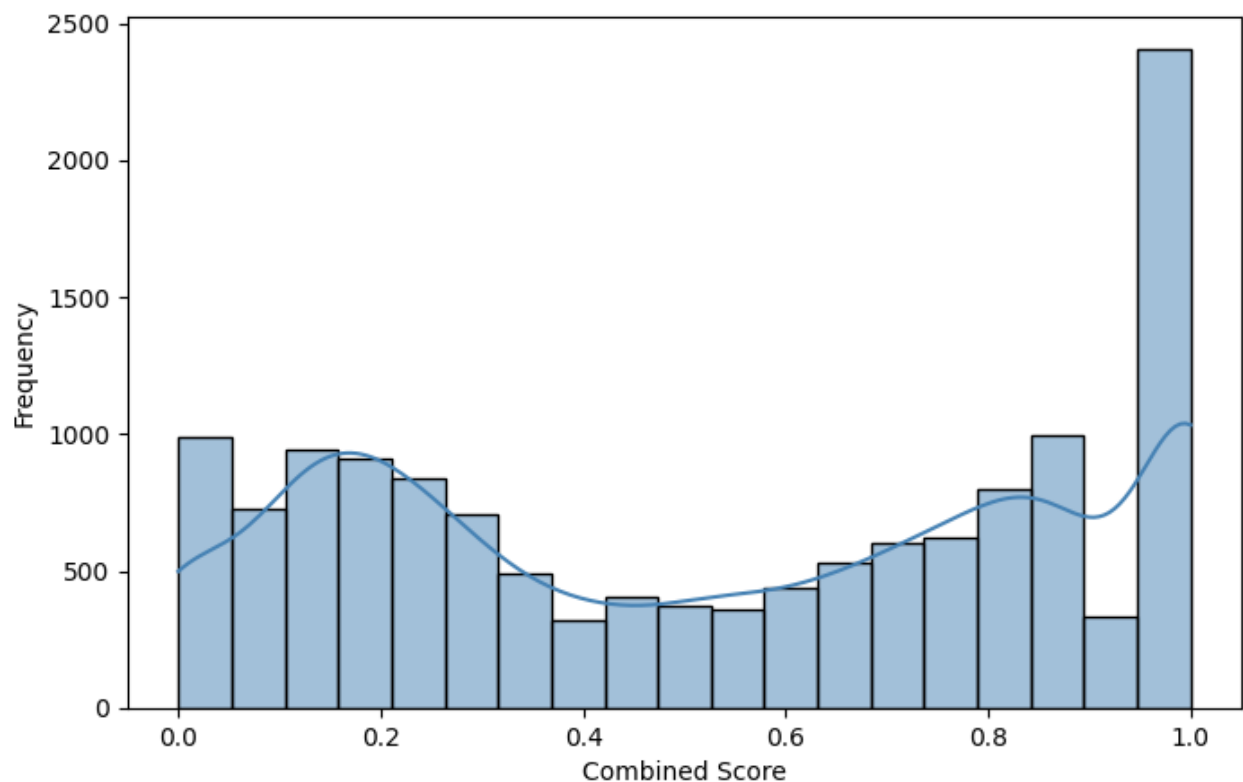


Figure 27 This histogram illustrates the distribution of the Combined Score across the dataset. The plot shows how frequently different values of the Combined Score occur, with the Kernel Density Estimate (KDE) curve overlaid to provide a smoothed representation of the data's distribution. The Combined Score is normalized and ranges from 0 to 1, with the histogram revealing the spread and concentration of scores. The plot suggests that most values of the Combined Score are clustered around the mid-range (0.5), with some higher and lower values forming the tails of the distribution

Figure 27 shows the histogram illustrating the distribution of the Combined Score across the dataset. The plot reveals how frequently different values of the Combined Score occur, with the Kernel Density Estimate (KDE) curve overlaid to provide a smoothed representation of the data's distribution. The Combined Score is normalized, ranging from 0 to 1, with the histogram showing the spread and concentration of scores. The data is grouped into 10 distinct ranges, with the frequency of occurrences as follows: the range 0.0 - 0.1 has 1532 occurrences, 0.1 - 0.2 has 1926 occurrences, and 0.2 - 0.3 has 1537 occurrences. The range 0.3 - 0.4 contains 797 occurrences, while 0.4 - 0.5 has 633

occurrences. In the higher ranges, 0.5 - 0.6 has 904 occurrences, 0.6 - 0.7 has 998 occurrences, and 0.7 - 0.8 has 945 occurrences. The range 0.8 - 0.9 has 1794 occurrences, with the highest frequency observed in the 0.9 - 1.0 range, which has 2741 occurrences. The plot suggests that most values of the Combined Score are concentrated in the higher ranges, particularly between 0.8 and 1.0, with the highest frequency occurring in the 0.9 - 1.0 range. Conversely, the lower ranges (0.0 - 0.2) have fewer occurrences, indicating that the scores tend to be more skewed towards the higher end of the scale. This pattern suggests a central tendency in the upper half of the score range, with fewer extreme low or high values, pointing to a relatively high concentration of scores near the maximum value.

The dataset under analysis reveals notable characteristics regarding the linguistic and computational measures, particularly focusing on Stem Similarity and Normalized Levenshtein Distance (Norm. LD), which serve as critical indices for evaluating the relationship between word forms in different languages or contexts.

The Stem Similarity variable displays a range from 0 to 1, with a mean value of approximately 0.57 and a standard deviation of 0.34. This suggests a moderate central tendency towards similarity between words stems across the dataset, although significant variability is present, as indicated by the relatively wide standard deviation. The 25th percentile (0.25) and 75th percentile (0.89) further illustrate that while most data points are concentrated around moderate similarity, the dataset includes both cases of minimal similarity (0) and near-perfect similarity (1). Additionally, the median value of 0.60 suggests that half of the observed instances exhibit a Stem Similarity of 0.60 or higher, which implies that a substantial portion of the words in the dataset share considerable lexical similarity.

Similarly, the Normalized Levenshtein Distance (Norm. LD) variable, which measures the dissimilarity between word pairs, spans from 0 to 1, with a mean of 0.50 and a standard deviation of 0.36. This indicates a balanced distribution, with data points equally spread across both high and low distances. The median value of 0.50 suggests that, akin to Stem Similarity, half of the word pairs have a dissimilarity close to the mid-range of the possible distance. The interquartile range (IQR) of Norm. LD spans from 0.17 to 0.83, highlighting the predominance of moderate distances but also revealing a subset of pairs with large dissimilarities. These results suggest that the dataset contains a diverse range of word pairs, from highly similar forms to those exhibiting considerable differences.

An examination of the Presence Counts by lexical source further enriches our understanding of the dataset. The variable for AWL Presence indicates that the majority of records (approximately 13055 of 13807) do not exhibit AWL presence, whereas only a small subset (752 records) does. This suggests that AWL words are less frequent in the dataset. In contrast, NBC-FL presence is a dominant feature, with 11127 records showing a NBC-FL presence and 2680 records not. This significant skew towards NBC-FL-based presence may have important implications for the contextual relevance of the words

in the dataset. Furthermore, the distribution of NVdB Presence, while also imbalanced, shows that 7227 records contain NVdB-related words, compared to 6580 that do not, suggesting that while NVdB presence is common, there is still a notable subset of records without it. When examining the Stem Similarity distribution by lexical group, we observe distinct differences between the records with NVdB Presence and those with NBC-FL presence. For records marked by NVdB Presence, the mean Stem Similarity is 0.52, with a standard deviation of 0.33. These values suggest that NVdB-related words exhibit moderate stem similarity, with significant variation around the mean. The distribution is skewed towards higher similarity, as evidenced by the interquartile range from 0.22 to 0.83. Conversely, records with NBC-FL presence show a slightly higher mean value of 0.58 and a similar standard deviation of 0.34. This indicates that words tied to NBC-FL-specific contexts tend to exhibit slightly greater similarity, with an interquartile range extending from 0.25 to 0.89, which overlaps with the NVdB group but with a shift towards higher similarity scores.

The Combined Score, which integrates various linguistic measures, presents an overall mean of 0.53 and a standard deviation of 0.35, signifying moderate performance across the dataset. The distribution of the Combined Score ranges from 0 to 1, with a median of 0.55, suggesting that most records fall within the mid-range. Notably, records without NVdB Presence exhibit a higher mean Combined Score of 0.59, compared to those with NVdB Presence, which have a mean of 0.48. This differential suggests that NVdB-related records tend to score lower on the combined metric, possibly due to differences in the lexical or contextual features associated with NVdB.

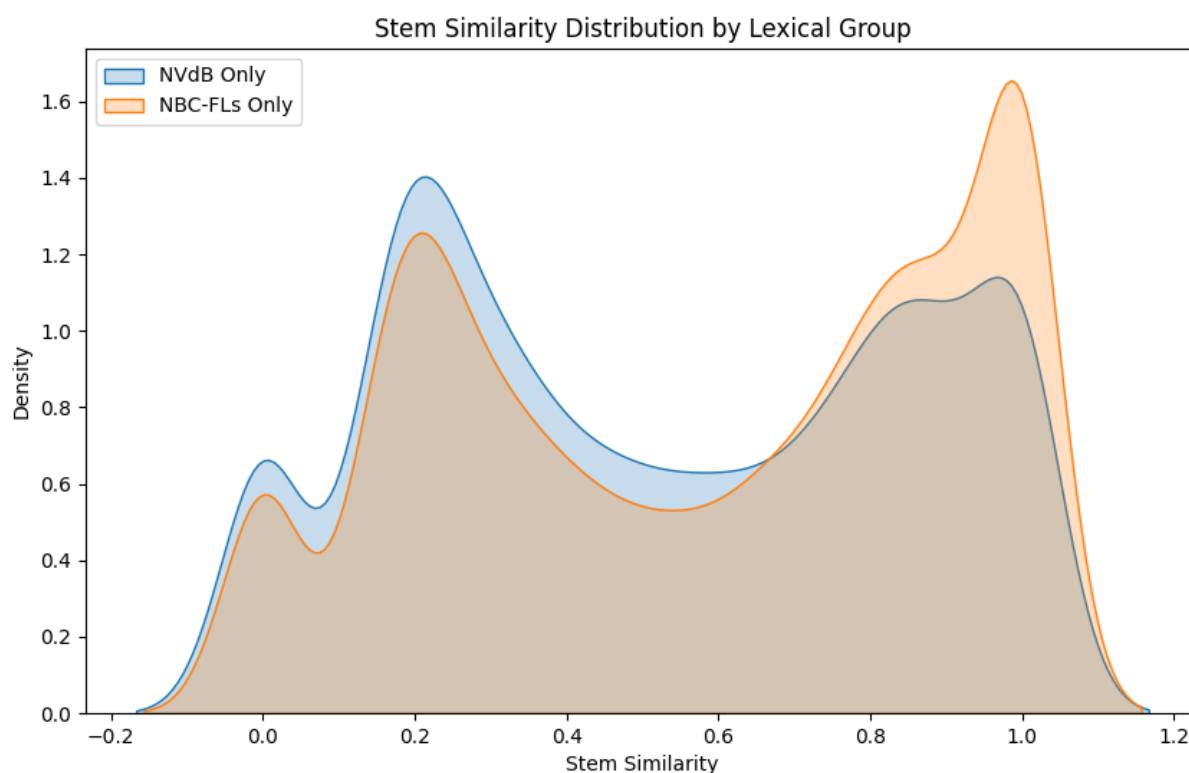


Figure 28 This plot presents the Kernel Density Estimate (KDE) of Stem Similarity for two subsets of the data: those with NVdB Presence (only) and those with NBC-FL presence (only). The KDE curves allow for the comparison of the distributions of Stem Similarity between these two groups, providing insights into how the presence of NVdB or Nation influences lexical similarity. The graph shows the smooth density of Stem Similarity for each group, with distinct peaks indicating the most common similarity values within each category. The plot may reveal that NBC-FL presence is associated with higher Stem Similarity on average compared to NVdB Presence, as indicated by the position and shape of the curves

Figure 28 provides a clear visualization of the distribution of Stem Similarity between two subsets of the dataset: those with NVdB Presence and those with NBC-FL presence. The Kernel Density Estimate (KDE) curves offer a smoothed representation of the distribution of Stem Similarity scores for each group, allowing for easy comparison of the frequency of different similarity values.

The KDE curves reveal that the distribution of Stem Similarity for terms with NBC-FL presence is shifted towards higher similarity values, with a distinct peak indicating that a significant portion of terms associated with NBC-FL-specific contexts share considerable lexical similarity. On the other hand, the NVdB Presence group displays a broader, more uniform distribution, with a less pronounced peak, suggesting a more diverse range of similarity values. This pattern indicates that terms in the NVdB tend to exhibit a wider variation in similarity, possibly reflecting the database's inclusion of a broader array of terms, some of which may not be as morphologically or orthographically similar.

The findings presented in Figure 28 support the earlier analysis of the dataset, where it was observed that terms tied to NBC-FL presence tend to exhibit slightly higher Stem Similarity on average compared to terms associated with the NVdB. This divergence may be attributed to the distinct selection criteria guiding the inclusion of terms in these two resources. NBC-FL presence likely prioritizes terms that are more structurally consistent and widely used, while NVdB Presence may incorporate a more varied set of terms, including those that are semantically or contextually significant, even if they are less similar in form.

The statistical analysis of Stem Similarity, Normalized Levenshtein Distance (Norm. LD), and the Combined Score highlights the complexity of lexical similarity within the dataset. The findings underscore the varying selection strategies employed by different lexical sources, such as the NVdB and NBC-FL presence. While terms tied to the NBC-FL presence tend to exhibit higher morphological similarity, terms within the NVdB showcase a more diverse range of similarity values, reflecting the database's broader focus on domain-specific or semantically rich vocabulary.

The use of both descriptive and inferential statistics, coupled with visualizations like Figure 26, Figure 27 and Figure 28 provides a comprehensive understanding of the interplay between string similarity and contextual relevance in the curation of lexical resources. This analysis emphasizes that lexical inclusion decisions are influenced by factors beyond simple morphological or orthographic similarity, pointing to a more nuanced approach in the compilation of language corpora. As a result, these insights are crucial for enhancing the development and refinement of specialized linguistic databases, especially in multilingual and domain-specific contexts.

4.4 Cognitive and Pedagogical Dimensions of Appil

The pedagogical strategies that underpin Appil build directly on the system's core features, translating them into meaningful learning experiences for users. At the heart of this design is the balance between overt and covert teaching processes, which allows learners to move naturally from explicit knowledge to intuitive fluency. In its overt dimension, Appil provides concise lessons where grammar, vocabulary, and phraseological units are introduced with clarity and transparency. These short, deductive sessions encourage learners to consciously engage with rules and structures, developing analytical awareness in line with Kahneman's distinction between controlled System 2 processing and automatic System 1 responses. Yet overt instruction alone cannot sustain fluency; for this reason, Appil complements explicit lessons with covert processes, in which the same items reappear in authentic, dynamically adapted texts. Here learners encounter patterns implicitly, guided not by direct explanation but by contextual repetition and subtle feedback. This two-phase cycle echoes Krashen's

distinction between learning and acquisition, ensuring that knowledge introduced overtly is later absorbed and consolidated through meaningful use.

A central principle in this process is noticing. Following Schmidt's Noticing Hypothesis and Swain's Output Hypothesis, Appil creates opportunities for learners to detect gaps in their knowledge and to test hypotheses about language use. Activities such as cloze tasks, sentence reformulations, or short production exercises invite learners to articulate their own language while receiving feedback that highlights errors indirectly, through reformulations or hints. In this way, feedback becomes a form of negotiation of meaning rather than a simple judgment of correctness. The process encourages learners to reflect on their interlanguage, refine their hypotheses, and gradually build more accurate and flexible representations of the target language.

At the same time, Appil pays close attention to cognitive load. Research by Nation (2014) and others suggests that the proportion of unknown items in a text should not exceed a minimal threshold if learners are to remain engaged rather than overwhelmed. For this reason, Appil uses advanced text processing to monitor lexical density, syntactic depth, and morphological complexity, ensuring that new material remains at an optimal level of difficulty. This design reflects cognitive load theory: by controlling extraneous demands, the system allows learners to focus their mental resources on processing and integrating new linguistic information. The gradual recycling of key vocabulary and structures across different contexts further strengthens memory traces, fostering long-term retention. Vocabulary learning in Appil does not take the form of isolated word lists but instead draws on the lexical approach, which emphasizes collocations, idioms, and phraseological patterns as the building blocks of fluent speech. By embedding lexical items within authentic discourse, the system exposes learners to language as it is naturally used, encouraging them to internalize formulaic sequences alongside individual words. This approach not only broadens vocabulary breadth but also deepens lexical knowledge, enabling learners to develop more natural and idiomatic competence.

Personalization reinforces these strategies by ensuring that learners engage with content that is meaningful and motivating. Appil continuously maps learner profiles, tracking progress and recurring difficulties in order to adjust input dynamically. For beginners, this might involve glossed texts with simplified structures, while advanced learners encounter authentic materials that challenge them with nuanced syntax and subtle lexical distinctions. By linking progress to transparent CEFR descriptors, the system also makes learners aware of their development in internationally recognized terms, thereby increasing motivation and providing a clear sense of achievement.

The way feedback is structured further strengthens the pedagogical design. Rather than presenting responses as right or wrong, Appil provides reformulations, contextual clarifications, or alternative phrasings that reflect how negotiation of meaning takes place in real-life communication. In this

respect, the system echoes Long's Interaction Hypothesis, treating feedback not as an endpoint but as part of the communicative process itself. Learners are encouraged to reconsider their output, adapt it to context, and progressively refine their control of the target language.

Taken together, these strategies reveal how the technical features of Appil translate into concrete learning pathways. Adaptive text processing becomes a mechanism for regulating cognitive load; overt and covert phases create a bridge between explicit knowledge and implicit fluency; lexical emphasis expands beyond individual words into phraseological competence; and personalization ensures that the learner's progress remains meaningful and transparent. Grounded in theories of second language acquisition and cognitive science, Appil does more than deliver content: it orchestrates an evolving, responsive learning experience that accompanies the learner from early awareness to confident, autonomous language use.

4.5 Conclusion

In sum, the pedagogical strategies embedded in Appil show how theoretical insights from second language acquisition, cognitive psychology, and lexicography can be transformed into practical mechanisms that support real learning. By balancing overt and covert processes, regulating cognitive load, and personalizing input through dynamic adaptation, the system creates a flexible yet structured environment where learners can progress with confidence. These strategies form the backbone of the application, but they gain their true effectiveness only when supported by a coherent technical design. The following chapter therefore turns to the system's architecture and implementation, examining how Appil's modules, data flows, and adaptive algorithms operationalize these pedagogical principles in practice.

5 Technical Infrastructure and Architecture

5.1 Introduction

The success of any digitally supported language learning initiative depends not only on the soundness of its pedagogical framework but also on the robustness, scalability, and responsiveness of its technological infrastructure. In the context of this project, the application, Appil, was developed to address a dual set of technical and educational challenges: it needed to deliver accessible and user-friendly language learning materials to a diverse cohort of municipal employees with varying levels of digital literacy, while simultaneously supporting complex, adaptive, and linguistically intelligent functionalities such as real-time learner tracking, personalized feedback, and natural language processing (NLP).

Importantly, the development of Appil began in 2022, at a time when large language models (LLMs) and generative AI technologies were neither as widely available nor as accessible as they are today. As development progressed, the team found itself in the midst of a profound transformation in the field of language technology. While this rapid evolution has led to a proliferation of generative tools, Appil remained grounded in a different trajectory, one shaped by pedagogical theory, open-access goals, and the technical and material resources realistically available to the researcher. In this sense, the learning platform may appear modest, or even outdated, when viewed in contrast to today's AI-driven tools. However, it represents a deliberate and principled approach: building a free, academically valid platform for language learning, guided by instructional theory and realized through available competence and self-funded infrastructure.

The core objective of the system was to operationalize a dynamic, pedagogically informed environment where users could engage with meaningful language input while receiving automated, adaptive support. Achieving this required a deliberate and iterative development process, one that evolved from rudimentary static implementations to a fully modular and extensible platform capable of integrating with advanced AI-driven components.

Through this technological trajectory, Appil exemplifies how language learning theory and computational infrastructure can be systematically fused to produce an effective and adaptable instructional tool, even under non-ideal and resource-constrained conditions.

All references to the source code, implementation details, and the final version of the Appil software correspond to the officially archived release available on Zenodo. All references to the codebase and final software version of Appil point to the archived release on Zenodo and Github (Pasquale Esposito, 2025, <https://doi.org/10.5281/zenodo.17490980>).

5.2 Early Steps: From Static Web to Scalable Vision

The development process began with modest steps, initially relying on a static web environment built on HTML, JavaScript, and CSS. This provided the necessary foundation for establishing the tool's basic functionality and for experimenting with the interaction flow. However, as the pedagogical vision expanded, it became clear that a static environment was insufficient. Its limitations in interactivity, scalability, and modularity made it incompatible with the long-term objectives of the project, particularly regarding dynamic personalization and integration of adaptive features.

At this stage, PHP was briefly considered as a candidate framework. Despite its extensive use in web development and its maturity as a technology, PHP revealed significant shortcomings in handling complex, real-time interactivity and microservices integration. These constraints also hindered future prospects of scaling the system toward more advanced features, such as adaptive natural language processing (NLP) modules. For this reason, PHP was ultimately deemed unsuited to the project's forward-looking design principles.

5.3 Early Web Prototyping (HTML, CSS, JS)

5.3.1 Evaluating PHP: A Short-Lived Contender

The development process of Appil commenced with modest technical foundations, initially relying on a static web environment constructed using standard web technologies: HTML, CSS, and JavaScript. This preliminary setup enabled the implementation of fundamental user interactions and provided a testbed for evaluating interface workflows and learner engagement strategies. Static prototypes proved valuable for rapidly iterating on the pedagogical design and for showcasing the feasibility of a browser-based interface. However, as the educational vision matured, encompassing real-time feedback, learner-specific adaptation, and NLP-supported interactivity, it became increasingly evident that the static environment lacked the architectural depth required for long-term sustainability.

The limitations of the static stack were manifold. First, the absence of a backend layer restricted the system's ability to store, process, and personalize user data across sessions. Second, the lack of

modular components meant that features could not be developed independently or reused efficiently. Most crucially, static environments are inherently ill-suited for tasks involving computational linguistics or real-time data manipulation, both of which are essential for the project's focus on adaptive language learning.

At this stage, the development team explored the potential of integrating a dynamic backend using PHP. As a widely adopted server-side language with a mature ecosystem, PHP appeared to offer a low-barrier entry point into dynamic content rendering and session management. It facilitated server-side scripting, form processing, and integration with relational databases, which were essential for basic personalization features.

However, despite these advantages, PHP exhibited critical limitations when evaluated against the project's long-term computational needs, particularly in the domain of advanced natural language processing. PHP lacks native support for high-performance scientific computing and offers minimal compatibility with modern machine learning libraries, which are typically developed in Python or Java. Managing tasks such as named entity recognition, syntactic parsing, or word-sense disambiguation would have required cumbersome workarounds or inefficient external service calls, leading to significant latency and complexity. PHP also falls short in supporting asynchronous processing and parallelism, both of which are essential when handling computationally intensive NLP tasks for multiple users in real time.

Moreover, PHP's monolithic design and limited support for microservices architectures posed challenges to modularization. The inability to isolate language processing logic from user interface logic, for example, would have hindered future efforts to scale or refactor the system as more sophisticated functionalities were introduced. These technical constraints ultimately rendered PHP unsuitable for the vision of Appil as an extensible, linguistically intelligent system.

Recognizing these limitations led to a decisive architectural pivot toward a more robust and future-proof framework, which will be described in detail in the following section.

5.3.2 Recognizing the Need for a Scalable Backend

The transition toward Java and Spring Boot represented a decisive turning point. Java, with its strong ecosystem and performance reliability, coupled with Spring Boot's ability to streamline microservices development, offered the modularity, scalability, and robustness required. This combination enabled the project to move beyond static content delivery and toward a dynamic architecture that could support real-time adaptation, continuous learner tracking, and advanced text analysis. The adoption of a microservices architecture further supported modularization, allowing individual components of Appil, such as text parsing, learner profiling, or feedback management, to be developed, deployed,

and updated independently. This flexibility ensured that the application could evolve alongside emerging pedagogical needs and technological innovations.

Python was integrated into the system to handle computationally intensive tasks, particularly those involving natural language processing. Leveraging its extensive libraries in machine learning and linguistic computation, Python became central to operations such as disambiguation, named entity recognition, and lexical profiling. A RESTful API mediated communication between Python-based modules and the Java/Spring Boot backend, ensuring smooth interoperability and preventing performance bottlenecks. This hybrid approach allowed Appil to benefit from the strengths of both ecosystems: Java's stability and scalability in backend management, and Python's agility in handling computational linguistics.

A robust and carefully structured database underpinned the entire system. The database was tasked with managing multiple streams of data: linguistic (grammatical structures, vocabulary entries), textual (user-generated input, processed materials), and infrastructural (user profiles, progress logs, interaction histories). Efficiency in data storage and retrieval was prioritized, as responsiveness is critical for maintaining learner engagement. The design ensured not only speed and stability but also the capacity to accommodate increasingly complex user interactions and large-scale deployments in the future.

The resulting application functions as a sophisticated web-based learning environment that operationalizes the didactic framework outlined earlier. Upon login, learners access a user interface designed with clarity and usability in mind, granting them entry into a suite of tools aligned with the principles of overt and covert instruction. The interface integrates features such as vocabulary tracking, grammatical analysis, adaptive reading modules, and feedback dashboards, while maintaining an intuitive navigation structure. Attention was paid to cross-platform compatibility, ensuring a consistent user experience across devices and browsers.

From the outset, development followed a user-centered design philosophy. Feedback was solicited from educators and learners at multiple stages, allowing iterative refinements to the interface, workflows, and feature set. This iterative cycle ensured that the application evolved in line with real-world classroom needs, bridging the gap between theoretical design and practical deployment. Comparisons with existing CALL and ICALL tools guided design decisions, helping to identify clear advantages such as Appil's stronger focus on adaptive lexical organization and its balanced integration of overt and covert instructional features.

Performance testing was also central to the development process. Stress tests and performance benchmarks ensured that the system could handle high user loads and computationally demanding

tasks without degradation. These measures reinforced Appil's suitability for both small-scale classroom integration and larger institutional deployments.

The technical challenges faced during development underscored the complexity of translating didactic innovation into operational reality. For instance, integrating Python-based NLP algorithms with the Java backend initially led to compatibility issues in data handling and processing speeds. The RESTful API solution, however, not only resolved this challenge but also created a flexible foundation for future expansions, including the potential incorporation of more sophisticated AI-driven modules. Similarly, ensuring a seamless user experience across devices demanded extensive testing and refinement, ultimately leading to a robust and reliable interface.

User evaluation provided further validation. Pilot testing with educators and learners revealed valuable insights into navigation ease, feature usefulness, and instructional clarity. This feedback informed refinements to both interface design and backend logic, ensuring the learning platform remained aligned with learner needs. Key performance indicators, including engagement rates, task completion, and user satisfaction, were established to track ongoing effectiveness.

In practice, the system translates into concrete educational workflows. For example, an educator preparing a reading comprehension task might upload a text to Appil, which then automatically processes it through named entity recognition and lexical profiling. The learning platform generates adaptive exercises tailored to learner proficiency, while students engage with the text and receive real-time feedback on vocabulary use, grammar, and comprehension. In this way, Appil not only operationalizes SLA theory but also empowers educators to create rich, responsive, and evidence-based learning experiences.

Looking ahead, the system is designed with scalability and futureproofing in mind. Its modular architecture ensures compatibility with emerging technologies, particularly AI-driven personalization tools. Planned enhancements include advanced adaptive learning algorithms, analytics dashboards for educators, and deeper integration with institutional learning management systems. The project's long-term vision is to extend Appil into a platform that not only supports individual learners but also provides institutions with actionable insights into learning progress, instructional effectiveness, and curriculum design.

Ultimately, the development of Appil represents more than a technical achievement: it is the materialization of a pedagogical vision into an operational system. By merging linguistic theory, cognitive principles, and advanced programming frameworks, Appil demonstrates how educational technology can be both scientifically grounded and technically robust. The challenges encountered and solutions devised throughout this process testify to the complexity of bridging theory and

practice. Yet, they also highlight the opportunities for creating innovative, scalable tools capable of reshaping language education in the digital age.

5.3.3 Initial Web Deployment via JSP and Servlet Architecture

In June 2024, the first operational version of Appil was deployed online via a dedicated web server. This version was built using JavaServer Pages (JSP) and Servlets, marking the project's transition from local prototyping to a publicly accessible, server-side application. The JSP-based model allowed for basic dynamic content rendering and form handling, enabling users to input text and receive rule-based feedback. While largely hard-coded and lacking modularity, this early deployment provided crucial insights into real-world usability and performance under network conditions. It also served as a proof of concept that confirmed the feasibility of deploying language-learning tools in a web-native environment.

However, the JSP/Servlet architecture quickly revealed its limitations. As the application began to incorporate more linguistically complex features, such as automatic feedback on user-generated texts and adaptive vocabulary scaffolding, the monolithic servlet model proved insufficient. Its limited support for asynchronous processing, modular componentization, and integration with modern computational libraries made it ill-suited for further expansion.

5.4 Architecture and Technology Stack

5.4.1 Microservices Architecture: Modularity and Scalability

The subsequent architectural transition to Java with Spring Boot marked a critical turning point in the development of Appil. Java's performance reliability and extensive ecosystem, combined with Spring Boot's support for microservices and dependency injection, enabled the creation of a modular and maintainable backend system. This shift allowed the application to decouple its core components, such as learner management, text processing, feedback generation, and analytics, into independent services that could be developed, deployed, and scaled individually.

Spring Boot also facilitated the creation of RESTful endpoints, which would later become essential for integrating Python-based NLP services. Furthermore, Spring's security modules provided the

necessary infrastructure for managing user authentication and authorization, ensuring compliance with institutional data privacy standards.

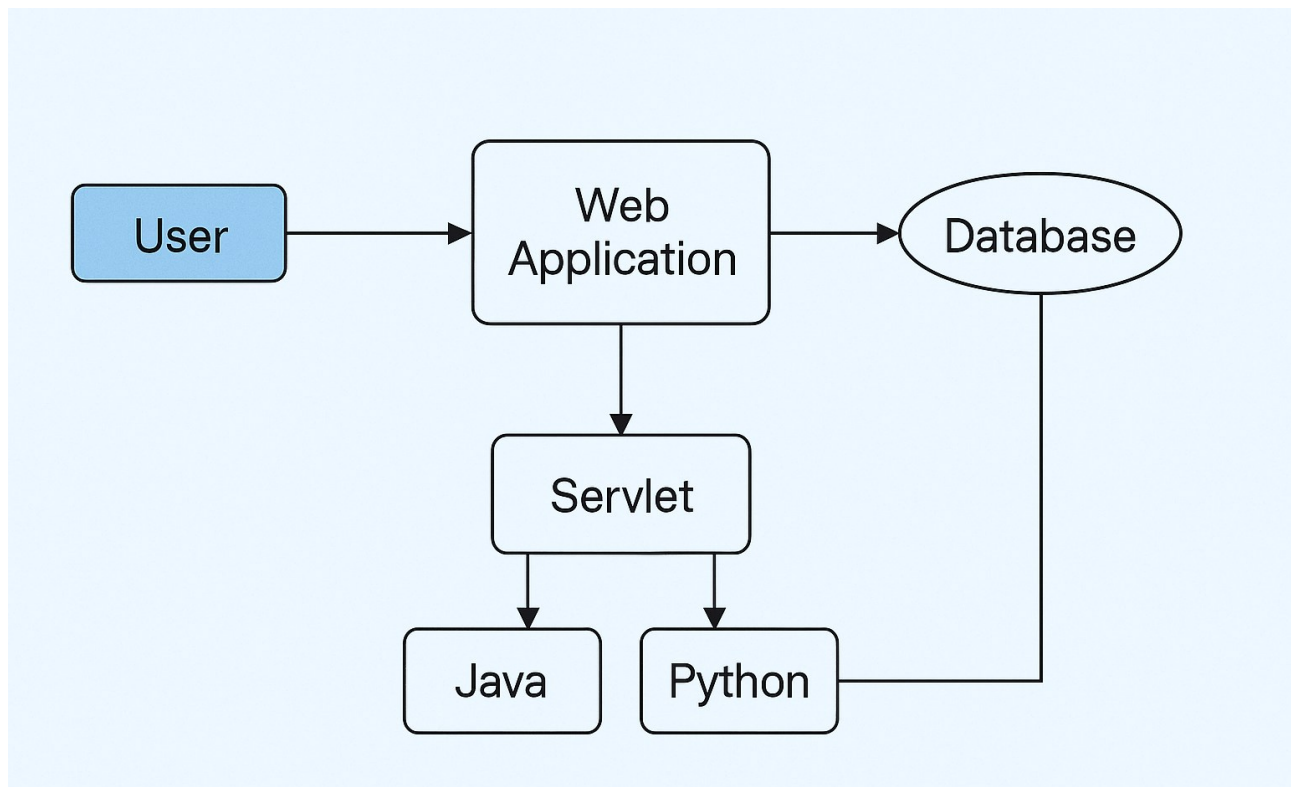


Figure 2 Application essential workflow

The subsequent architectural transition to Java with Spring Boot marked a critical turning point in the development of Appli. Java’s performance reliability and extensive ecosystem, combined with Spring Boot’s support for microservices and dependency injection, enabled the creation of a modular and maintainable backend system. This shift allowed the application to decouple its core components, such as learner management, text processing, feedback generation, and analytics, into independent services that could be developed, deployed, and scaled individually.

Spring Boot also facilitated the creation of RESTful endpoints, which would later become essential for integrating Python-based NLP services. Furthermore, Spring’s security modules provided the necessary infrastructure for managing user authentication and authorization, ensuring compliance with institutional data privacy standards.

5.4.2 Python Integration: NLP and Computational Linguistics

To support linguistically intelligent features, Python was integrated into the system architecture as a dedicated processing layer. Its extensive ecosystem of natural language processing libraries, such as spaCy, NLTK, and scikit-learn (Bird et al., 2009; Jurafsky & Martin, 2023; Pedregosa et al., 2011; Python Software Foundation, 2022), made it particularly well suited for computationally intensive

linguistic analysis. Within this layer, Python handled a range of core tasks including named entity recognition, part-of-speech tagging, lexical profiling, readability estimation, syntactic disambiguation, and vocabulary-level classification.

Rather than embedding Python logic directly within the Java application, the system adopted a service-oriented architecture that encapsulated Python modules as independent microservices. Each module exposed its functionality through a RESTful interface, allowing the main Java backend to delegate specific linguistic computations without creating tight interdependencies. This separation of concerns not only improved maintainability and modularity but also allowed Python-based components to be scaled or optimized independently. Furthermore, it enabled the deployment of these services in specialized computational environments when required, ensuring that the linguistic processing layer remained flexible, efficient, and responsive to evolving analytic and pedagogical needs.

The integration between the Java-based backend and the Python-based NLP modules was orchestrated through carefully designed RESTful APIs, which ensured smooth and asynchronous communication between components. For example, when a learner submitted a text, the Java backend would forward the content via an HTTP request to a Python service endpoint. The processed data, annotated for linguistic features and enriched with adaptive feedback, would then be returned as a structured JSON response.

This architecture not only facilitated interoperability between Java and Python but also enabled scalability by allowing each service to be containerized and deployed independently. The RESTful interface abstracts computational logic from the frontend, enabling future flexibility in replacing or upgrading individual services without disrupting the overall application.

5.4.3 Modular, Scalable, and Extensible Design

This chapter presents the system architecture underlying Appli, a second-language learning platform designed to balance didactic rigor with computational scalability. The learning platform adopts microservices based architecture in which discrete services, User Management, Linguistic Analysis, Instructional Logic, and Feedback Rendering, communicate via RESTful APIs. Services are loosely coupled and independently deployable, allowing the learning platform to evolve without disruptive redesigns. This architectural decision supports adaptive content delivery, near-real-time feedback, and smooth integration with external systems while preserving high availability under variable loads

A microservices approach was selected to satisfy four requirements central to Appli's mission. First, scalability: each component can be replicated horizontally according to observed demand, so additional NLP workers can be provisioned during peak activity without altering other subsystems.

Second, extensibility: new capabilities such as audio input or AI-driven feedback can be introduced as separate services behind stable interfaces, avoiding core re-engineering. Third, maintainability: faults and performance regressions remain confined to the affected service; recovery actions, rollback, hotfix, or targeted scaling, do not ripple through the platform. Finally, flexibility: services can be deployed on distinct servers or clouds in line with their computational profile and compliance needs, from on-premise research labs to public-cloud accelerators.

Appil is organized around clear domain boundaries. The User Management service handles authentication, authorization, learner profiles, and privacy controls. Its interfaces expose sign-up and sign-in flows, token refresh, and profile management through REST endpoints. Data are stored in a relational database with encryption at rest and minimal retention of personally identifiable information.

The Linguistic Analysis service ingests texts and produces structured annotations, tokenization, lemmatization, morphological tagging, syntactic parsing, lexical coverage estimation, and complexity metrics. It offers synchronous REST endpoints for small inputs and asynchronous job submission for corpora or batch workloads. Models and feature extractors are versioned to ensure reproducibility, and a feature store caches frequently requested analyses to reduce latency.

The Instructional Logic service transforms analyses and learner profiles into pedagogically coherent tasks. It selects and sequences activities, manages session state, and issues adaptive recommendations through RESTful interfaces. Its data layer combines graph-like relations among skills and tasks with transactional records of learner progress; rules and policies are versioned and auditable.

The Feedback Rendering service converts instructional decisions into user-facing feedback. It formats lexical suggestions, form-focus prompts, and textual enhancement highlights for web and mobile clients, and can stream progressive hints for longer exercises. Templates, localization bundles, and feedback logs support experimentation and A/B testing.

Inter-service communication follows an HTTP/REST style using JSON payloads. A gateway mediates client access, providing routing, authentication via OAuth2/JWT, rate limiting, and consistent observability. Public clients interact only with the gateway; east–west traffic remains within the protected network boundary. API versioning through URI prefixes (for example, /v1 and /v2) enables gradual evolution while maintaining backward compatibility. Operations subject to retries use idempotency keys, and long-running analyses are exposed as asynchronous jobs with token-based status queries. Internally, a lightweight message queue can be employed when burst smoothing or event fan-out is beneficial, but this does not affect public contracts.

Each service owns its private data store, preventing tight coupling through a shared schema and allowing independent scaling and migration. Transactional concerns such as identities and roles favor

a relational model, whereas document caches and analysis artefacts benefit from document or key-value stores. Schema changes are rolled out with forward-compatible migrations and blue-green deployment to minimize interruption. Caching is layered: static assets are served at the edge, immutable GET endpoints are cached at the gateway, and hot features are cached in-service. Data governance measures include pseudonymization of learner identifiers across analytical pipelines, retention schedules aligned with institutional policy, and comprehensive audit trails for access to sensitive records.

The architecture supports single-tenant institutional deployments on private infrastructure, multi-tenant cloud deployments for smaller cohorts, and hybrid models in which computing intensive NLP workers run on burstable public-cloud nodes while learners data remain on secure institutional servers. Containerization and orchestration allow horizontal scaling and rolling updates with minimal downtime. Environment configuration follows twelve-factor principles, enabling predictable promotion from development to production.

Scalability is achieved through stateless service replicas, autoscaling policies linked to latency, CPU, or queue depth, and separate scaling planes so that increasing the number of NLP workers does not impact the footprint of the user subsystem. Extensibility is realized by introducing new capabilities such as services or plugins behind existing interfaces, with feature flags and capability negotiation enabling progressive rollout and controlled experiments. Maintainability benefits from clear ownership boundaries, contract tests, and consumer-driven API testing; a degradation in feedback rendering does not affect authentication or profile retrieval. Flexibility arises from the ability to deploy services across heterogeneous environments and to adapt resource allocation to changing computational needs without code changes.

Performance tuning focuses on pre-warming model caches for linguistic analysis to avoid cold starts, batching high-volume requests, and streaming partial results for long-running jobs. Static assets for enhancement and feedback are rendered at the edge to shorten the critical path from request to interaction. Payloads are rightsized for constrained devices to reduce bandwidth demands without compromising pedagogical value.

Continuous integration and delivery pipelines automate building, testing, and security scanning. Releases use canary or blue-green strategies with automatic rollback based on service-level objectives. Configuration and secrets are externalized and stored securely with rotation policies.

Interoperability is addressed through REST connectors for learning platforms, import and export endpoints for learner data and results, and webhooks for event notifications such as activity completion or assessment availability. These integration points enable Appil to serve as a classroom

tool and, when required, to operate as a scalable platform for large-scale institutional deployment without sacrificing modularity or data governance.

The microservices design described here positions Appil as a robust and adaptable learning platform. By separating concerns across independently deployable services, using stable and well-documented interfaces, and enforcing strong operational discipline, the system can scale with demand, accommodate new instructional modalities, and maintain reliability even under diverse deployment conditions. This foundation supports both small-scale pilots and large institutional roll-outs while preserving the flexibility required for ongoing innovation in second-language pedagogy.

5.5 Database Design and Data Management

5.5.1 Data Streams and Structures

A critical foundation of Appil's technological infrastructure lies in its database architecture, which was designed to sustain the interplay between pedagogical goals, user interaction, and computational linguistics. As the platform expanded to deliver personalized feedback, track learner progress, and adapt content dynamically, data management had to be robust and efficient while remaining scalable and extensible. The solution adopts a modular, entity-driven relational schema capable of supporting multiple concurrent data streams without sacrificing transactional integrity or query performance.

The first stream concentrates on linguistic data. Lexical entries, grammatical annotations, frequency lists, and word-level metadata are normalized into machine-readable structures so that NLP modules can operate deterministically and reproducibly. Specific care is taken to preserve provenance, variant forms, and usage constraints, enabling analyses that distinguish lemmas from surface forms and register- or variety-specific realizations. The second stream organizes textual data, including learner-submitted productions, curated instructional materials, and the processed artefacts produced by the analysis pipelines. Documents are versioned and linked to their derived annotations, allowing audits of changes over time and facilitating reprocessing when models or rules evolve.

The third stream addresses user data. Profiles, learning histories, skill assessments, and user-generated feedback are captured with explicit temporal semantics to support longitudinal analysis and fine-grained personalization. Referential integrity across identities, sessions, and artefacts ensures that the system can reconstruct the pedagogical context for any interaction. The fourth stream captures interactional data. Fine-grained event logs record session activity, task completion metrics, and timestamps suitable for analytics and inference. These records are indexed for fast aggregation and aligned with privacy requirements through pseudonymization and controlled retention windows.

A central design commitment was the transformation of vocabularies for second-language learning into electronic objects amenable to computational processing for specific pedagogical purposes. Conventional wordlists and glossaries were converted into structured lexical resources with stable identifiers, typed relations, morphological and phonetic features, frequency and dispersion statistics, usage examples, and licensing metadata. This conversion enables programmatic tasks such as lexical coverage estimation, adaptive selection of target items, text enhancement aligned with learning objectives, and the generation of context-sensitive feedback. Data quality is maintained through validation rules, reproducible import pipelines, and versioning policies that make updates traceable and reversible.

Across all streams, the schema separates hot operational tables from analytical replicas to balance OLTP responsiveness with research-grade reporting. Caching and materialized views are used to accelerate recurrent queries, while write paths remain minimal and strongly consistent. By combining clear entity boundaries, explicit temporal modelling, and machine-processable lexical resources, the database architecture provides a durable substrate on which adaptive instruction, rigorous analytics, and future extensions can reliably operate.

5.5.2 Efficiency in Storage and Retrieval

The relational schema is deliberately normalized (3NF with selective BCNF in high-churn areas) to minimize redundancy while preserving predictable, low-latency joins across user, task, submission, and analytics domains; core vocabulary lives in a central lexicon table keyed by stable lemmas and language codes, with morphological variants and frequency statistics in child tables, while learner-specific familiarity is captured as a many-to-many *learner_vocab_familiarity* relation that records user ID, lexeme ID, familiarity level, evidence source, and a precise event timestamp so longitudinal progress can be reconstructed without duplicating lexicon rows; writes are append-only for auditability and to support temporal queries (e.g., ‘familiarity as of t’), and surrogate keys (UUIDv7) are used to improve locality and sortability; soft deletes and validity ranges (*valid_from*, *valid_to*) permit schema evolution without losing history. To avoid recomputation in downstream services, each user submission has a companion *text_artifacts* record that stores canonicalized text plus a layered bundle of NLP outputs: tokenized text with character offsets, sentence boundaries, POS tags, dependency parses, named entities with spans and confidence scores, readability indices with method provenance, grammar/style matches with rule IDs and suggested fixes, embeddings for semantic similarity, and detector flags (e.g., L2 interference patterns); artifacts are versioned (*artifact_schema_version*, *nlp_model_version*) so pipelines can run multiple model generations in parallel, and a slim *artifact_index* table exposes common filter keys (language, domain, length bucket,

processing status) to keep hot queries on narrow indexes; large or binary payloads (parse forests, attention maps) are stored outside the relational core in an object store with content hashes, referenced via opaque URIs, while the database holds normalized metadata and cross-references, enabling cheap joins for most reads and streaming access for heavyweight assets; this hybrid approach scales horizontally (more users and texts over time) because the core tables remain narrow and index-friendly, and vertically (richer annotations) because new artifact layers can be added by incrementing the artifact version and inserting extension rows rather than altering existing contracts; finally, strict data contracts (column types, nullability, enum domains) are enforced at the database level, with migration scripts that are backward compatible and include pre-deployment checks (row counts, constraint validation) to keep continuous delivery safe.

5.5.3 Performance Considerations and Query Optimization

Because learners anticipate immediate feedback following the submission of written tasks, the system's architecture was optimized for sub-second responsiveness under high concurrency conditions. Latency targets at the ninety-fifth and ninety-ninth percentiles were treated as key service-level objectives, achieved through a multi-layered optimization strategy that spans physical data design, runtime control, and continuous observability. At the data layer, indexing and partitioning strategies were employed to reduce retrieval latency and ensure bounded index growth. Frequently filtered columns were indexed individually, while compound and partial indexes supported time-based queries and selective retrieval of active data. For very large datasets, temporal partitioning by month constrained index size and facilitated efficient retention management. Query plans were continuously validated through explain-based analysis integrated into the testing pipeline, ensuring that execution costs and buffer utilization remained within thresholds. Periodic maintenance routines and tuned autovacuum parameters preserved stable performance statistics over time.

Higher in the stack, read-intensive operations were optimized through precomputation and caching. Materialized views were introduced for stable aggregation patterns such as recent learner artifacts or cumulative mastery indicators, refreshed incrementally to balance data freshness and computational cost. For critical read paths requiring low latency, a denormalized read model was selectively employed, retaining a normalized source of truth while minimizing multi-join fan-out. On the application side, performance improvements relied on a combination of caching, pooling, and asynchronous execution. In-memory caches stored slow-changing instructional assets with explicit time-to-live policies and version-aware keys to prevent stale retrievals. Negative caching mitigated request stampedes on missing entries, while background refresh routines ensured smooth cache turnover without imposing burst load.

Connection pooling was tuned to amortize handshake overhead while preserving stability during deployment or failover events. Pool configurations were designed with bounded concurrency, controlled lifetimes, and randomized eviction intervals to prevent synchronized reconnect surges. Computationally expensive natural language processing tasks were decoupled from the synchronous request cycle and delegated to background workers through durable queues. Each job was designed to be idempotent, allowing safe retry under transient failures. The processing pipeline employed exponential backoff and transactional outbox patterns to ensure exactly-once persistence of results. Intermediate outputs, such as tokenization and sentence segmentation, were stored incrementally so that dependent modules could begin processing prior to pipeline completion. User-facing responsiveness was supported through progress events that allowed the interface to display intermediate states rather than waiting for full computation. This design minimized cognitive uncertainty and maintained user engagement during processing delays. Observability practices were embedded throughout the system, combining structured logs, correlation identifiers, and quantitative metrics on request rate, error frequency, and duration. These signals fed continuously updated service-level dashboards and automated alerts, enabling proactive performance management. Regular load tests and fault-injection exercises verified that latency objectives were maintained even under degraded conditions such as node failure or cold-cache states. Through this layered optimization framework, the learning platform achieved high responsiveness and reliability while preserving the interpretability and pedagogical coherence of feedback during peak operational demand.

5.5.4 Futureproofing and Data Portability

To support integration, research, and forthcoming AI personalization at scale, the schema emphasizes interoperability, governance, and evolution: foreign-keyed boundaries and clear data ownership (user, content, analytics, ops) make it straightforward to expose stable views to LMS platforms and analytics dashboards, while export pipelines produce CSV and JSON Lines with embedded schema/version headers plus optional Parquet for columnar analytics; change-data-capture (CDC) streams mirror selected tables into a lakehouse for offline modeling, with lineage metadata (source table, transformation commit, artifact version) recorded to ensure reproducibility; event logs capture timestamped learner actions (opens, edits, submissions, feedback interactions) in an append-only structure with consistent event schemas, high-resolution timestamps, and device/context fields so behavioral sequences can be reconstructed for predictive models (e.g., next-best-exercise, error-type forecasting, disengagement risk), and a light feature registry enumerates derived features (rolling accuracy, time-on-task, lexical diversity) with owners and backfills to keep online/offline parity; schema evolution follows semver-like rules (additive first, deprecations flagged with sunsets),

migrations are blue/green with dual-write or write-redirect where necessary, and compatibility tests confirm that downstream consumers keep functioning across versions; finally, privacy and compliance are first-class: PII is separated into a minimal profile table with encrypted-at-rest columns and strict role-based access, analytics tables use pseudonymous user keys with reversible mapping guarded behind a data steward workflow, retention and erasure (GDPR/CCPA ‘right to be forgotten’) (Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation, 2016) are implemented via keyed tombstones and partition-prune deletes, and exports are scoped by purpose with audit trails; by combining modular schema boundaries, portable data formats, CDC to analytic stores, and rigorous governance, Appil sustains today’s instructional objectives while creating a durable runway for richer annotations, model-driven personalization, and cross-institutional interoperability without costly rework.

5.6 Interface Design and User Experience

The effectiveness of any educational technology depends not only on the strength of its backend architecture but also on the clarity, usability, and pedagogical alignment of its front-end interface. In Appil, interface design was a first-class concern from inception and was guided jointly by user-centered design practice and second language acquisition (SLA) theory. Because the interface is the learner’s primary contact with the system’s instructional affordances, it is intentionally intuitive, minimal, and goal-directed: core tasks are foregrounded, friction is reduced at every step, and the linguistic insights generated by the feedback engine are surfaced in a way that is understandable at a glance yet expandable on demand. The result is a UI that helps learners do the right thing by default, makes progress visible, and turns complex analyses into actionable suggestions without cognitive overload.

5.6.1 Clarity, Accessibility, and Cross-Platform Compatibility

Clarity begins with information architecture and language choices: labels are written in plain, learner-facing terms; task flows are linearized wherever feasible; and the same concepts appear with the same names in the same places across the product. Accessibility was designed in from the start rather than retrofitted. Interfaces comply with WCAG 2.1 AA (World Wide Web Consortium, 2018, 2023) through high-contrast modes, scalable typography with a readable base size and generous line height, keyboard navigability and visible focus states, ARIA roles and landmarks for assistive technologies,

consistent target sizes and spacing for touch input, and motion-reduced alternatives for users with vestibular sensitivities. Form components announce errors and success states with programmatic labels, not color alone, and all interactive elements can be reached and operated without a pointing device. For users with limited digital literacy, including municipal employees unfamiliar with e-learning, the visual system relies on a small, consistent set of patterns (tabs, accordions, steppers) and progressive disclosure: advanced controls remain available but are tucked behind explicit ‘More detail’ affordances, so novices stay oriented while experts move quickly.

Cross-platform compatibility is delivered through responsive, mobile-first design and a resilient layout system. Flexible grids, fluid containers, and component-level breakpoints ensure that the same activity can be completed seamlessly on a desktop during a training session, on a tablet in the field, or on a phone during a commute. Interaction idioms adapt to the device, hover reveals become tap-to-expand, keyboard shortcuts have touch equivalents, while preserving cognitive continuity across contexts. Performance budgets keep perceived latency low: above-the-fold content streams first, heavy assets defer until interaction, and long NLP results load incrementally so users can begin reading feedback before the full analysis arrives. Offline-tolerant behaviors (local draft storage, retry queues for submissions) protect work in unreliable network conditions, and internationalization/localization support (string catalogs, bidi layout safety, locale-aware formatting) lets institutions deploy Appil in multilingual settings without UI regressions.

5.6.2 Iterative Design Based on User Feedback

The interface is the product of continuous discovery rather than a single blueprint. Early stages used low-fidelity wireframes and clickable prototypes tested with both educators and learners to validate navigation, terminology, and mental models. Findings drove structural changes, for example, the original feedback panel was dense with technical terms; it was reworked into a tiered design that surfaces a short, learner-friendly summary first, then allows expansion into rule explanations, examples, and references for those who want depth. Subsequent rounds introduced mid- and high-fidelity prototypes instrumented for task success metrics (time on task, completion rate, error rate) and standardized measures such as SUS and SEQ to quantify perceived usability and ease of key flows.

As features matured, controlled experiments and behavioral analytics guided refinement. A/B tests evaluated alternative button placements, copy variants for prompts, and ordering of feedback categories; heatmaps and scroll-depth analyses identified content that was overlooked or overwhelming; and cohort analyses compared first-session activation and week-one retention for learners exposed to different onboarding sequences. Importantly, design decisions were checked

against SLA principles to ensure that improvements in clicks or time-to-complete did not increase cognitive load or reduce learning value. For instance, when a more aggressive highlight scheme raised click-through on vocabulary tips but increased abandonment of writing tasks, the team tuned salience and batching to preserve the ‘noticing’ effect without fragmenting attention. Feedback loops are institutionalized: in-product micro-surveys capture moment-of-use impressions, support tickets feed directly into a triage board that tracks severity and frequency, and periodic educator councils review roadmap items to align classroom practice with product affordances.

5.6.3 Usability Testing and Workflow Optimization

The UI deliberately operationalizes both explicit and implicit instructional strategies. Explicit instruction appears as visible, just-in-time feedback, grammar highlights with short rationales and examples, vocabulary annotations that connect to the central lexicon, tooltips that explain rules in plain language, and ‘why this suggestion?’ links that disclose rule provenance and model confidence where appropriate. Implicit instruction is embedded in the structure of tasks and sequencing: input enhancement (subtle emphasis on target forms), guided noticing through comparison examples, and practice activities that nudge retrieval and variation without heavy meta-language. The dashboard closes the loop by visualizing progress on dimensions that matter for learning, recent texts analyzed, distribution of feedback types, vocabulary exposure and spaced-repetition status, time-on-task streaks, so learners and educators can spot patterns quickly and decide what to do next.

Workflow optimization focuses on keeping cognitive load within an optimal band while maintaining a sense of control and momentum. Feedback arrives in coherent batches rather than a stream of interrupts; suggestions are grouped by pedagogical purpose (form, meaning, usage) and sorted by expected impact; and each suggestion offers three symmetric affordances, apply, learn more, or dismiss with reason, so the system can adapt future feedback to learner preferences. Color, iconography, and micro-interactions serve pedagogy rather than decoration: unknown vocabulary is consistently emphasized with a single hue and a predictable icon; lexical repetition flags include immediate alternatives drawn from the learner’s known-word zone to avoid gratuitous difficulty spikes; collapsible panels remember user state between sessions; and lightweight animations signal cause-and-effect without stealing focus. When errors occur (time-outs, missing artifacts), recovery paths are explicit and forgiving, drafts are never lost, and the UI explains what will happen next in clear, non-technical language.

End-to-end usability is supported by instrumentation and guardrails. Latency SLOs are set for key flows (submitting a text, opening feedback, applying suggestions); if the system anticipates a delay, the UI switches to progressive states (skeleton screens, percentage-free spinners with contextual ETA

hints) and offers an immediate path to other productive tasks (reviewing prior feedback, practicing flagged vocabulary). Personalization artifacts are transparent: learners can see why a recommendation was made (*based on repeated article-noun errors last week*) and can adjust difficulty, pacing, or feedback density from a single preferences panel, balancing autonomy with guidance. Educators get aggregated views that mirror learner dashboards at class scale, with privacy-aware summaries and export options aligned to institutional reporting needs.

5.7 Pedagogical Alignment and Feature Integration

A defining feature of Appil is its ability to turn second language acquisition (SLA) theory into a responsive, interactive, and data-driven learning experience. The application is not a passive container for digital content but a pedagogically intelligent system in which interface patterns, feedback logic, and data models are explicitly aligned with instructional principles. This alignment is realized through interlocking features that jointly support explicit instruction (clear, just-in-time explanations), implicit acquisition (structured exposure and guided noticing), adaptivity (content and feedback calibrated to evolving proficiency), and metalinguistic awareness (tools that help learners articulate what they are doing and why). Every technical component, NLP pipelines, recommendation services, dashboards, and micro-interactions, serves an instructional purpose, so that the system not only delivers content but also mediates a rich set of learning interactions tuned to each learner's needs.

5.7.1 Mapping SLA Theory to Application Features

The learning platform operates key SLA tenets in concrete, repeatable ways. Noticing (Schmidt, 1990) is supported through calibrated salience: the UI draws attention to patterns and potential errors using a limited palette of visual cues, short rationales in plain language, and optional 'expand for detail' affordances; highlights are batched to avoid attentional scatter, and their density adapts to task phase (drafting vs. revising) so that noticing increases when it is most useful. Input enhancement is applied to authentic texts via subtle emphasis on target forms (e.g., articles with count nouns, verb-tense/aspect pairs, collocations), with intensity adjusted by learner level and prior exposure; the enhancement is informative rather than decorative, and learners can toggle it off to reduce cognitive load. Form-focused instruction appears inside meaningful communicative tasks: grammar guidance is anchored to the learner's own production or to the discourse of assigned readings; rules include minimal terminology, canonical examples, and contrasts with near-miss forms, and each suggestion links to short practice items that recycle the same structure in varied contexts. Interactionist approaches are realized through input-output-feedback loops: learners read or listen, produce

language in low-stakes prompts, and receive contingent feedback; the loop is tightened by micro-tasks (e.g., reformulate a sentence, choose between two plausible forms, rank alternatives by register) that require uptake rather than passive acknowledgment. Concretely, when a learner uploads a text, NLP pipelines segment, tag, and diagnose common L2 challenges (word order, tense/aspect, article choice, prepositions, agreement). The feedback layer then scaffolds engagement: each detection is framed as a hypothesis ('Check article choice before singular count noun'), supported by a brief explanation and one or two contrastive examples drawn from the learner's recent work or level-appropriate corpora; the learner can accept, revise with guidance, or dismiss with a reason, and these actions feed back into the profile so future feedback becomes more targeted. Importantly, corrections are never presented as end points; they are invitations to notice patterns, test alternatives, and consolidate emerging rules.

5.7.2 Adaptive Reading and Vocabulary Modules

The adaptive reading module adjusts texts and tasks to proficiency in real time. Lexical profiling estimates difficulty using frequency bands and CEFR-aligned lists; the system computes coverage and dispersion, identifies spikes in low-frequency items, and proposes supports (glosses, paraphrases, or alternative sentences) when predicted comprehension falls below a threshold. Named entity recognition isolates domain concepts and proper nouns to reduce spurious difficulty: entities receive concise, context-aware tooltips or image cues, and repeated entities collapse into summary chips so that support remains present without cluttering the page. Cloze generation produces fill-in-the-blank items that target actual gaps: the system selects candidates based on part-of-speech patterns, collocational strength, and the learner's exposure history; difficulty scales by distractor quality (minimal pairs, morphological neighbors, register confounds), gap density, and whether answers come from the text or the learner's known-word zone. WordNet (Miller, 1995) integration ties reading to durable vocabulary growth: unknown items can be saved with a single click; each entry stores lemma, sense, example drawn from the learner's text, and pronunciation; review scheduling uses spaced-repetition with difficulty updates derived from response latency and self-ratings; mastery is not binary but probabilistic, so items resurface after decay or when new contexts reveal instability. Together, these mechanisms support incidental learning, encountering words in context, while scaffolding comprehension and fluency: learners see just enough support to keep moving, are nudged to notice recurring forms, and receive follow-up practice at the right interval and level.

5.7.3 Feedback Dashboards and Learner Profiling

Adaptive behavior rests on a robust, privacy-aware learner profile that aggregates interaction data into actionable signals. Profiles track vocabulary familiarity and usage (items seen, produced, and successfully retrieved; collocation strength and coverage across themes), grammar challenge zones (error types normalized by opportunities, recent regressions, and responsiveness to feedback), fluency measures (time on task, burst length, revision cycles, hesitation proxies in typed production), and topic engagement (which genres and prompts sustain attention and yield higher-quality output). These signals feed the recommendation layer and also drive transparent visualizations. For learners, the dashboard presents a concise, longitudinal view of progress, recent texts analyzed, top three stabilized patterns, top three emerging challenges, and vocabulary consolidation status, plus goal-setting tools that translate diagnostic findings into weekly, concrete targets (‘practice article–noun combinations with count nouns in descriptions,’ ‘consolidate these 12 work-related collocations’). For educators, class-level views summarize distribution of challenge types, vocabulary coverage by theme, and engagement patterns, with drilldowns to anonymized exemplars and exportable reports aligned to institutional requirements.

Feedback delivery is designed to inform and guide without overwhelming. Inline annotations surface at the point of need (definitions, collocation hints, grammar flags with minimal meta-language); summary boxes at the end of a task synthesize lexical density, readability, and challenge mix, compare the current submission with recent work, and propose one or two focused next steps; dynamic recommendations connect diagnostics to action by suggesting readings, micro-exercises, or revision prompts that recycle target forms in fresh contexts. Control stays with the learner: they can tune feedback density, choose simplified vs. detailed explanations, and defer non-critical suggestions to a ‘later’ queue that the system revisits during practice sessions. Every acceptance, revision, or dismissal updates the profile and closes the loop, so future interactions become better calibrated, fewer redundant prompts, more attention to patterns that are both frequent and pliable for that individual.

5.8 Performance and Reliability Testing

5.8.1 Load Testing and Computational Demands

To assess robustness under pressure, we executed scenario-based stress tests that mirrored realistic institutional usage rather than synthetic microbenchmarks; traffic models combined bursts of concurrent logins from dozens of users on mixed devices, simultaneous text submissions for NLP analysis, and rolling dashboard refreshes with vocabulary updates, so that authentication, write-heavy ingestion, and read-heavy analytics contended for the same shared resources. Test orchestration used Apache JMeter to drive concurrent virtual users with stepped, spike, and soak patterns (e.g., step to

50/100/150 active users every two minutes; 30-second spikes at $2\times$ baseline; two-hour soak at $1.2\times$ baseline) while Spring Boot Actuator exposed live application metrics (thread pool saturation, request latency histograms, heap/non-heap memory, GC pauses) and custom counters for queue depth, artifact processing time, and cache hit ratios; we complemented these with database telemetry (connection pool utilization, lock waits, slow query logs) so that symptoms (e.g., rising p95 latency) could be tied to specific causes (e.g., index misses, hot partitions, connection thrashing). The most salient early finding surfaced during coordinated text submissions: Python-based NLP workers showed latency spikes as request concurrency increased, which we traced to two compounding issues, sequential choke points inside the pipeline (tokenization $\rightarrow$ tagging $\rightarrow$ parsing $\rightarrow$ diagnostics running in serial even when data dependencies allowed parallelism) and unoptimized data I/O between services (large JSON payloads, synchronous RPCs, and repeated reads of the same artifacts). We introduced task queues and asynchronous job handlers backed by RabbitMQ so that the web tier could enqueue work and return immediately while workers pulled jobs at their own pace; we also restructured the pipeline to parallelize independent stages (e.g., NER and POS on separate processes), warmed model weights on worker startup to avoid cold-start stalls, switched high-volume hops to a compact binary payload, and cached immutable resources (rulesets, frequency lists) in-memory with versioned keys. Under the same heavy-load scenarios, median and p95 response times for submission $\rightarrow$ first-feedback dropped by up to 60%, throughput increased without proportionate CPU growth, and queue depth stabilized with backpressure rather than overflowing. To prevent resource collapse, we added load-shedding and graceful degradation: when worker saturation or queue length breached thresholds, non-critical feedback layers (e.g., style suggestions) were deferred while core grammar and vocabulary diagnostics continued; circuit breakers and timeouts prevented head-of-line blocking and retries used exponential backoff with idempotent job keys and a dead-letter queue for inspection. Finally, we tuned the database path by adding covering and partial indexes on hot filters, partitioning time-series tables to keep indexes small, and right-sizing connection pools (max size, max lifetime, jittered eviction) to avoid thundering herds during redeployments.

5.8.2 Ensuring Seamless Multi-Device Operation

Ensuring a consistent and equitable learner experience across devices and contexts requires both technical robustness and human-centered validation. The evaluation of Appil's interface was guided by HCI principles emphasizing perceptual consistency, accessibility, and situated usability (Norman, 2013). The learning platform was examined across major browsers and operating systems, Chrome, Firefox, Safari, and Edge on Windows, macOS, iOS, and Android, to ensure that users encountered the same affordances and feedback patterns regardless of hardware or environment. Manual

exploratory testing was combined with automated validation routines to identify interaction breakdowns and ensure behavioral predictability across devices.

Each core interaction, authenticating, submitting text, reviewing feedback, saving vocabulary, and accessing progress dashboards, was tested under multiple viewport configurations to confirm continuity of layout and information flow. Visual stability audits identified deviations in alignment, color contrast, and spatial rhythm, ensuring that perceptual cues remained stable as the system evolved. Accessibility verification involved keyboard-only navigation, focus-order checks, and screen-reader testing with NVDA, VoiceOver, and TalkBack. These evaluations confirmed that dynamic instructional feedback remained perceivable and operable without reliance on sight or pointer input, aligning with WCAG 2.1 AA and forthcoming WCAG 2.2 standards (W3C, 2018, 2023).

Environmental and situational robustness were assessed under ecological conditions reflective of real-world use: intermittent connectivity, fluctuating bandwidth, memory constraints, and partial uploads were simulated to evaluate continuity of interaction. Under these conditions, Appil exhibited fault-tolerant and user-centered recovery behaviors. Drafts persisted locally to prevent loss of cognitive effort; queued submissions retried automatically with exponential backoff to reduce learner frustration; long-running NLP analyses provided progressive visual feedback, maintaining user engagement during computational latency; and temporary backend slowdowns triggered graceful fallbacks that paused nonessential modules (such as adaptive hints) while retaining essential correctness checks. These behaviors reflect the HCI principle of resilient interaction design (Hollnagel, 2014), ensuring that users experience system continuity and trust even in degraded states. Interaction ergonomics were standardized to minimize motor and cognitive strain. Touch targets and gesture areas were harmonized across devices, hover-dependent behaviors were complemented with tap-based alternatives, and animations were adapted to respect user “reduce motion” preferences. These practices align with inclusive interaction design (Clarkson et al., 2013), ensuring accessibility for diverse physical and cognitive needs. Internationalization testing verified linguistic and cultural inclusivity: language packs were validated for bidirectional text safety, locale-aware formatting, and translation accuracy, ensuring that learners in different linguistic contexts received comparable instructional experiences. Performance thresholds were set to guarantee prompt visibility of above-the-fold content on mid-range mobile devices, reducing waiting times and cognitive load during interaction. Collectively, these efforts produced a platform that is predictable, perceivable, and inclusive, delivering a coherent sense of flow across learning environments and ensuring that usability remains integral to pedagogy.

5.8.3 Debugging Integration Bottlenecks

Diagnosing cross-service performance bottlenecks was approached as a problem of interactive system transparency, an HCI concern as much as a software engineering one. Achieving smooth, low-latency interactions requires that users experience minimal friction between intention and system response (Card, Moran, & Newell, 1983). To that end, the Appil architecture was instrumented for observability-driven development, allowing system events to be traced seamlessly across layers and correlated with user-facing feedback timing. Distributed tracing via server telemetry, assigned a unique identifier to each learner action, enabling its passage to be monitored from browser input through the gateway, API, message queues, NLP workers, and databases. This end-to-end traceability supported causal reasoning about perceived latency and reinforced design accountability by linking backend behavior directly to user experience outcomes.

Aggregated structured logs and performance metrics provided a multidimensional view of system health, exposing bottlenecks that degraded interactivity or responsiveness. Performance data revealed redundant database calls in the feedback compositor, verbose payloads between services, and unclosed connection pools under heavy load. Each of these was systematically addressed through iterative testing and optimization: query batching and read-optimized views reduced database latency; compact serialization and compression streamlined data exchange; and connection handling routines were refined to prevent resource exhaustion.

System stability during deployment was maintained through non-disruptive schema migration strategies, online index creation, shadow tables, and feature-flag rollouts, that preserved continuous learner access. Complementary profiling techniques, including flame graph analysis and heap tuning, distinguished computational bottlenecks from I/O waits, while refined garbage-collection policies on the Java side improved latency consistency. All optimizations were validated through canary deployments and monitored under quantitative thresholds such as submission acknowledgment below 800 milliseconds and sustained worker utilization under seventy-five percent.

Post-resolution analyses were integrated into operational runbooks and alerting mechanisms anchored to the same metrics that originally surfaced anomalies. This reflexive design practice embodies resilience engineering (Hollnagel, 2014) and the human-centered DevOps perspective (Kim et al., 2016), in which technical incidents become structured opportunities for organizational learning and design refinement. By framing observability as an instrument of user experience, Appil transforms low-level performance data into a feedback system for interaction quality. The result is an infrastructure where stability, transparency, and pedagogical responsiveness are not secondary properties but fundamental affordances of the user experience.

5.9 Technological and development limitations

While the current architecture of Appil reflects deliberate design choices aimed at modularity, scalability, and pedagogical alignment, it is also the product of a constrained development context. These constraints, particularly related to time, personnel, and existing technical knowledge, shaped not only the pace of development but also certain foundational decisions in the system’s design. Most notably, the decision to build the core backend architecture using Java and Spring Boot was influenced less by a comparative technical evaluation of all possible frameworks, and more by the practical realities of the development timeline and the pre-existing expertise of the programmer-researcher leading the project. Java was chosen because it allowed rapid prototyping and deployment within a limited timeframe, leveraging familiarity with the ecosystem to accelerate implementation without the steep learning curve associated with newer or more specialized platforms. This approach enabled the team to meet the urgent requirement of deploying a functional prototype by mid-2024, while still supporting the project’s long-term vision of modularity and integration. However, it also introduced some limitations. For example:

Aspect	Description
Service Optimization	Java is not natively optimized for certain types of stateless, lightweight services commonly used in high-concurrency environments.
Framework Weight	The Spring framework, while powerful, can be relatively heavy for rapid iteration compared to more minimal microframeworks such as Node.js (JavaScript) or Flask (Python).
NLP Integration	NLP-related tasks had to be offloaded to Python services due to Java’s limited ecosystem for modern computational linguistics.

Table 13 Limitations Introduced by the Java-Based Implementation

In addition to these technical factors, the development and deployment infrastructure was almost entirely provided by the researcher himself. All materials, server resources, and hosting arrangements were self-financed and managed directly by the project lead. This arrangement ensured full autonomy and rapid decision-making but also introduced constraints on budget, scaling capacity, and technical support that would not typically be present in institutionally funded projects. For example, server configuration, monitoring, and maintenance had to be handled manually, limiting opportunities to deploy more complex cloud-based architectures or external managed services.

In this sense, architecture reflects a strategic compromise: prioritizing rapid deployment and system stability over theoretical optimality. The system is functional and robust, but not necessarily the leanest or most elegant in every domain. These limitations do not detract from the viability of the learning platform, but they do contextualize its current form. Crucially, the system has been built in such a way that future refactoring, such as containerizing services, adopting more lightweight frontends, or migrating certain services to specialized stacks, is entirely feasible when resources and time allow.

5.10 Future Directions and Scalability

5.10.1 Planned Enhancements (AI, LMS Integration, Analytics)

The next iteration focuses on deepening adaptivity while tightening institutional interoperability and evidence-based decision making. On the AI side (still within our deterministic/explainable ethos), we will expand diagnostic coverage and sequencing: improved error taxonomies for article–noun agreement, prepositions, and tense–aspect will feed a lightweight mastery model (Bayesian knowledge tracing/IRT-style estimates) that calibrates difficulty and practice spacing per learner; non-generative speech components (ASR with constrained scoring, forced alignment) will enable pronunciation feedback that reports segmental and prosodic issues with transparent criteria; automatic discourse and cohesion checks will surface organization signals (connectives, referential clarity) as interpretable metrics rather than opaque scores, and an enhanced recommendation engine will use multi-objective ranking (learning gain, novelty, relevance to goals, time cost) to propose the next text, exercise, or review set with visible ‘why this?’ rationales. For LMS integration, we will implement standards-based touchpoints to reduce IT overhead and keep data consistent across systems: LTI 1.3/Advantage for single-click launch, deep linking, and assignment/grade services; OIDC/SAML for SSO; OneRoster for roster sync; and learning telemetry via Caliper or xAPI so institutions can pipe events into their analytics stacks without custom adapters; grade passback will map Appil rubrics to LMS gradebook categories with configurable weighting, while section/term provisioning will automate cohort setup and archival. Analytics will move from descriptive dashboards to a proper learning-analytics pipeline: a warehouse fed by CDC streams will power longitudinal views of vocabulary consolidation, error reduction per 1,000 words, and time-on-task adherence; instructors will see cohort heatmaps of challenge areas, early-risk flags based on engagement and accuracy trends, and outcome-aligned reports that can be exported in CSV/Parquet for research; an experiments framework will support A/B and interleaving tests with pre-registered metrics (e.g., revision uptake, delayed post-test gains), guardrails for equitable exposure, and automatic power checks; finally,

platform enhancements, PWA (Microsoft Edge Dev, 2025; Mozilla MDN Contributors, 2025) offline drafts, improved mobile breakpoints, expanded localization packs, and admin audit trails, round out operational readiness without increasing cognitive load for end users.

5.10.2 Deliberate Constraints on Generative AI and LLM

The deliberate exclusion of Large Language Models (LLMs) at this stage of development constitutes a principled design decision rather than a technical constraint. This choice reflects an intentional alignment with the values of interpretability, reliability, and pedagogical accountability, foundational dimensions of both explainable artificial intelligence (XAI) and responsible educational technology. The objective is not merely to limit computational complexity, but to ensure that every system output can be traced to a specific rule, linguistic feature, or dataset. Such traceability supports epistemic transparency (Floridi & Cows, 2019), enabling developers, educators, and learners to understand the provenance and rationale behind the system's feedback.

From a technical perspective, architecture favors targeted, deterministic NLP pipelines over generative models. Word Sense Disambiguation (WSD) employs domain-trained classifiers with explicit sense inventories and calibrated confidence thresholds, ensuring controlled ambiguity resolution. Named Entity Recognition (NER), lexical profiling, and syntactic parsing are implemented using supervised or rule-augmented models, selected for their bounded variance and measurable error profiles. Feedback generation follows a templated and parameterized approach: messages are constructed from reusable linguistic primitives, guaranteeing consistency across learners and sessions. Each corrective action is linked to a concise pedagogical rationale and one or two contrastive examples, illustrating both the incorrect and correct forms, thus operationalizing principles of form-focused instruction and metalinguistic awareness.

This deterministic posture offers several systemic advantages. It minimizes hallucinations, bias drift, and stylistic inconsistency, problems well-documented in current-generation LLMs (Bender et al., 2021). It also simplifies data governance and privacy compliance: since no user-generated content is repurposed for model retraining, all learning data remains under institutional control. Each decision, prediction, or correction can be audited against a fixed ruleset and model version, ensuring accountability and reproducibility. Moreover, the bounded behavior of deterministic systems facilitates predictable latency and resource utilization, a critical factor for scalability and equitable access in distributed educational contexts.

Pedagogically, this architecture privileges transparency over automation. Learners are provided with scaffolded feedback that promotes explicit reflection on linguistic form and usage, reinforcing noticing (Schmidt, 1990) and supporting metalinguistic development (DeKeyser, 2007). By contrast,

opaque AI-generated corrections risk undermining learner agency by offering prescriptive edits without intelligible explanation. The Appil approach instead cultivates trust and interpretive engagement: students can see why a correction was made and how it aligns with curricular outcomes. Instructors, in turn, retain authorship and control over pedagogical tone and emphasis, preserving coherence across courses and cohorts.

From an ethical and operational standpoint, this design aligns with the principles of trustworthy AI (European Commission, 2021), namely, transparency, fairness, and human oversight. Determinism and explainability are not austerity measures but deliberate pedagogical and ethical scaffolds: they sustain trust by ensuring accountability, promote equity by preventing hidden bias, and foster transfer by encouraging learners to generalize linguistic insights beyond the system. In sum, the decision to exclude LLMs during this developmental phase reflects a commitment to human-centered AI in language education, one where clarity, accountability, and pedagogical intent take precedence over opacity and automation.

5.10.3 Planned Integration of Generative AI

Generative AI will be introduced where its flexibility clearly augments, not replaces, structured pedagogy, and only behind guardrails that preserve accuracy, tone, and learner agency. Candidate use cases are narrow and well-bounded: (1) text simplification and controlled rewriting that obey readability and CEFR profiles, domain lexicons, and style constraints; (2) dialogue simulations and task-based prompts that seed creativity while remaining within topical and register limits; (3) elaborated explanations or hints that sit on top of rule-based detections, offering clearer paraphrases or additional examples without inventing new rules. Technically, integration will follow a safety-first architecture: requests are pre-filtered (PII scrubbing, profanity/bias screening), prompts are grounded with retrieval from our curated lexicon and rule bank, decoding is constrained (temperature caps, length bounds, banned-phrase lists), and outputs pass post-filters for toxicity, factual alignment to retrieved sources, and readability targets; low-confidence generations fall back to deterministic templates, and every AI-assisted element is labeled with provenance ('AI-assisted, grounded to rule X and source Y'). Deployment will proceed in shadow mode (generate but do not show), then opt-in pilot with human-in-the-loop review, before general availability with per-institution toggles; evaluation will combine automatic checks (hallucination rate vs. references, coverage/preservation of key information, readability deltas) and human ratings (adequacy, fluency, pedagogical fit), plus learning outcomes from controlled studies. Governance will codify data boundaries (no training on user content by default; narrow, time-boxed retention for safety audits), rate limits and latency budgets to protect core flows, and educator controls to set tone, register, and disallowed transformations. The

north star remains constant: use generative systems to make good pedagogy easier to deliver and understand, never to shortcut careful instruction or obscure how and why feedback is given.

5.10.3.1 Frontend Refactoring and Framework Modernization

Another high-leverage area identified for the next phase of development concerns the refactoring of the presentation layer, which currently relies on JSP views supplemented by ad hoc JavaScript. While this legacy stack has been serviceable, it now constrains responsiveness, maintainability, testability, and the integration of modern interaction and analytics paradigms. The forthcoming modernization will adopt component-based architecture implemented through a contemporary frontend framework, such as React (React Team, 2025) or Vue.js (Vue.js Core Team, 2025), in conjunction with TypeScript and structured routing and data-fetching libraries (e.g., React Router and React Query, or Vue Router). This transition aims to modularize user interface (UI) concerns, isolate side effects, and standardize data management, thereby improving the overall reliability and scalability of the interface layer.

A design system grounded in accessibility and consistency will be implemented to unify the visual and interaction language of the learning platform. Frameworks such as Tailwind CSS (Tailwind Labs, 2025) or Material Design (MUI, 2025) will provide utility-first tokens for spacing, color, and typography, supporting responsive layouts, coherent patterns (forms, steppers, dialogs), and dark-mode compatibility. This systematic approach will ensure that pedagogical affordances, such as progressive disclosure and scaffolded feedback, are conveyed through stable, interpretable visual cues.

The refactoring will also extend to progressive web application (PWA) capabilities, enabling offline learning and robust operation in constrained connectivity contexts. Technologies such as Service Workers (Workbox) and IndexedDB will manage background synchronization, draft caching, and submission reliability, ensuring learners can compose and review feedback seamlessly across devices. Performance optimization will be addressed through code splitting, route-level chunking, deferred script loading, and intelligent asset prefetching. For public-facing resources requiring visibility and fast first render, server-side rendering (SSR) or static site generation (SSG) will be selectively employed, while authenticated learning flows will remain client-rendered to balance security and performance.

A Backend-for-Frontend (BFF) layer will formalize data contracts between microservices and the presentation layer, ensuring that APIs deliver UI-ready payloads, enforce authentication and authorization (OIDC/SAML), and centralize error handling. This architectural mediation reduces coupling, mitigates network overhead, and shields the client from backend volatility. State management will be pragmatic, distinguishing between server state, managed through query-caching

utilities (e.g., React Query, RTK Query, or Vue Query), and UI state, handled through lightweight stores for interface controls and feature flags. Form validation will employ schema-based approaches (Zod or Yup) with accessible error messaging and screen-reader support, ensuring consistency with WCAG 2.1 AA accessibility standards (World Wide Web Consortium, 2018, 2023).

Accessibility and inclusivity constitute baseline requirements. The refactored interface will adhere to the Web Content Accessibility Guidelines (WCAG), implementing focus management, ARIA landmarks, reduced-motion alternatives, and full keyboard navigability. Automated accessibility checks (axe) will complement periodic manual audits using screen readers (NVDA, VoiceOver, TalkBack) to maintain compliance and usability for all learners. Internationalization (i18n) will be treated as a first-class concern, employing ICU message formatting, locale-aware handling of numerical and temporal data, and bidirectional layout safety, while all interface text will be externalized into translation catalogs to enable rapid localization.

Modernization will proceed incrementally through a strangler-fig pattern, whereby the existing JSP application remains operational while new routes and modules are mounted as micro-frontends. Early migrations will target self-contained components (e.g., login, vocabulary lists, feedback panels), followed by more complex, high interaction flows such as writing and feedback loops or educator dashboards. A continuous integration and delivery (CI/CD) pipeline will govern deployments, incorporating static analysis, type checking, unit and component testing (Jest, Vitest, Storybook), end-to-end validation (Playwright, 2025), and visual regression monitoring with defined performance budgets. Controlled release mechanisms, feature flags, canary deployments, and A/B testing, will ensure iterative refinement without service interruption.

Security and reliability will be treated as integral architectural dimensions rather than post hoc additions. Measures will include Content Security Policy (CSP), Subresource Integrity (SRI), token rotation, and rigorous cookie management (SameSite and HttpOnly), complemented by dependency scanning and redaction of sensitive data in logs. Telemetry will instrument pedagogically significant user actions (e.g., text submission, feedback application, vocabulary revision), measuring rate, error, and duration (RED metrics) to align interface optimization with learning outcomes rather than superficial engagement metrics.

The success of the refactoring effort will be evaluated through both technical and pedagogical indicators: measurable reductions in latency and interaction errors ($\geq 30\%$ improvement in time-to-interactive on mobile mid-tier devices), decreased support requests related to navigation and clarity ($\geq 40\%$), and enhanced learner engagement and revision uptake as evidenced by usability studies. Potential risk framework version churn, bundle size inflation, and double rendering under SSR, will

be mitigated through the use of long-term support (LTS) releases, strict performance budgets, and empirical validation of SSR benefits prior to adoption.

Collectively, this modernization initiative strengthens Appil's alignment with human-computer interaction principles, pedagogical transparency, and sustainable software engineering practices. The refactored presentation layer is thus conceived not merely as a technical enhancement, but as a structural enabler of equitable, accessible, and cognitively informed language learning experiences.

5.10.3.2 Institutional Integration and Scaling

In parallel, this research will advance the institutional integration of Appil, ensuring responsible scalability across cohorts, languages, and disciplinary domains. Integration with institutional Learning Management Systems (LMS) will adhere to established interoperability standards, thereby minimizing bespoke development and facilitating seamless IT onboarding. Specifically, Learning Tools Interoperability (LTI 1.3/Advantage) will be employed to enable secure launches, deep linking, and assignment-grade services; OIDC/SAML will support single sign-on; OneRoster will manage roster synchronization; and learning telemetry will be streamed through Caliper or xAPI, allowing educational institutions to incorporate Appil event data directly into their analytical infrastructures without custom adapters.

Assessment interoperability will be ensured through grade passback mechanisms that map Appil's rubric structures to LMS gradebook categories with configurable weightings and late or missing submission policies. Administrative automation, such as section and term provisioning, roster updates, and archival routines, will sustain data consistency across systems. Educator dashboards will evolve from descriptive visualizations to actionable analytics, providing cohort-level heatmaps of linguistic challenge areas (e.g., articles, prepositions, tense/aspect), thematic vocabulary coverage, and engagement trajectories with early-risk indicators. Drill-down capabilities will include anonymized exemplars and exportable datasets (e.g., CSV, Parquet) for institutional research. In alignment with data governance frameworks, the analytics layer will separate personally identifiable information (PII) from analytical identifiers, enforce role-based access controls (RBAC), and maintain comprehensive audit trails for sensitive operations.

The learning platform's language and domain expansion strategy will follow a repeatable, modular pattern. Language packs will define tokenization, morphological rules, and CEFR-aligned lexicons, while domain packs (e.g., municipal services, healthcare, hospitality) will encapsulate controlled vocabulary, collocational patterns, register constraints, and genre-specific templates. Both artifacts will be version-controlled and independently testable, enabling rapid extension without the need for re-engineering core systems.

Appil's microservices architecture provides the foundation for incremental rollout and sustainable scaling. New NLP modules and language/domain packs will be deployed as independent services with semantic versioning, backward-compatible APIs, and continuous health monitoring. A Backend-for-Frontend (BFF) layer abstracts internal service complexity, ensuring frontend stability as internal modules evolve. Change Data Capture (CDC) streams will aggregate operational data into a centralized warehouse to support cross-institutional analytics. Scalability will be achieved through horizontally replicated stateless services behind autoscaling groups, elastic work queues (e.g., RabbitMQ or Kafka) dynamically scaled according to queue depth and service-level objectives (RabbitMQ Team, 2025), and database partitioning strategies (temporal and tenant-aware sharding) complemented by read replicas for analytical workloads. System safeguards will include per-tenant quotas, request shaping, and adaptive degradation policies to maintain pedagogical continuity under high load.

Compliance and data privacy will remain central to the learning platform's design. Regional data residency, configurable retention windows, purpose-limited data exports, and adherence to FERPA and GDPR requirements (including erasures through keyed tombstones and partition pruning) will ensure ethical data stewardship. Authentication and provisioning integrations will respect institutional claims for RBAC, defining granular permissions for learners, educators, coordinators, and administrators, with immutable audit logs for critical operations.

Finally, an integrated experimentation framework will allow institutions to pilot features safely through opt-in cohorts, power analyses, and equitable exposure checks. Observability infrastructure, comprising distributed tracing, structured logs, and service-level objective (SLO) dashboards, will ensure that integration and performance issues are traceable across environments. Collectively, these design principles position Appil as a pedagogically grounded, interoperable, and ethically responsible learning platform, one that modernizes the learner experience, integrates seamlessly with institutional ecosystems, and scales to new linguistic and disciplinary domains without architectural disruption.

5.11 Conclusion

At the current stage of development, efforts are underway to extend Appil into a more adaptive and mobile-responsive form designed specifically for smartphone use. This ongoing iteration builds directly upon the architectural and methodological foundations outlined in this chapter, integrating many of the technologies already discussed. The new mobile version employs a re-engineered presentation layer developed in React, enabling dynamic rendering, offline readiness, and improved accessibility across mobile interfaces. Core linguistic and adaptive components, such as the Python-

based NLP services, RESTful APIs, and distributed observability stack, are being refactored to support lightweight communication and seamless synchronization in constrained environments. The overarching objective of this next phase is to preserve the pedagogical transparency and cognitive adaptability of Appil while ensuring that learners can engage with its intelligent feedback and vocabulary systems fluidly across contexts and devices. In doing so, the project continues to evolve toward a unified, device-agnostic learning ecosystem that bridges theoretical rigor with practical usability in real-world language learning scenarios.

6 *Fieldwork: Study Design, Testing, and Analysis*

6.1 Introduction

Language proficiency in English has become an essential component of effective public administration, particularly in contexts where institutions increasingly interact with international stakeholders and digital platforms. Despite this, employees in many local government offices often have limited opportunities for structured language development. Addressing this gap requires initiatives that not only provide accessible training but also generate research-based insights into adult language learning within institutional settings.

The present project responds to this need through pilot intervention with the Municipality of Fisciano (Salerno). It brings together principles of language pedagogy, digital innovation, and national policy priorities, with the dual aim of enhancing professional skills and contributing empirical evidence to the field of adult education.

The initiative was framed within Italy's *Piano Nazionale di Ripresa e Resilienza* (PNRR), which emphasizes innovation, resilience, and collaboration between academic institutions and public administrations. In particular, the project aligned with Mission 4 (Education and Research) and Mission 1 (Digitization, Innovation, Competitiveness, Culture), both of which prioritize the creation of educational models that are scalable, replicable, and integrated into institutional settings.

From a pedagogical perspective, the project responded to well-documented challenges in adult second language education, especially for learners with fragmented exposure to foreign languages and varying degrees of digital competence. By situating the intervention within a municipal office, the research also aimed to address a broader gap in the literature: the relative scarcity of studies that investigate workplace-based language training in the public sector.

Thus, the project served a dual mission: to enhance the professional development of local government personnel and to contribute empirical evidence to the field of adult education, particularly concerning the integration of digital resources, institutional priorities, and learner diversity.

6.1.1 Project Details

The Municipality of Fisciano (Salerno) was chosen as the pilot site for both logistical and strategic reasons. Its proximity to the University of Salerno facilitated academic collaboration, while the municipality's explicit interest in staff development and modernization provided a supportive context

for implementation. The medium size of the administration ensured that the intervention could be realistically scaled and monitored.

In consultation with municipal officials, twenty employees were selected to participate, ranging from 36 to 65 years. Selection criteria included professional role, degree of public interface, and willingness to engage in digital learning environments. This demographic composition was particularly significant from a research perspective, as it reflected the diversity typical of adult education contexts, including variation in schooling, foreign language exposure, and technology use (Milroy & Gordon, 2003). The central hypothesis guiding the intervention was that basic English competence represents both an achievable and valuable outcome for public administration personnel when instruction is adapted to their cognitive, professional, and motivational profiles. Structured over a six-month period (June–December 2024), the program combined face-to-face diagnostic activities with continuous interaction through a custom-built digital platform. Importantly, the learning platform remained accessible beyond the pilot phase, ensuring that the municipality retained a living, adaptive resource to support ongoing professional development and future staff training.

6.2 Planning phase

The project design was informed by a structured, multi-phase planning process to ensure pedagogical rigor, contextual relevance, and alignment with established best practices in adult language education (Knowles et al., 2014; Macalister & Nation, 2019). The initial phase, a comprehensive Needs Analysis, involved consultations with municipal leadership to identify institutional expectations and assess existing language competencies within the organization (Richards, 2001). These discussions highlighted a strong institutional emphasis on the development of English language skills directly applicable to professional contexts, such as interpreting administrative documents, responding to routine inquiries from international stakeholders, and participating in regionally networked initiatives, particularly those related to EU funding mechanisms. Following this, a Participant Profiling process was undertaken, during which each learner completed an intake form capturing demographic information, linguistic background, self-assessed digital literacy, and previous exposure to English instruction. This data was critical in shaping both the customization of curricular content and the selection of appropriate instructional modalities. It was assessed with the on-purpose generated profiling questionnaire, designed in accordance with established needs analysis frameworks (Richards, 2001; Macalister & Nation, 2019). The integral version is available in the appendix I ‘A at *Profiling questionnaire*’ section, page 424.

Building on the findings from these initial phases, the curriculum was subsequently aligned with the Common European Framework of Reference for Languages (CEFR), using the 2020 Companion Volume as a foundational guide. This version was selected due to its updated descriptors and explicit integration of digital mediation competencies, features particularly pertinent to adult and professionally engaged learners. Concurrently, a digital platform was chosen based on criteria that considered both technological accessibility and pedagogical functionality. Key selection requirements included an intuitive user interface suitable for individuals with limited digital proficiency, interactive content capabilities (such as quizzes and vocabulary practice tools), mechanisms for participation tracking and progress monitoring, and compatibility with CEFR-aligned materials.

In parallel with the learning platform selection, an assessment instrument was designed to measure both lexical knowledge and learners' attitudinal orientation toward English language learning. This tool, adapted from Nation's Monolingual Vocabulary Size Test (Nation, 2001), was modified to incorporate a five-point Likert scale, thereby enabling the capture of both cognitive outcomes and affective dimensions, including learner confidence and perceived relevance of English to professional responsibilities. The implementation phase commenced with two onboarding sessions, one conducted in person and one online, which introduced participants to the learning platform and outlined CEFR-based learning objectives. The instructional content was initially organized into six thematic modules, each approximately one month in duration, focusing on foundational workplace vocabulary, reading comprehension of administrative texts, email writing, basic grammatical structures (e.g., simple present tense, prepositions of place), listening comprehension of clear, slow-paced English, and basic mediation tasks such as translating brief notices or agendas. The thematic organization of modules was informed by principles of thematic analysis, ensuring that each unit reflected coherent clusters of meaning and communicative function relevant to participants' professional contexts (Braun & Clarke, 2006).

To support sustained engagement and reinforce learning habits, well-documented predictors of success in adult education, weekly check-ins were conducted, involving short assignments and micro-assessments, with ongoing monitoring by the researcher. At the conclusion of the program, participants completed a Final Reflective Questionnaire aimed at evaluating vocabulary development, self-reported confidence, perceived professional utility of English, and overall satisfaction with the intervention.

In its finalized form, the instructional plan was reorganized into eight discrete lessons, each designed to reflect the specific institutional needs identified in the initial consultation with municipal leadership. This refinement represented both a pedagogical enhancement of the original modular structure and a strategic response to the municipality's request for a fully digitized, continuously

accessible learning experience. Substantial effort was directed toward adapting and reformatting the instructional content for seamless delivery via the selected digital platform, ensuring that all learning materials, including multimedia resources, interactive exercises, and assessment instruments, were available online at all times. This digital-first configuration facilitated learner autonomy and flexibility, enabling participants to progress at their own pace while balancing professional commitments. The resulting intervention thus fulfilled the municipality's objective for a scalable, modern, and professionally relevant language training program tailored to the specific needs of its workforce.

6.3 Profiling Learners: Diagnostics for Learner-Centered Instruction

A critical component of the instructional design process was the comprehensive profiling of participants, which served as the foundation for tailoring pedagogical strategies to specific needs, to accurately capture the experiences, motivations, and linguistic profiles of the learner group, the project adopted a data-driven approach to needs analysis, particularly mindful of the inherent heterogeneity typical of adult education cohorts. Variability in prior linguistic exposure, digital competence, and learning expectations necessitated a structured intake and diagnostic phase prior to instruction. Central to this process was a multi-section questionnaire, designed to elicit both quantitative and qualitative insights.

The instrument was developed with reference to the Common European Framework of Reference for Languages (CEFR), specifically the 2020 Companion Volume, and integrated select items adapted from Nation's Monolingual Vocabulary Size Test, thereby blending self-perception measures with performance-based diagnostics. This hybrid design allowed for a nuanced profiling of both the cognitive (e.g., lexical knowledge, grammatical awareness) and affective (e.g., confidence, motivation) dimensions of learner readiness.

The first section gathered essential demographic and biographical data: age, gender, nationality, languages spoken at home, previous exposure to English, and time elapsed since last formal instruction. These contextual variables were crucial for identifying factors such as linguistic interference, educational background, and plurilingual awareness, elements highlighted in the CEFR's Learner Biography component.

Subsequently, learners completed a series of reflective self-assessment items targeting their perceived needs and priorities across key communicative domains (speaking, listening, reading, writing). These items were aligned with CEFR-aligned can-do descriptors and organized across proficiency bands (A1–C2), enabling participants to self-position along a recognized proficiency continuum. A five-

point Likert scale (Likert, 1932) was employed to gauge their level of confidence in typical communicative tasks. For example, listening descriptors ranged from understanding basic, clearly articulated speech to interpreting fast, complex discourse, while writing descriptors moved from producing simple phrases to constructing extended argumentative texts.

The vocabulary task, adapted from Nation's test (Nation, 2001), required learners to identify synonyms of bolded words in contextualized sentences (e.g., selecting hours as closest in meaning to *a lot of time*), assessing both lexical recognition and conceptual accuracy. The grammar section featured a short narrative about Albert Einstein, where participants were prompted to supply the correct past-tense forms of regular and irregular verbs (e.g., developed, was, gave), targeting A1–B1 level grammatical structures. Together, these tasks provided tangible, performance-based indicators of both receptive and productive skill areas in need of reinforcement.

While the initial instrument did not explicitly assess digital mediation or spoken interaction, its design was nonetheless informed by the CEFR's expanded communicative competence model. Future iterations will aim to incorporate these dimensions more directly, further aligning the intake tool with the CEFR's holistic vision of language use.

Finally, the diagnostic findings played a pivotal role in shaping a learner-centered, differentiated curriculum. The diversity in lexical range, grammatical competence, and self-reported confidence underscored the necessity of designing a learning experience that was both accessible and cognitively stimulating. The insights gleaned from this process informed not only content development and instructional pacing but also served as a baseline for ongoing formative assessment, ensuring that the course addressed both institutional goals and the expressed needs of the learners.

6.3.1 Demographic and Linguistic Profile

The data collected reveal a cohort of adult learners mostly in the 35–50 age range, with a balanced gender distribution. Almost all respondents declared Italian nationality, and Italian was the predominant language spoken at home, sometimes in combination with local dialects. This confirms a context of limited multilingual exposure, in which English is perceived primarily as a foreign language studied in formal education rather than integrated into daily linguistic repertoires.

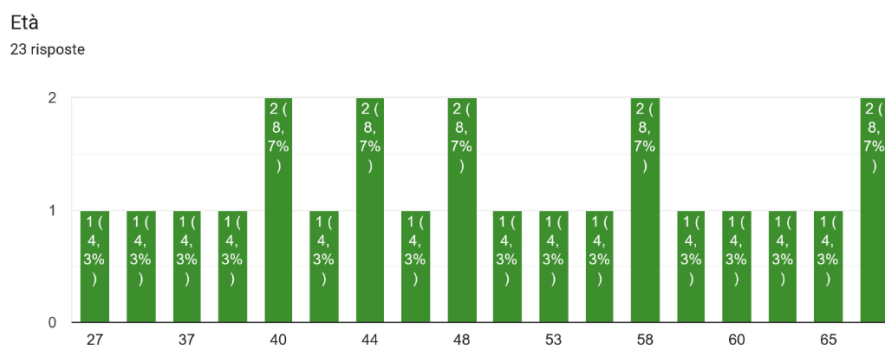


Figure 30 Age statistics

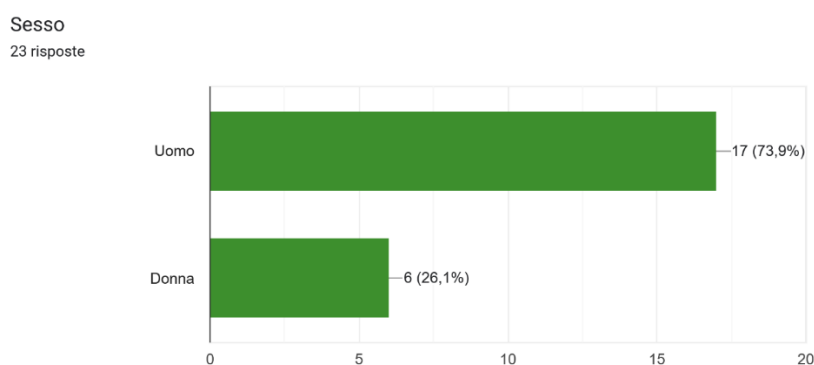


Figure 31 Male/ Female statistics

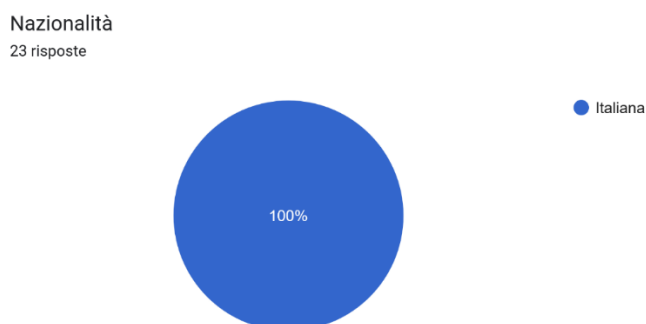


Figure 32 Nationality statistics

In terms of biographical background, learners reported having studied English mainly during secondary school or at university, with occasional additional training courses. The time elapsed since their last formal study varied widely, from 6 years to over 30 years. Such discontinuity represents a potential fragility, as previously acquired knowledge may have fossilized or been lost due to lack of practice.

Lingue parlata in famiglia:

23 risposte

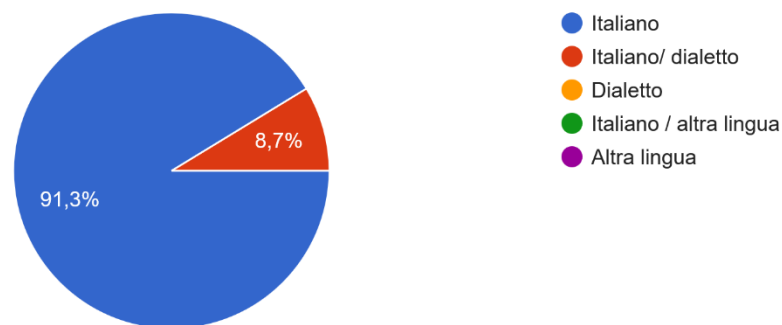


Figure 33 Preferred everyday spoken language

This questionnaire aims to better understand their current English skills and learning preferences. Knowing how well you understand spoken English helps us identify the best strategies and materials to improve their comprehension, while assessing their writing ability allows us to design activities that strengthen their expression and accuracy. By asking which skill you feel needs the most improvement, we can focus on the areas that matter most to you. We also want to know whether you enjoy learning through topics of personal interest, since motivation plays a key role in successful language learning. Understanding when you last studied English helps us gauge how recently you have practiced the language and how much review you might need. Finally, identifying the language you find easiest to communicate in allows us to consider their linguistic background and tailor the learning experience accordingly. Your honest answers will help us create a personalized and engaging learning path that truly supports your progress.

Quale è stata l'ultima occasione in cui hai studiato l'inglese?

23 risposte

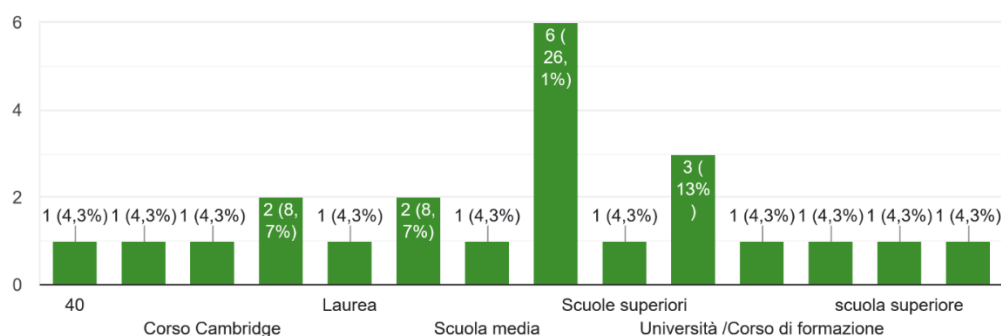


Figure 34 When was the last time you studied English?

Attraverso quale lingua ti riesce più facile comunicare?

23 risposte

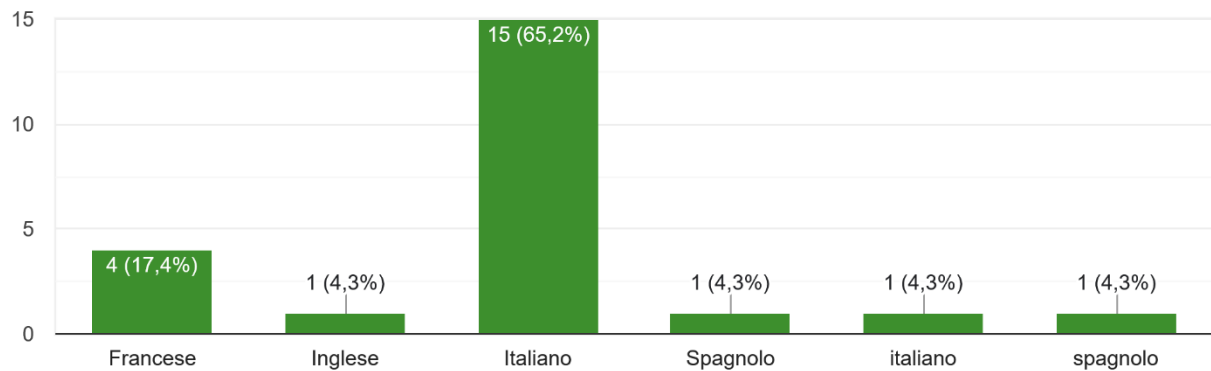


Figure 35 Which language do you find easiest to communicate in?

6.3.2 Perceived Needs and Self-Assessment

Most participants identified listening comprehension as the primary area requiring improvement, followed by speaking, while writing was considered less urgent. Some respondents reported needing to improve across all skills. This orientation highlights the importance of prioritizing oral and interactive activities, consistent with the CEFR's action-oriented approach, which frames learners as social agents engaged in real-world tasks.

Responses to the self-assessment section, structured around CEFR *can-do* descriptors (A1–C2), indicate a heterogeneous distribution:

- Listening: most learners placed themselves at A2–B1, reporting an ability to understand the gist of clear speech on familiar topics, but difficulty in following fast or complex input.
- Writing: the majority reported being able to produce short or simple texts (A2–B1), while only a minority claimed competence in advanced writing (B2–C1).
- Reading: many indicated that they could decode short, everyday texts, while some reported understanding more complex articles and professional documentation (B1–B2).

Overall, the self-assessment data suggest an average proficiency level between A2 and B1, with some variation toward A1 and B2.

In quale abilità pensi di dover migliorare il tuo inglese?

22 risposte

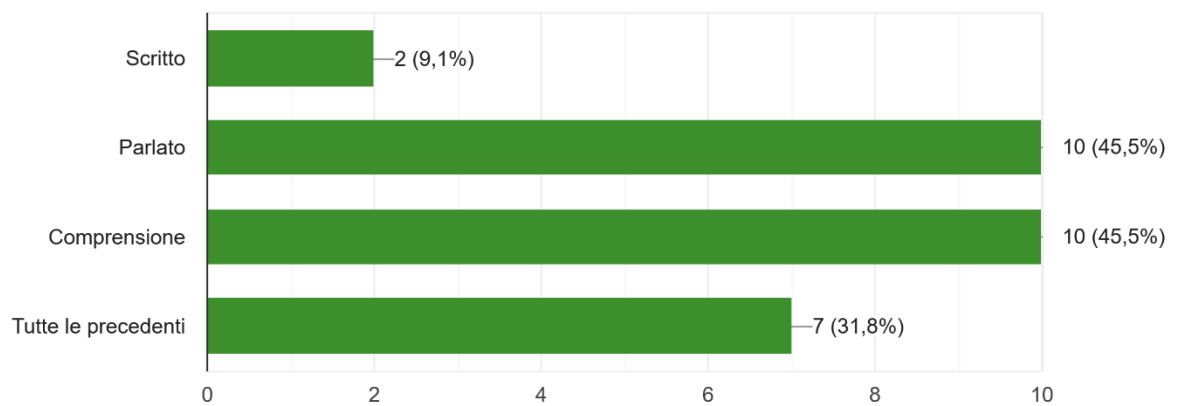


Figure 36 Speaking and listening are the main areas participants wish to improve.

Apprendere da contenuti di tuo interesse e in lingua inglese, sarebbe di tuo gradimento?

23 risposte

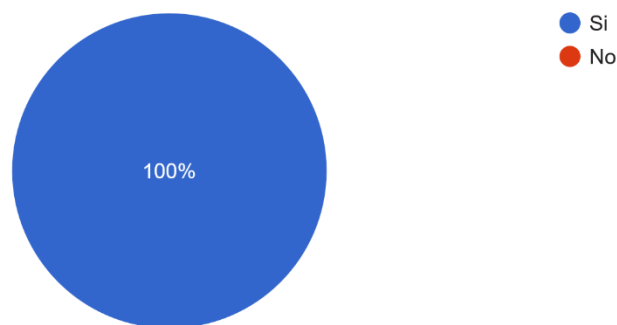


Figure 37 Would you enjoy learning through topics of personal interest in English?

Capire l'inglese: come definisci la tua capacità?

21 risposte

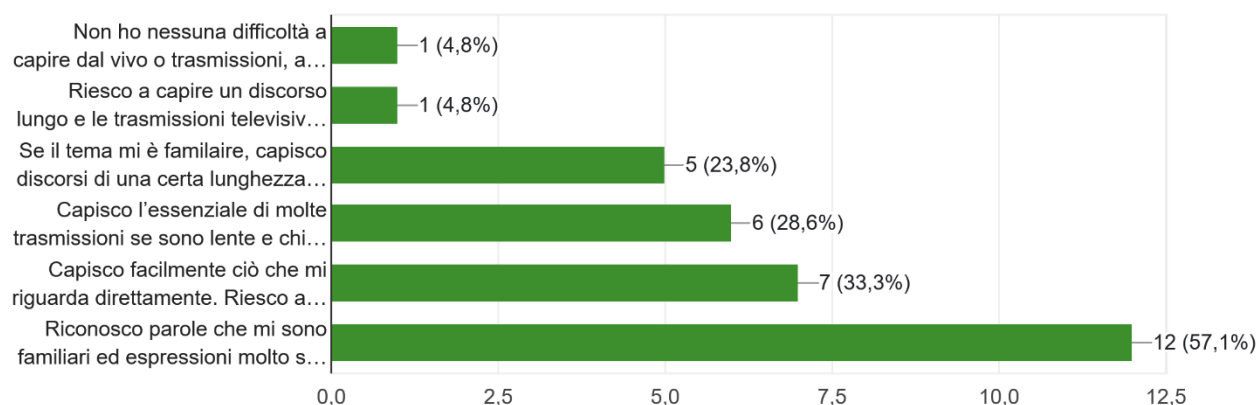


Figure 38 Understanding English: How do you define your ability?

Scrivere in inglese: come definisci la tua capacità?

19 risposte

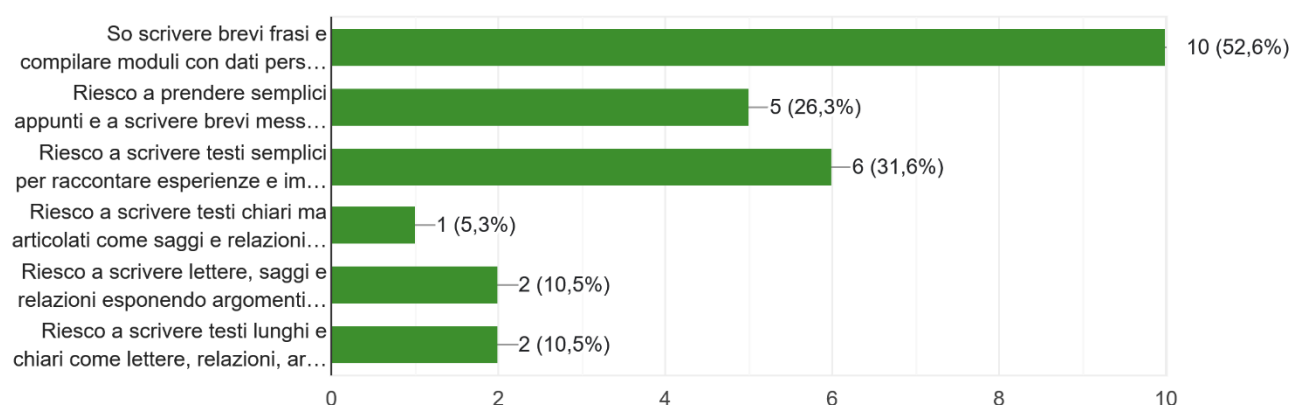


Figure 39 Writing in English: How do you define your ability?

The writing and understanding questions were developed in alignment with the Common European Framework of Reference for Languages (CEFR) principles, ensuring that the assessment reflected internationally recognized descriptors of language competence. In particular, the tasks were designed to elicit evidence of communicative ability across key dimensions such as range, accuracy, coherence, and comprehension. By grounding the questionnaire in CEFR benchmarks, the diagnostic process provided reliable insights into learners' receptive and productive skills, situating their performance within a transparent and comparable framework. This approach also ensured coherence between the diagnostic phase and subsequent instructional planning, allowing for a curriculum that responds to learners' actual proficiency levels and progression pathways as defined by the CEFR scale.

6.3.3 Vocabulary Section

The vocabulary test, adapted from Nation's framework , produced uneven results. Certain items were correctly identified by the majority of participants, such as '*They have a lot of time*' (answered as *hours*, 70%), '*I have many*' (answered as *a lot*, 61%), and '*Where is your shoe?*' (answered as *the thing you wear on your foot*, 56%). However, several items proved considerably more challenging. Only 39% of respondents provided the correct answer for '*Where did you go?*' (*to what place*), and just 26% recognized '*unhappy*' as a synonym of '*I am upset*.' The lowest performances were observed for '*Is this the right figure?*' (*number*, 22%) and '*I can deal with it*' (*fix*, 30%).

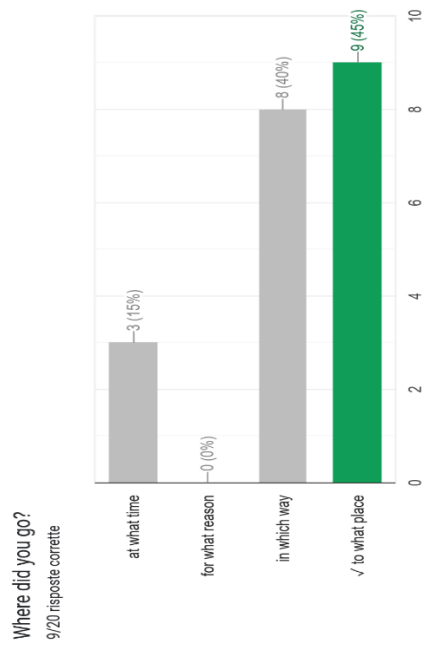


Figure 40 Vocabulary Section Question 1

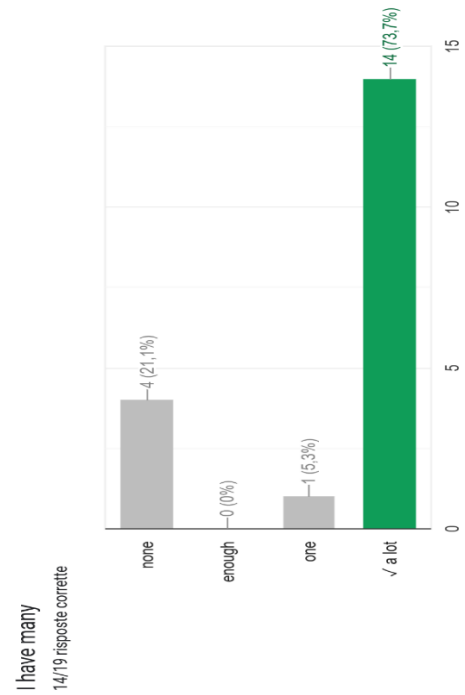


Figure 41 Vocabulary Section Question 2

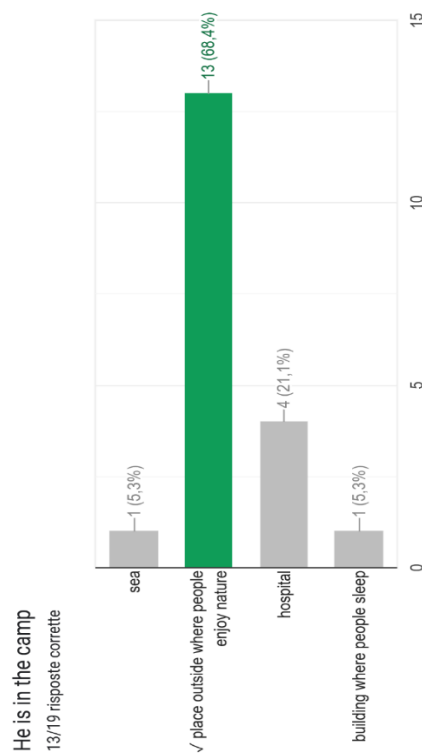


Figure 42 Vocabulary Section Question 3

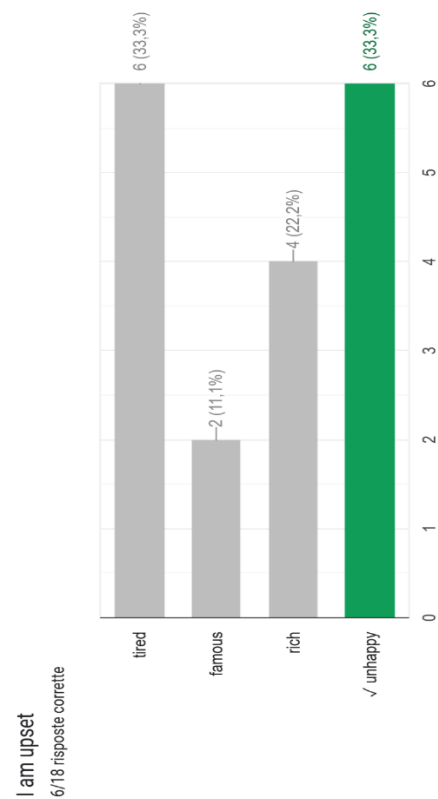


Figure 43 Vocabulary Section Question 4

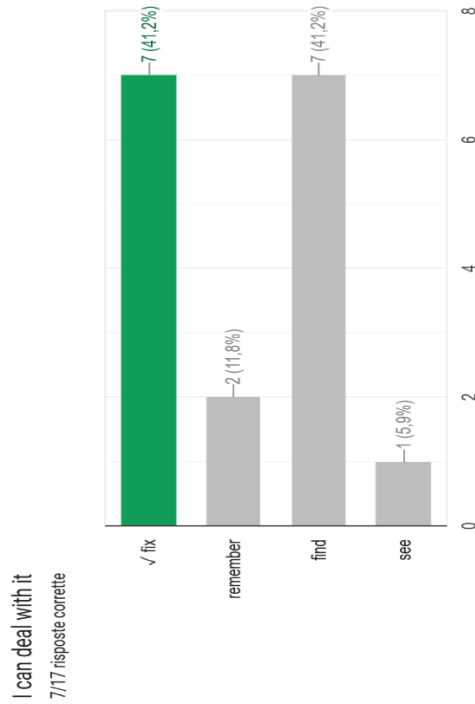


Figure 44 Vocabulary Section Question 5

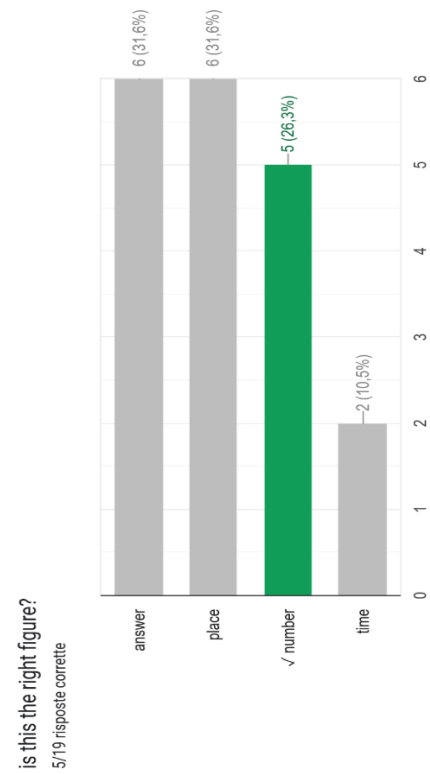


Figure 45 Vocabulary Section Question 6

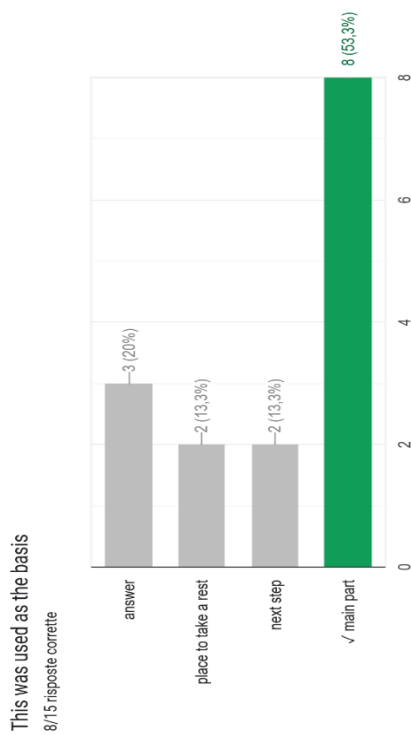


Figure 46 Vocabulary Section Question 7

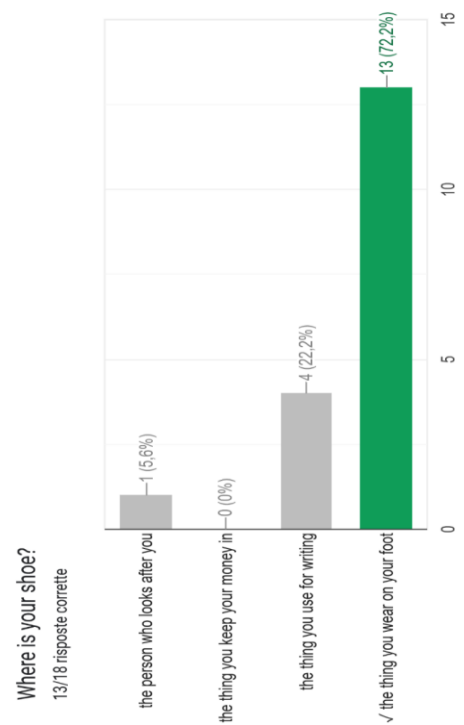


Figure 47 Vocabulary Section Question 8

These findings suggest that while basic vocabulary is generally well recognized, there are substantial gaps in learners' knowledge of intermediate and idiomatic lexis, particularly concerning polysemy, collocations, and figurative meanings. This unevenness indicates that participants may rely heavily on literal translation and word-for-word associations rather than developing a nuanced understanding of contextual meaning. Difficulties with items such as '*deal with*' and '*figure*' reveal limited exposure to multi-word expressions and semantic flexibility, features that are central to real-world language use.

Moreover, the results point to a vocabulary range restricted largely to high-frequency words, with reduced familiarity with mid-frequency and low-frequency items that often appear in authentic communication. This pattern is consistent with learning histories centered on traditional instruction, where vocabulary is presented in isolation rather than through contextual or communicative tasks. Addressing these lexical gaps will therefore require systematic exposure to natural language input, focus on lexical chunks, and opportunities for noticing and using new words in meaningful contexts. Overall, the test results underscore the importance of integrating explicit vocabulary instruction into the course design, ensuring that learners move beyond recognition toward active, flexible, and context-sensitive lexical competence.

6.3.4 Grammar Section

In contrast, the grammar task, based on a short narrative about Albert Einstein, showed uniformly high performance. All learners correctly selected the required verb forms (was, developed, wrote, The use of verbs such as gave, spent, changed, believed, passed away, and continues suggests a solid mastery of basic verb structures, particularly the simple past tense in both its regular and irregular forms. It is plausible that this competence stems from traditional school instruction, which has historically emphasized formal grammar and verb conjugation patterns. The combined analysis of self-assessment data and diagnostic tasks highlights a marked asymmetry between grammatical competence and lexical knowledge. While grammar appears consolidated and relatively homogeneous across the group, vocabulary, especially at the intermediate, idiomatic, and metaphorical levels, emerges as a shared area of weakness. This discrepancy points to the need for a pedagogical shift away from an exclusive focus on formal accuracy and toward the development of authentic communicative competence. Such an approach should incorporate systematic vocabulary expansion, consistent exposure to authentic input, and guided opportunities for meaningful interaction. In addition, demographic and experiential data underline the importance of addressing affective factors in the learning process, fostering confidence and motivation alongside linguistic proficiency. Overall, the diagnostic phase produced a comprehensive profile of the learner cohort,

revealing an average proficiency level situated between A2 and B1, with well-consolidated grammatical competence contrasted by significant lexical gaps, perceived difficulties mainly in listening and speaking, and prior learning experiences that were both fragmented and distant in time. These insights guided the design of a differentiated curriculum that balances accessibility with cognitive challenge, in alignment with CEFR principles. Furthermore, the questionnaire fulfilled a dual function as both a diagnostic instrument and a baseline for formative assessment, ensuring that the instructional pathway would respond not only to the strategic objectives of the municipality but also to the concrete needs and aspirations of the learners themselves.

Per ogni spazio vuoto scegli la forma corretta del verbo, selezionando la casella giusta: Albert Einstein (1) _____ (be) a famous physicist. He (2) ... (11) _____ (continue) to influence modern physics.

14/23 risposte corrette

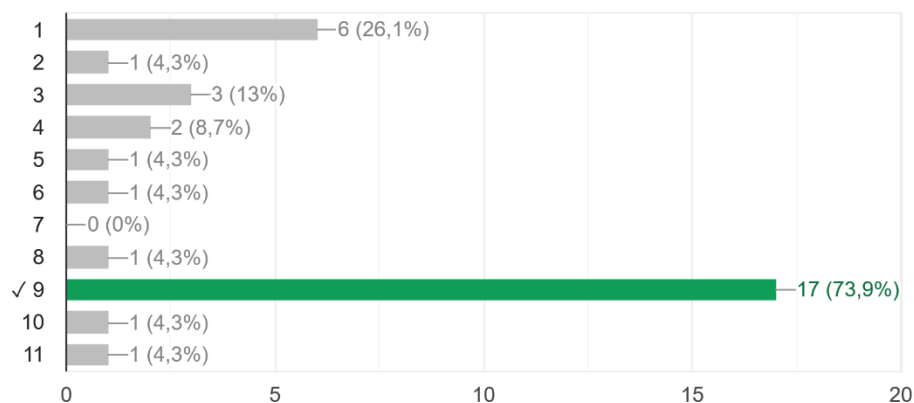


Figure 48 Grammar test results

6.4 Platform setup

A defining characteristic of the learning platform setup process was its iterative, dialogic nature, grounded in continuous interaction with participants and responsive adaptation to their emerging needs and interests. Rather than adopting a static, one-size-fits-all approach, the development and refinement of the digital learning environment unfolded as an ongoing process of negotiation between the project team and the learners. From the initial onboarding sessions through the full implementation of the curriculum, participant feedback played a central role in shaping both the content and the functionality of the learning platform. This approach was underpinned by the recognition that adult learners bring diverse professional experiences, varying degrees of linguistic competence, and differing levels of digital literacy which necessitate a dynamic and responsive pedagogical framework.

Day-to-day engagement with participants revealed evolving preferences regarding content relevance, learning modalities, and platform usability. These insights were collected through informal check-ins, platform usage analytics, in-platform feedback tools, and regular communication via messaging functions and synchronous sessions. In many cases, learners articulated specific thematic interests or highlighted practical challenges they encountered in navigating the learning platform or understanding specific linguistic concepts. Such input directly informed micro-level adjustments in real time, for instance, the addition of targeted vocabulary exercises related to administrative correspondence, the simplification of platform navigation structures, or the incorporation of multilingual glossaries and visual aids to support comprehension.

Moreover, the continuous feedback loop allowed the instructional team to adapt the pedagogical sequencing and instructional strategies to better align with participants' learning rhythms and professional responsibilities. Modules were periodically reordered, expanded, or condensed based on collective progress and individual learner trajectories. For example, when several participants expressed difficulty with email writing in professional contexts, an additional set of scaffolded writing tasks was integrated into the module sequence, accompanied by annotated exemplars and interactive revision activities. Similarly, based on user activity data indicating reduced engagement with certain resource formats, content was reformatted into shorter, more interactive units to enhance accessibility and motivation.

This adaptive methodology was not limited to content updates but extended to the affective dimension of learning as well. The evolving platform served as a space for learners to express their growing confidence, shifting attitudes toward English language use, and reflections on their professional identity as it intersected with linguistic development. In response, motivational messaging, peer-to-peer encouragement features, and opportunities for informal sharing of work-related language tasks were incorporated to foster a sense of community and learner agency.

Ultimately, the learning platform became a co-constructed space, shaped not only by institutional goals and pedagogical theory but by the lived experiences and input of the learners themselves. This participatory design approach contributed to high levels of engagement, perceived relevance, and learner satisfaction, underscoring the value of responsiveness and flexibility in digital language education for adult professionals. The project thus demonstrates that meaningful integration of learner voice, not merely at the outset but throughout implementation, is critical for fostering a learning environment that is both pedagogically sound and contextually attuned.

6.5 The Didactic Plan Behind Appil

The course within the application is specifically designed to cater to absolute beginners who are entirely new to the English language. Recognizing the unique challenges faced by novice learners, the curriculum introduces foundational concepts in a clear and accessible manner, ensuring that users can begin their language learning journey with confidence (Ellis, 2008). The initial lessons are crafted to impart essential vocabulary, basic grammatical structures, and fundamental communication skills, enabling learners to grasp the core elements of the language. These lessons are deductively based, where rules and concepts are explicitly presented, ensuring that learners build a clear and systematic understanding from the start (Richards & Rodgers, 2001). To complement this, inductive activities are integrated into the curriculum, allowing learners to discover language patterns on their own through examples and practice, fostering deeper comprehension (Ellis, 2008).

Many of the design choices applied to the course are directly justified by the need to tailor the product for absolute beginners, in compliance with the guidelines set forth by the PhD framework of external collaboration. A comprehensive assessment of learners' needs revealed that the target audience required a curriculum that started from the very bottom, addressing the lack of prior exposure to English. This insight guided the development of a platform that provides not only the essential building blocks of language learning but also a pedagogical structure that ensures gradual, steady progression (Macalister & Nation, 2019). The collaboration with external stakeholders further emphasized the importance of aligning the learning platform's features with the specific needs of beginners, ensuring that the educational experience is robust, inclusive, and supportive from the outset. By offering a structured yet flexible introduction to the language, the application empowers beginners to build a solid linguistic foundation, upon which they can continue to develop their skills through the more advanced activities and features of the program.

As a consequence, we designed a didactic plan specifically tailored for beginner-level students of English, targeting learners at the A1 level of the Common European Framework of Reference for Languages (CEFR, 2020). The primary goal of this plan is to guide students towards achieving level A2. This didactic plan incorporates foundational elements necessary for early language acquisition, emphasizing a structured approach to developing essential skills in vocabulary, grammar, and basic communication. By systematically introducing core concepts and gradually increasing complexity, the plan ensures that learners build a solid foundation while progressing steadily through the language learning continuum. The application's features are aligned with this plan to provide a cohesive and supportive learning experience, fostering advancement from A1 to A2 proficiency. The course aims to develop basic language skills in grammar and phonetics, enrich essential vocabulary for everyday

communication, and enhance comprehension skills. This didactic plan is designed for beginner English language learners, with an emphasis on active and engaging learning through the use of various teaching methodologies and technological tools. The plan is divided into 8 lessons, each lasting 45 minutes, and covers fundamental topics of English grammar and phonetics, such as the alphabet, nouns and articles, the verbs to be and to have as auxiliaries, the verb to have with full meaning, the present simple and present continuous, the negative and interrogative forms of verbs. In the following sections, we will highlight some of the salient points of the course to provide a clearer understanding of its structure and objectives. This course is meticulously crafted to address the needs of absolute beginners in English, providing a robust foundation upon which learners can build their language skills. By integrating both overt and covert teaching methods, the course ensures a balanced approach that supports explicit learning while fostering intuitive understanding. The structured didactic plan, with its focus on fundamental grammar, phonetics, and vocabulary, is designed to engage learners actively and facilitate their progression from the A1 to the A2 level. Through its use of various teaching methodologies and technological tools, the application not only delivers essential language content but also adapts to learners' evolving needs, ensuring that each user can advance at their own pace. Ultimately, this comprehensive and adaptive approach aims to equip beginners with the confidence and competence needed to navigate the English language effectively and continue their learning journey with a solid, self-directed foundation.

However, adhering to a strictly pre-defined syllabus of content, materials, and texts could potentially undermine our intentions. We aimed to create an application that maintains high motivation levels by engaging learners through content that resonates with their interests. The challenge of this approach is compounded by the variability in learner engagement, especially when the model is made available online to a broader audience. While the experimental setting allows for controlled testing, a wider online release introduces the unpredictability of learner engagement. Consequently, we have strived to blend traditional, deductive learning with personalized content, creating a flexible and responsive educational experience. This approach ensures that learners are not only grounded in essential language skills but also remain motivated and engaged through content tailored to their individual interests.

In crafting this course, we have found a strategic compromise by integrating both technical and didactic skills. This innovative approach utilizes advanced programming capabilities alongside proven pedagogical strategies to create a balanced and effective learning environment. By utilizing linked data and dynamic content generation, we ensure that the educational material remains relevant and engaging, while the structured didactic methods guarantee a systematic and coherent learning experience. This fusion of technical prowess and educational theory allows us to offer a flexible and

adaptive course that responds to learners' individual needs, maintaining high levels of motivation and facilitating a more personalized learning journey.

In this section, we will explore each lesson's main topic, and the specific methodologies employed to teach it. The approach to each lesson has been thoughtfully selected to maximize engagement and effectiveness, whether through direct instruction, contextual analysis, or interactive activities. By employing a range of teaching strategies, such as multi-sensory engagement, visual aids, communicative practice, and gradual skill development, the plan aims to provide a comprehensive learning experience that addresses the diverse needs of beginner learners.

The following commentary will provide detailed insights into how each lesson's topic is introduced, taught, and reinforced, highlighting the rationale behind the chosen methodologies and their impact on the language acquisition process.

6.5.1 Methodology and Didactic Approach

The didactic plan employs an integrated approach that effectively combines inductive and deductive methods to cater to diverse learning needs and preferences. The inductive approach facilitates language acquisition by guiding students to discover grammatical rules and linguistic patterns through the examination of contextualized examples. This method promotes critical thinking and deeper understanding, as students derive rules from authentic language use rather than being presented with them in isolation. Conversely, the deductive approach complements this by providing explicit explanations and structured presentations of grammatical rules and concepts. This method ensures that students receive clear and direct instruction, which helps to solidify their understanding and application of language rules.

Lesson	Main Objectives	Methodology	Introductory Phase	Interactive Phase	Consolidation Phase
1	First introduction to the English alphabet. Learn basic vocabulary, including numbers from 1 to 20, days of the week, countries and nationalities, personal information, and common objects.	Multi-sensory approach involving listening	Introducing the alphabet, showing each letter with its pronunciation and an associated example word.	Sound-letter matching exercises. Letter-image matching exercises. Fill-in-the-blank exercises with the initial letter of a word represented by an image. Recording and comparing one's pronunciation with the correct one. Map: nation-nationality association	Phonetic crossword puzzle. Memory game with cards showing letters and their sounds.

2	Know and correctly use the verb <i>to be</i> in the present tense. Learn vocabulary related to common adjectives, adverbs <i>quite</i> and <i>very</i> , and the concept of date.	Deductive approach with explicit presentation of verb forms and grammatical rules.	Explanation of the forms of the verb <i>to be</i> (<i>am, is, are</i>) and their use in affirmative, negative, and interrogative sentences.	Cloze exercise of texts where the verbs <i>to be</i> have been omitted. Interactive flashcards for matching subjects and correct forms of the verb <i>to be</i> .	Creation of dialogues or short texts using the verb <i>to be</i> to describe oneself or others.
3	Nouns and articles in English. Present Simple, positive and negative form. Learn contextual vocabulary related to irregular plurals, jobs, family.	Inductive approach starting from the analysis of an authentic text.	Presentation of nouns (concrete and abstract) and articles (definite and indefinite) with examples of use.	Cloze exercise of a text where articles have been omitted. Interactive flashcards for matching nouns and articles.	Grammar treasure hunt with clues that include nouns and articles. Creation of stories or descriptions using nouns and articles appropriately. Role-playing game where students describe objects or people using nouns and articles.
4	Know and correctly use the verb <i>to have</i> in the present tense. Learn vocabulary related to free time activities, phrases related to love stories, musical genres and musical instruments.	Similar structure to the lesson on <i>to be</i> , combining direct instruction with guided practice and autonomous application.	Explanation of the forms of the verb <i>to have</i> (<i>have</i> and <i>has</i>) and their use in affirmative, negative, and interrogative sentences.	Cloze exercise of texts where the verbs <i>to have</i> have been omitted. Interactive flashcards for matching subjects and correct forms of the verb <i>to have</i> .	Game <i>Guess Who?</i> in which students describe a person using sentences with the verb <i>to have</i> . Game of the goose with questions or sentences to be completed with the verb <i>to have</i> .
5	Past Simple Word formation (suffixes, prefixes) Past time expressions (yesterday, last week, ago). Verbs <i>go, have, get</i> in past simple form. Learn vocabulary related to words and expressions related to daily routines.	Gradual approach, starting from the presentation of the rules for forming the Past Simple and proceeding with cloze exercises with scaffolding.	Explanation of the rules for forming the Past Simple for affirmative, negative, and interrogative sentences.	Cloze exercise with scaffolding, in which students choose between multiple options or complete incomplete sentences. Exercise of transforming affirmative sentences into negative ones. Autonomous consolidation activity in which students complete sentences in the Present Simple without hints.	Description of one's daily routines using the Present Simple. Personal presentations in which students describe their hobbies, interests, and activities using the Present Simple. Creative writing of a short paragraph on a topic of interest using the Present Simple.
6	Learn vocabulary related to houses and furniture, prepositions of place, places in a city. First 5 lessons recap	Direct instruction of grammatical rules, followed by guided practice and consolidation activities.	Explanation of the rules for forming the negative form of verbs, using the structures <i>do not</i> (<i>don't</i>) and <i>does not</i> (<i>doesn't</i>).	Multiple-choice activity in which students select the correct negative form of a verb.	<i>True or False</i> game with sentences in the Present Simple, some affirmative and others negative. Correction of errors in sentences with incorrect negative forms. Debate on a topic of interest, in which students are encouraged to use the negative form of verbs to express disagreement or opposing opinions.
7	Subject and object pronouns in English (Learn vocabulary related to food,	Inductive approach, presenting subject pronouns in meaningful contexts	Presentation of subject pronouns through an animated video or an	Matching exercise: Students will have to match subject pronouns	Listening exercises: Students will listen to short dialogues or descriptions and will

	countable and uncountable nouns, drinks, holidays.	through the use of dialogues and images.	interactive presentation.	to the corresponding images. Sentence Completion: Students will have to complete incomplete sentences by inserting the correct subject and object pronoun. Sentence Creation: Students will have to create original sentences using subject pronouns.	have to identify the subject pronouns used. Writing activities: Students will have to write a short text about themselves or someone they know, using subject pronouns correctly. Interactive quizzes.
8	Understand and correctly use prepositions of place in English Learn vocabulary related to personality adjectives, weather, adventures, common adverbs.	Visual and kinesthetic approach, using images, maps, and physical movements to illustrate the meaning of prepositions of place.	Presentation of prepositions of place through a video or presentation showing images of objects or people in different positions.	Matching Exercise: Students will have to match prepositions of place to the corresponding images. Multiple-choice exercises: Images will be presented, and students will have to choose the correct preposition of place to describe the position of objects or people in the image. Sentence Completion: Students will have to complete incomplete sentences by inserting the correct preposition of place.	Image Description: Students will have to describe images of objects or people in different positions using prepositions of place. Map Creation: Students will have to create a map of a familiar place (home, school, neighborhood) and describe the position of objects or people using prepositions of place. Listening exercises: Students will listen to descriptions of places or positions and will have to identify the prepositions of place used.

Table 14 Appil lessons' plan overview

Central to the teaching methodology is an emphasis on active and engaging learning experiences. Students are encouraged to participate actively in the learning process through a variety of interactive activities designed to foster both engagement and practical application. These activities include role-playing games, which simulate real-life language use, guided discussions that facilitate critical thinking and communicative skills. By combining these methodologies, the didactic plan aims to create a dynamic and supportive learning environment that addresses the needs of all students and enhances their overall language acquisition process.

6.5.1.1 Introductory phase

At the start of each lesson, students engage in an Introductory Phase, which involves a structured presentation of the grammatical or phonetic topic for that session. This phase is designed to provide students with a foundational understanding of the key concepts before delving into more complex applications.

During this phase, students are introduced to new material through a combination of multimedia resources, interactive presentations, and concise explanations. These resources might include visual aids, audio recordings, and sample sentences that illustrate the core principles of the topic. The goal

is to present the essential elements in a clear and accessible manner, allowing students to grasp the fundamental concepts independently.

To accommodate different learning styles and paces, the Introductory Phase is designed to be self-paced. Students can interact with the materials at their own speed, revisit explanations as needed and explore examples to build their understanding. This flexibility supports individual learning needs and encourages students to take ownership of their learning process.

Additionally, the Introductory Phase may include initial formative assessments, such as quick quizzes or interactive exercises, to gauge students' preliminary understanding and provide immediate feedback. This allows for the identification of any area needing further clarification before moving on to more complex activities.

Overall, the *Introductory Phase* serves as a critical foundation, setting the stage for deeper exploration and practice in subsequent parts of the lesson. By equipping students with a solid grasp of the topic from the outset, this phase enhances their readiness for active participation and application of the new concepts.

6.5.1.2 Interactive phase:

In the Interactive Phase, students apply the knowledge they have gained through engaging exercises available on the online platform. This phase is crucial for reinforcing learning and providing practical experience with the new grammatical or phonetic concepts introduced in the lesson.

Students participate in a variety of interactive exercises, including cloze completion tasks, word-image matching activities, and flashcard drills. These exercises are designed to offer immediate and personalized feedback, which supports active learning and enables students to self-correct and refine their understanding in real-time. The immediate feedback mechanism helps students to recognize and address errors promptly, fostering a more effective learning process.

Each lesson is equipped with its own set of innovative exercises and games tailored to the specific content covered. These include interactive activities such as drag-and-drop exercises, matching games, and simulation tasks that make learning more dynamic and enjoyable. The incorporation of such games aims to maintain high levels of motivation and engagement, despite the potentially challenging nature of the numerous grammatical notions being introduced.

By providing a variety of interactive and game-based learning tools, the Interactive Phase not only reinforces the material but also keeps the learning experience stimulating and motivating. This approach ensures that students remain engaged and enthusiastic about acquiring new language skills, facilitating a more enjoyable and productive learning journey.

6.5.1.3 Consolidation phase

In the consolidation Phase, students deepen their understanding and solidify their learning through a range of oral and written production activities. These tasks include creating dialogues, describing images, and writing short texts, which encourage students to apply their knowledge creatively and engage in authentic communication. By actively using the language in context, students reinforce their grasp of the grammatical and phonetic concepts covered in the lesson.

The online platform plays a pivotal role in this phase by providing a suite of resources and interactive tools that support the consolidation process. The learning platform offers features such as interactive writing prompts, voice recording tools, and peer review mechanisms that facilitate productive practice and feedback. Immediate feedback on these production activities allows students to assess their performance, make necessary corrections, and refine their language skills.

Furthermore, the platform enhances the learning experience by fostering communication and collaboration between students and teachers. Discussion forums, group projects, and real-time messaging options enable students to seek clarification, share ideas, and receive guidance from their instructors. This connectivity contributes to a stimulating and supportive learning environment, where students can engage more fully with the material and benefit from collaborative learning experiences. Overall, the Consolidation Phase, supported by the robust features of the online platform, ensures that students not only practice but also internalize and apply their language skills effectively, setting the stage for continued improvement and proficiency.

6.5.2 Lessons' plan

The lesson plans outlined in this course are designed to guide learners through a comprehensive and structured journey of English language acquisition. Each lesson is thoughtfully crafted to build on foundational knowledge, progressively introducing new concepts and skills. Our approach combines the best of traditional language learning with the flexibility of personalized content, ensuring that learners receive both a systematic framework and the opportunity to explore topics that interest them. The lessons are structured to cater to the needs of absolute beginners, starting with essential elements such as phonetics, the alphabet, and basic grammar. As learners advance, they will encounter progressively more complex language structures, including verb usage, sentence formation, and vocabulary acquisition. Throughout each lesson, students will engage in active learning through interactive exercises and guided practice, ensuring they can apply what they've learned in practical contexts.

In addition to the core content, the course offers flexibility through the Content Management System (CMS), which allows learners to select materials that align with their personal interests. This dynamic

approach ensures that the learning experience remains engaging, relevant, and motivating. Ultimately, these lesson plans aim to equip learners with the foundational language skills necessary to confidently progress through more advanced stages of English proficiency.

6.5.3 Lessons overview

The instructional design developed in this project represents an integrated and sequential pathway for the acquisition of English as a second language. The complete set of lessons, covering topics from phonetics and basic vocabulary to core grammatical structures and functional communication, was conceived as a unified curriculum grounded in Second Language Acquisition (SLA) theory, cognitive psychology, and contemporary pedagogical practice. Rather than functioning as isolated modules, each lesson contributes to a cumulative process of language development, in which linguistic forms, lexical items, and communicative functions are revisited, reinforced, and extended through progressive stages of complexity.

The sequence begins with the introduction of English phonetics and the alphabet, where students become familiar with sound–symbol correspondence and basic vocabulary items such as numbers, days of the week, countries, and common objects. This stage relies on a multisensory approach combining auditory and visual input to strengthen phonological awareness and orthographic recognition. Repetition, sound-letter matching, and contextual listening tasks promote memorization and accurate pronunciation. Learners engage with interactive digital exercises that integrate sound files, images, and feedback mechanisms, allowing them to self-correct and refine their perception of English phonemes. These activities draw on Dual Coding Theory and Krashen’s Input Hypothesis, which together emphasize the cognitive benefits of combining verbal and non-verbal information and of providing comprehensible input slightly beyond the learner’s current level ($i+1$).

Building upon this foundation, subsequent lessons introduce core grammatical structures, beginning with the verbs to be and to have, and later extending to nouns and articles, the Present and Past Simple tenses, and the negative form of verbs. Grammar instruction follows an inductive and communicative model: learners encounter examples in meaningful contexts before generalizing the underlying rules. Through guided discovery, students analyze authentic or semi-authentic sentences, complete cloze and transformation exercises, and then apply the forms in controlled and free production tasks. This methodology reflects constructivist principles and discovery learning theory, encouraging active engagement and deeper cognitive processing.

Each grammar topic is systematically reinforced through thematic vocabulary domains, daily routines, jobs, family, housing, food, holidays, and city life, ensuring that grammatical competence evolves alongside lexical and functional proficiency. This integration follows the Lexical Approach,

in which grammar and vocabulary are treated as interdependent aspects of communicative competence. Learners practice forming sentences, role-playing dialogues, and producing short written texts or oral presentations that situate new structures within realistic communicative settings. Interactive and gamified activities play a central role throughout the curriculum. Exercises such as pronunciation recording, matching games, treasure hunts, memory cards, and board games like Guess Who? or Game of the Goose transform repetition into dynamic engagement. These activities are supported by Self-Determination Theory, which identifies autonomy, competence, and relatedness as key motivational drivers. Immediate feedback from the online platform enables learners to monitor progress, foster self-regulation and intrinsic motivation.

Vygotsky's view of the Zone of Proximal Development (ZPD), where knowledge is co-constructed through dialogue and shared experience (Vygotsky, 1978), underpins the pedagogical design of the lesson activities. Cultural and intercultural dimensions are intentionally embedded within the lesson content. Topics such as countries and nationalities, global festivals, and social relationships provide opportunities to discuss cultural diversity and foster Intercultural Communicative Competence (ICC) (Byram, 2021). By associating language learning with real-world cultural contexts, learners not only acquire vocabulary and grammar but also develop awareness of cultural conventions, empathy, and communicative appropriateness in cross-cultural situations. The design of the activities also integrates principles from cognitive psychology related to attention, memory, and cognitive load management. Tasks are sequenced from simpler recognition exercises to more complex production activities, minimizing overload and ensuring gradual knowledge consolidation. Scaffolding is embedded at all stages: explicit guidance is provided during the introduction of new material, then gradually withdrawn as learners gain confidence and autonomy. Consolidation phases typically include reinforcement games, creative writing, debates, and reflection tasks that strengthen long-term retention through spaced repetition and meaningful use.

From a methodological standpoint, the curriculum combines direct instruction, guided practice, and autonomous application within a blended-learning model that exploits the affordances of digital technology. Interactive exercises, multimedia presentations, and online quizzes provide multimodal exposure to input and facilitate real-time correction. The learning platform's digital tools are designed to enhance learner autonomy, allowing individual pacing and adaptive feedback. By engaging multiple sensory channels, visual, auditory, and kinesthetic, the lessons cater to diverse learning styles and ensure inclusivity.

Across all lessons, grammar and vocabulary are presented as tools for meaningful communication rather than isolated knowledge. Activities that require learners to describe images, narrate past experiences, or debate everyday topics link linguistic form with pragmatic function. This

communicative orientation reflects the principles of Communicative Language Teaching (CLT) and the Task-Based Learning (TBL) approach, which view interaction as both the means and the goal of language learning.

The curriculum also integrates metacognitive elements by encouraging reflection on errors and learning strategies. Error-correction activities prompt students to identify and amend mistakes, cultivating analytical awareness and promoting learner independence. Creative tasks, such as story writing and personal presentations, allow students to transfer knowledge from controlled drills to expressive language use, reinforcing both confidence and competence.

Overall, the pedagogical framework represents a synthesis of linguistic theory, cognitive science, and technology-enhanced instruction. Each component, phonetic training, vocabulary expansion, grammar practice, cultural exploration, and interactive consolidation, contributes to a cohesive progression from foundational literacy to communicative fluency. The combination of inductive learning, scaffolding, and gamified interaction ensures that learners are not passive recipients of information but active participants in constructing linguistic knowledge.

By aligning instructional design with established SLA principles and embedding it within a flexible digital environment, the lesson series fosters sustained motivation, learner autonomy, and cognitive engagement. It transforms language instruction from a static transmission of rules into an adaptive, participatory, and culturally rich learning experience, one that equips learners not only with grammatical accuracy and lexical range but also with the communicative competence necessary to navigate diverse linguistic and social contexts.

6.5.3.1 A hybrid implant: expectations from the Semi-Traditional Course Component

The semi-traditional course component of our web application is designed to ensure a structured and effective learning experience for absolute beginners. This component provides a clear, systematic progression through essential language concepts, starting with foundational elements such as phonetics, the alphabet, and basic grammar. Each lesson builds upon the previous one, incorporating explicit instruction and structured presentations of key topics. Learners engage with a series of progressively challenging lessons, each focused on a specific aspect of the English language, such as verb usage, sentence structures, and vocabulary acquisition. The semi-traditional format includes periodic assessments and reviews to monitor progress, reinforce learning, and address any gaps in understanding. This approach ensures that learners receive a solid grounding in fundamental language

skills, with clear goals and measurable outcomes. Additionally, the course incorporates interactive elements and guided practice to facilitate active learning and application of concepts.

The integration of inductive and semi-traditional elements in our web application addresses the need for continuous, personalized support in language learning. This approach functions similarly to an ever-present tutor, offering constant guidance and feedback. The semi-traditional component provides structured instruction and systematic progression, akin to a tutor's direct teaching, ensuring that learners acquire foundational knowledge and skills in a clear and organized manner. Meanwhile, the inductive, free learning course allows learners to explore language concepts independently, much like a tutor encouraging exploration and self-discovery. By combining these aspects, the application creates a supportive environment where learners receive both the structure and flexibility they need. The continual scaffolding offered through this method ensures that learners are not left without guidance, even when engaging in self-directed study. This dual functionality bridges the gap between structured lessons and independent learning, providing a seamless educational experience that mimics the presence of a tutor while enabling learners to progress at their own pace and revisit concepts as needed.

Our approach to language learning is distinguished by its unique integration of inductive, free learning experiences with a scaffolded semi-traditional didactic plan. Unlike conventional methods that may rely solely on structured instruction or purely exploratory learning, our approach combines these elements to offer a more holistic educational experience. The inductive, free learning course component allows learners to engage in self-directed exploration of the language, encouraging them to discover grammatical rules and patterns through authentic language use and contextual examples. This method promotes deeper understanding and active problem-solving skills. Complementing this, the semi-traditional didactic plan provides a structured framework that supports learners by offering explicit instruction, guided practice, and systematic progression through essential language concepts. This dual approach ensures that learners benefit from the flexibility and autonomy of free learning while receiving the necessary guidance and reinforcement through structured lessons. The result is a comprehensive and innovative learning environment that caters to diverse learning needs, facilitates deeper engagement with the language, and supports learners in achieving well-rounded proficiency in English.

Additionally, the flexibility to not only choose exercises but also dynamically generate them is a product of extensive research and integration across several domains. This capability is achieved through a sophisticated blend of linked data technologies, programming expertise, and pedagogical principles. Linked data provides a rich repository of interconnected information, allowing the application to offer a diverse range of authentic and relevant texts. The programming infrastructure

ensures that these texts can be efficiently selected and customized according to individual learner needs, while also supporting the creation of tailored exercises that align with specific learning objectives. The integration of didactic strategies ensures that these dynamically generated exercises are pedagogically sound, promoting effective learning through structured yet flexible activities. Behavioral insights guide the design of these exercises to engage learners in a manner that aligns with their individual preferences and learning styles, thereby enhancing motivation and retention. This cross-disciplinary approach enriches the learning experience and ensures that it remains adaptable, responsive, and aligned with both educational standards and learner needs. By combining these elements, the application successfully bridges the gap between structured instruction and personalized learning, offering a dynamic and engaging educational environment.

Having outlined the expectations and benefits of the semi-traditional component of our web application, we now turn our attention to the innovative area of the course. This section is designed to offer learners a dynamic and flexible learning experience that complements the structured foundation provided earlier. The innovative component incorporates modern, interactive elements that support self-directed exploration and engagement with the language. This aspect of the course is integral to our approach, as it fosters autonomy and encourages learners to apply their skills in diverse contexts. In the following sections of the manuscript, we will delve into the specifics of this innovative area, detailing its unique features and methodologies. Additionally, we will present the results and impact of this component, demonstrating how it enhances the overall learning experience and contributes to the effectiveness of the course.

One of the key features of our course is the ability for learners to choose the content of each lesson according to their preferences. This flexibility ensures that learners remain engaged and motivated by studying topics that are relevant and interesting to them. The input texts for each lesson can be drawn from a vast library of materials, spanning a wide array of genres, themes, and topics. Whether learners are interested in current events, literature, pop culture, or specific fields of study, they can select texts that match their individual tastes and goals. This personalized approach not only keeps the learning process enjoyable but also makes it more meaningful, as learners can connect language learning to their personal interests and real-world applications. By incorporating this level of customization, the course fosters a deeper, more immersive learning experience that enhances both motivation and retention.

Despite the need for a traditional phase in the course, which forms the backbone of the learning platform's structure, we have successfully integrated a strategy through the Content Management System (CMS) that allows for the delivery of user-chosen content. This means that while learners progress through the structured and foundational elements of the course, they can still tailor their

learning experience by selecting content that aligns with their interests. The CMS enables the dynamic delivery of personalized materials, ensuring that learners can engage with texts and topics of their choice within the larger framework of the course. This integration strikes a balance between maintaining a systematic, traditional learning path and providing the flexibility to explore and apply language concepts in ways that resonate with each learner's unique preferences.

6.6 Participant Response: Interpreting Feedback in Light of Access Patterns

While this chapter offers a comprehensive discussion of user feedback and system usability, the analysis is focused strictly on the functional validation of the platform. Because the evaluation phase encountered logistical and technical hurdles, the planned administration of final questionnaires and post-tests could not be completed. Consequently, the study relies on internal usage metrics specifically implemented within the platform to generate usage scores and performance assessments. These metrics provide a direct window into user behavior and system interaction, though they primarily reflect perceived effectiveness and functional engagement rather than traditional summative learning gains.

By explicitly noting the lack of direct measures for learning gains, we aim to strengthen the methodological transparency of the study and prevent any overinterpretation of the results. These findings primarily reflect user experience and the system's perceived effectiveness, rather than a measurable impact on language acquisition. A more detailed account of these scope constraints and the data collection process is provided in

The learning platform was deployed on a dedicated server administered exclusively by the researcher and remained operational for the entire duration of the program, from June 21, 2024, to May 22, 2025. The server was continuously monitored and maintained to ensure stability, data integrity, and uninterrupted access for all participants. Regular maintenance activities included performance optimization, security patching, database backups, and log monitoring to prevent service disruptions. Over the course of the deployment, the application evolved through several updated versions, each incorporating incremental improvements based on technical feedback and user experience reports. These updates addressed minor interface issues, optimized content delivery, and introduced additional features designed to enhance learner engagement and accessibility.

During this period, the complete sequence of eight instructional lessons was progressively released through the learning platform, following a carefully staged schedule aligned with the program's pedagogical timeline. After each lesson delivery, participants were invited to provide structured

feedback through online questionnaires and informal comments. This feedback loop was instrumental in refining both the technical and instructional components of the system. The researcher maintained a hands-on role throughout the entire implementation phase, remaining available to participants for real-time assistance with login difficulties, navigation problems, or content-related inquiries. This continuous presence not only ensured technical reliability but also fostered a supportive learning environment, reinforcing participants' confidence in using the digital learning platform and sustaining engagement across the program's duration. Participant response to the digital learning program was continuously monitored using both qualitative and quantitative tools. These included structured weekly feedback forms, mid-point and final reflective questionnaires, and real-time engagement data collected by the learning platform. The multi-source evaluation approach provided a nuanced understanding of how learners interacted with the program, what they valued, and where challenges emerged.

In general, participant feedback was highly positive during the early phases of the program. Many learners expressed enthusiasm for the clarity and structure of the content, especially highlighting how the platform supported their day-to-day responsibilities. The flexibility of the asynchronous format was frequently cited as a major advantage, allowing participants to engage with the material at their own pace and according to their personal schedules. The modular structure was also well received, as it allowed learners to focus on specific topics in a targeted and manageable way. The incorporation of real-life administrative scenarios and case-based examples was viewed as particularly relevant and useful for translating language skills into practical workplace applications.

6.6.1 Usage metrics

Participant engagement with the digital learning platform was assessed using a blend of qualitative instruments, structured feedback forms, reflective surveys, and quantitative methods, including real-time analytics and server access logs. This multi-source strategy enabled triangulation between self-reported experiences and behavioral indicators observed in log data.

This study adopts a structured methodology to analyze web server access logs in order to uncover patterns of user interaction, with a specific focus on dynamic and interactive web pages such as login and educational content. The log file used for analysis, named `access.log`, contains HTTP request records generated by the web server. Each log entry includes the client's IP address, timestamp, HTTP method, requested resource path, protocol, response status code, and response size.

The initial stage of the analysis involved parsing the raw log file into a structured format using regular expressions. Each valid line was decomposed into its constituent elements and stored in a Pandas DataFrame for further processing. The datetime values were converted into standardized, timezone-

neutral Python datetime objects to ensure consistency in temporal analysis, and any malformed entries were discarded.

To focus the analysis on meaningful user behavior, all requests for static content such as images, scripts, style sheets, and fonts were filtered out. These resources are typically requested automatically by the browser and do not reflect intentional user actions. This filtering was implemented by excluding requests based on known file extensions and URL path prefixes associated with static assets.

With a cleaned dataset in place, general descriptive statistics were computed. These included the total number of requests and the number of unique IP addresses observed during the analysis period. The data was then aggregated by IP address to examine the number of requests per user and the timespan between their first and last visits. In parallel, aggregation by requested path was performed to determine which resources were most frequently accessed, how many unique users visited each path, and which HTTP status codes were most commonly returned for each resource.

A more focused analysis was conducted on dynamic user-facing pages by further excluding static content. For these user-oriented endpoints, additional metrics were calculated, such as the number of visits per path, the number of unique visitors, and the average number of visits per unique IP. These figures provide insight into user engagement and the frequency of repeated interactions with specific pages. Furthermore, the distribution of HTTP status codes per path was analyzed to detect potential errors or redirection issues affecting user access.

Special attention was given to endpoints related to login functionality and educational content, specifically those paths containing the keywords ‘login’ or ‘lesson.’ These paths were isolated for detailed examination to better understand usage patterns around authentication and resource consumption. For each of these endpoints, the number of visits, the number of unique IP addresses, and the average visits per user were calculated and ranked to highlight their relative popularity and intensity of use.

Finally, all results were exported to a multi-sheet Excel file to support reproducibility and facilitate further interpretation. The spreadsheet includes sheets dedicated to visit counts by IP, access times, overall statistics, detailed path-level metrics, and focused data on user-centric and key functional pages. This structured output serves as a comprehensive report that encapsulates both high-level trends and granular behavioral insights extracted from the log data.

Early qualitative feedback highlighted high levels of enthusiasm, particularly regarding the clarity, modularity, and asynchronous design of the program. Participants appreciated the flexibility to engage with material according to their schedules, citing the platform’s alignment with their professional responsibilities. Many respondents noted that the real-world administrative scenarios

and language tasks were particularly applicable, helping them translate communication theory into workplace action.

These reflections were mirrored in the web access logs. During the first 6–8 weeks, daily HTTP request data revealed a consistent pattern of high engagement, with most participants accessing content multiple times per week.

Several learners also mentioned that the program served as a catalyst for professional reflection and skill refinement, especially in areas like email composition and report writing. For some, it functioned as a refresher; for others, it was an opportunity to adopt new strategies. These subjective reports were supported by high interaction volumes with grammar review modules and vocabulary exercises, which appeared frequently in the request logs under content endpoints labeled `/lesson/grammar` and `/lesson/vocab`.

Despite the overall positive reception, later-stage feedback identified challenges in sustaining motivation and maintaining consistent engagement. Participants managing full-time workloads reported difficulty balancing learning with professional obligations. This was corroborated by a gradual tapering of request frequency across the second half of the program, as captured in the log dataset. Login counts and content access declined steadily, transitioning from routine weekly activity to more deadline-driven spikes surrounding assessments.

Digital literacy gaps also emerged as a barrier, particularly during the onboarding phase. While most learners adapted quickly to the platform, access log analysis showed disproportionately high access rates to help-related pages and static documentation (e.g., `/support/faq`, `/guide/onboarding`) during the first two weeks. These requests were predominantly generated by a subset of IP addresses associated with lower overall content engagement, indicating users needing additional assistance.

Taken together, these insights suggest that future program iterations would benefit from enhanced scaffolding, both instructional and technical. Early-stage friction could be alleviated with onboarding tutorials and just-in-time support tools, while sustained engagement might be improved through adaptive pacing mechanisms, such as dynamic content release or optional synchronous check-ins.

6.6.2 Usage Metrics: Behavioral Insights from Web Server Logs

An analysis of platform usage data collected over a 334-day period, spanning from June 21, 2024, to May 22, 2025, provides a nuanced perspective on patterns of user behavior, system engagement, and potential implications for instructional design. Across this timeframe, the platform registered a cumulative total of 92,959 HTTP requests originating from 5,670 unique IP addresses. These interactions correspond to an estimated 200 hours of effective usage time. Taken together, these

figures suggest that while the platform does not experience intensive or high-volume utilization, it nonetheless sustains a moderate and consistent level of engagement across its user base.

When disaggregated by endpoint, the data reveal a pronounced concentration of activity around a relatively small number of access points. The login page, for example, alone accounted for more than 12,000 requests, underscoring its role as a critical entry node within the system. Beyond authentication, the highest levels of activity were directed toward the root homepage and the introductory lesson module (/lesson1), which collectively recorded more than 2,100 visits and attracted over 1,000 unique users. This pattern of engagement suggests that users are particularly oriented toward initial access and entry-level instructional content, while subsequent or more advanced modules may receive comparatively less attention.

From an instructional design standpoint, these findings raise several noteworthy considerations. The strong emphasis on preliminary components implies that users may disproportionately engage with the earliest stages of the learning sequence, either as a point of orientation or as a threshold beyond which fewer learners progress. Such a pattern is consistent with broader trends observed in digital learning environments, where attrition often occurs after introductory exposure. Accordingly, the results underscore the importance of designing initial learning experiences that are not only accessible and inviting but also scaffold learners effectively toward sustained engagement with more advanced instructional materials.

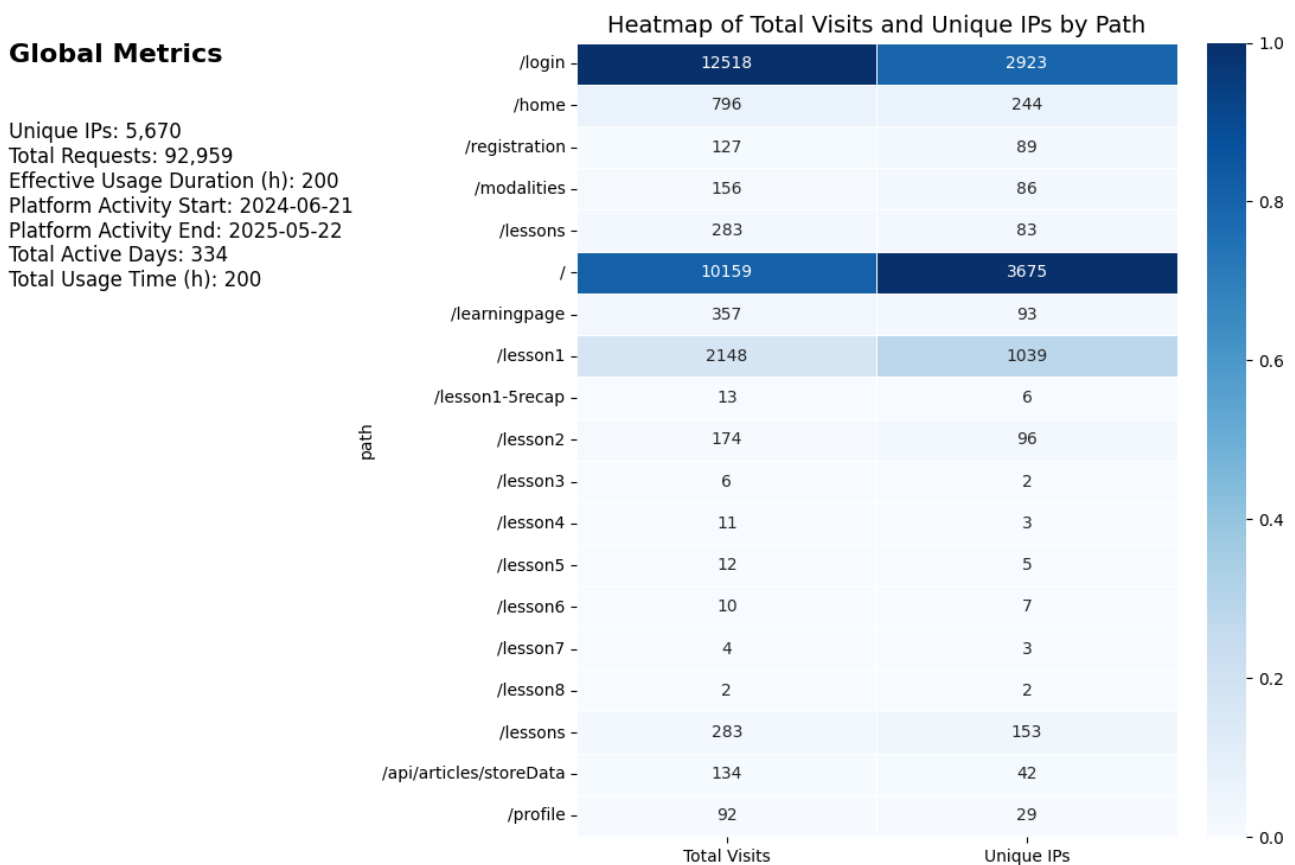


Figure 49 11-month usage-metrics heatmap

The usage data further reveal a pronounced decline in user interaction beyond the first instructional module. Whereas the introductory lesson (/lesson1) attracted substantial engagement, successive modules, particularly those positioned deeper within the instructional sequence (e.g., /lesson3, /lesson8), exhibited sharply reduced visit counts, in some cases registering only single-digit interactions. This tapering effect is characteristic of many online learning environments, particularly those that rely heavily on asynchronous progression and learner self-regulation. Such attrition may be attributable to a variety of factors, including diminishing user motivation, limited perceived relevance of the content, or challenges related to the accessibility and usability of the platform. While the introductory units appear to be effective in capturing initial attention, the subsequent decline suggests potential misalignment in the design, pacing, or scaffolding of later materials, thereby raising important questions about the overall coherence of the learning pathway.

From a temporal perspective, the aggregate usage of approximately 200 hours distributed across nearly one calendar year underscores a pattern of brief and intermittent interaction rather than sustained, daily engagement. On average, this equates to less than one hour of activity per day across the user base, suggesting that the platform may be functioning more as a resource for short, task-oriented sessions than as a space for extended, immersive learning. This observation has significant

implications for instructional design, particularly in relation to structuring content that encourages persistence and supports deeper, long-form engagement.

Additional insight can be drawn from the activity logs associated with auxiliary endpoints. Notably, the `/api/articles/storeData` route was triggered 134 times during the observation period, representing article processing events. When considered against the defined active user pool of approximately 20 individuals, this figure corresponds to an average of 6 to 7 articles engaged with per user. While this level of interaction demonstrates that users are indeed accessing and processing written materials, the overall frequency appears moderate. This raises further questions regarding the discoverability and relevance of article-based resources, as well as the potential cognitive demands associated with engaging with text-heavy instructional modalities. In combination with the declining engagement observed across successive lesson modules, these findings suggest a need for closer examination of the platform's capacity to sustain learner motivation and promote consistent interaction with diverse instructional formats over time.

In conclusion, these quantitative indicators provide a foundational basis for further inquiry into the quality of user experience and the effectiveness of pedagogical strategies implemented on the platform. The data reveal strong initial engagement but also hint at potential obstacles to sustained participation and deep learning. To address these issues, future efforts might explore the implementation of adaptive pathways, gamification mechanisms, or automated feedback loops designed to maintain user motivation. Complementary qualitative methods, such as surveys or structured interviews, would also be beneficial to better understand the motivations and barriers shaping user behavior, thus enabling a more holistic evaluation of the learning environment.

A notable development during the observation period has been the significant expansion of the platform's lexical database. The pool of known nouns has increased from 13,809 to 20,789, reflecting a substantial enhancement in the breadth of linguistic resources available to users. For each newly incorporated noun, the platform has systematically sought to retrieve and associate relevant image information and supplementary descriptive content whenever possible. This process has not only broadened the lexical coverage of the platform but has also enriched the quality of individual entries by providing multimodal resources, such as illustrative images and extended descriptions, that support more comprehensive understanding and engagement. The expanded database thus offers users access to a deeper and more visually supported lexicon, facilitating richer exploration and learning opportunities across the system.

In addition to general usage metrics, the platform's collaborative processing features yield further insights into content creation and learner engagement. Over the 334-day observation period, users collectively processed 150 texts through the system's text enhancement tools, as captured in the

dedicated processed table. This collaborative activity resulted in the automatic generation of 2134 flashcards; each derived from user-contributed abstracts and enriched content. These flashcards serve as reusable learning objects, supporting spaced repetition and personalized study strategies within the platform. The steady rate of text processing and flashcard production demonstrates an active, albeit moderate, participation in the platform's knowledge-building functions. This not only highlights the value of user-generated content but also indicates a level of investment in the co-construction of learning resources, which is foundational for fostering both retention and deeper engagement.

6.6.3 Lexical Distribution and Transparency: Interpretation of Similarity on Frequency Metrics

The lexical component of the platform was originally designed with a strong emphasis on frequency-driven exposure, drawing on established assumptions that repeated encounters with high-frequency vocabulary support efficient acquisition (Ellis, 2002; Nation, 2013). In line with this approach, lexical selection and progression were implemented in a loop-based manner, privileging quantity of exposure and iterative reinforcement across learning cycles. The underlying assumption was that frequent lexical items would constitute an accessible entry point for Italian learners of English, and that extensive interaction with such items would facilitate meaning construction through repeated guessing and contextual inference (Grabe & Stoller, 2020). However, analysis of participant interactions revealed that the effectiveness of this strategy was not primarily determined by frequency alone. Instead, user behavior and processing patterns indicated that lexical transparency played a more decisive role in comprehension and engagement, consistent with findings on cross-linguistic influence in lexical processing (De Groot & Keijzer, 2000; Pavlenko, 2008). In practice, the focus on a very large set of extremely high-frequency words resulted in increased reliance on guessing strategies, as many of these items offered limited formal or semantic cues for Italian learners. This shift in empirical evidence prompted a re-evaluation of the initial frequency-centered design, highlighting the need to balance lexical inundation with cross-linguistic transparency considerations. An additional insight emerging from the fieldwork concerns patterns of lexical development and cross-linguistic distribution between Italian and English. Server-side metrics indicate that user interaction and system processing were concentrated predominantly on lexical items within two distinct frequency ranges: high-frequency words in the 1–3,500 band and low-frequency words above approximately 7,500, reflecting commonly adopted frequency thresholds in lexical research (Nation, 2006). Mid-frequency vocabulary, which is often assumed to play a central role in progressive language acquisition, was comparatively underrepresented. When these findings were examined in relation to the custom readability index developed for the platform and the results of word-similarity

analyses, a notable paradox emerged. Lexical items traditionally classified as part of an “easy” beginner lexicon for Italian learners of English, primarily high-frequency words, were in fact associated with greater processing difficulty. This effect appears to be linked to limited lexical transparency between Italian and English at lower frequency bands, where formal similarity and cognate relationships are sparse (De Groot & Keijzer, 2000). Conversely, less frequent lexical items above the 7,500 range often exhibited higher degrees of morphological or semantic transparency, facilitating comprehension despite their lower corpus frequency. These findings suggest that frequency-based notions of lexical difficulty may be insufficient when applied to closely examined bilingual contexts, and that cross-linguistic transparency plays a critical role in shaping perceived and actual lexical complexity (Ellis & Larsen-Freeman, 2006). From a pedagogical perspective, this underscores the importance of integrating transparency-aware lexical metrics into instructional design and adaptive sequencing strategies.

To further investigate the relationship between lexical frequency, user progression, and cross-linguistic transparency, the annotated lexical dataset was subjected to a two-dimensional exploratory analysis. Specifically, we conducted a comparative examination between the custom readability index introduced in Assessing Text Complexity (4.3.15) and the user interaction metrics collected during platform usage. This approach allowed us to explore how text difficulty, as predicted by readability and lexical transparency measures, aligned with actual learner behavior and developmental progression. The analysis was designed to highlight whether the platform’s frequency-based design corresponded to observable user engagement patterns, and to identify whether transparency and similarity between Italian and English lexemes better explained processing ease or difficulty.

A heatmap visualization was constructed to examine how mean lexical similarity scores (*combined score* as reported in Table 11) distribute across Nation frequency bands and levels of the *Vocabolario di Base* (VDB), which represents a pedagogically motivated classification of core vocabulary commonly introduced in early stages of language learning. This representation provides a compact overview of how lexical transparency behaves at the intersection of corpus-based frequency and pedagogical lexical centrality. Rather than treating frequency and pedagogical status as independent predictors of lexical difficulty, the heatmap makes it possible to observe their combined effects and to identify non-linear patterns that may not be evident through unidimensional analyses. As such, the visualization serves as an empirical bridge between the platform’s original frequency-driven design assumptions and the usage-based evidence emerging from learner interaction data.

Heatmap of Lexical Similarity Across Frequency and Developmental Levels

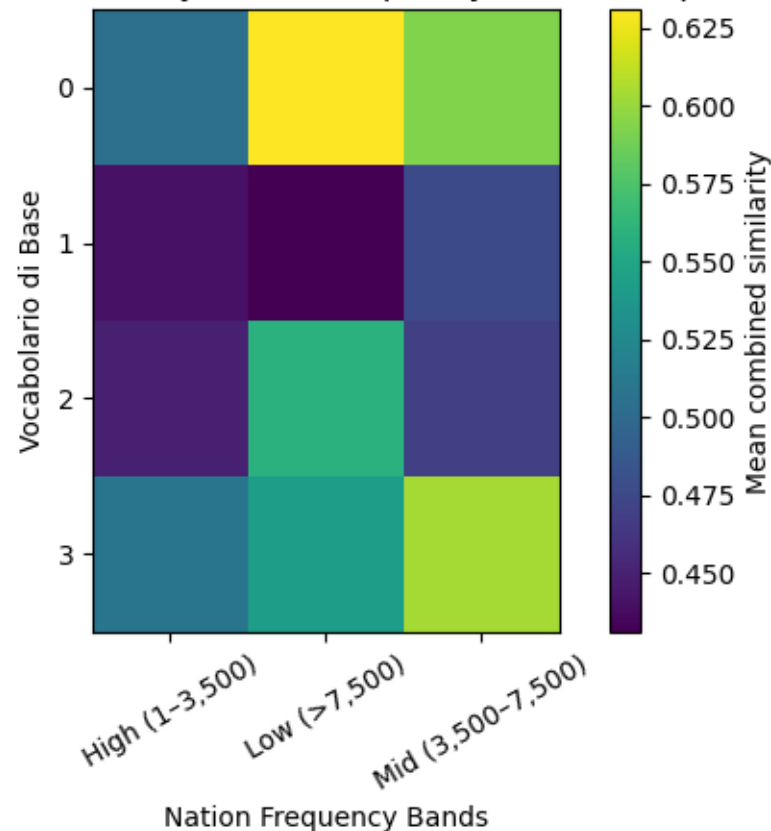


Figure 50 The heatmap provides an exploratory visualization of how lexical similarity behaves across frequency and developmental dimensions within the annotated lexicon. While the analysis is descriptive in nature, the observed patterns suggest systematic interactions between frequency-based and transparency-based measures of lexical complexity.

6.6.3.1 Overview of the dataset and methodological framing

The analysis of the lexical component of the platform, based on 4,500 English lexical items processed during fieldwork, reveals a consistent pattern of interaction that is heavily shaped by frequency distribution and cross-linguistic transparency. The similarity on frequency contingency table provides a quantitative overview of how learner interactions are distributed across three frequency bands (high, mid, and low) and three similarity bands (low, medium, high). In this context, frequency is operationalized using the Nation lexical levels (Nation, 2006), while similarity is measured by the combined score produced by the platform's algorithmic assessment of morphological and phonological resemblance between Italian and English word pairs.

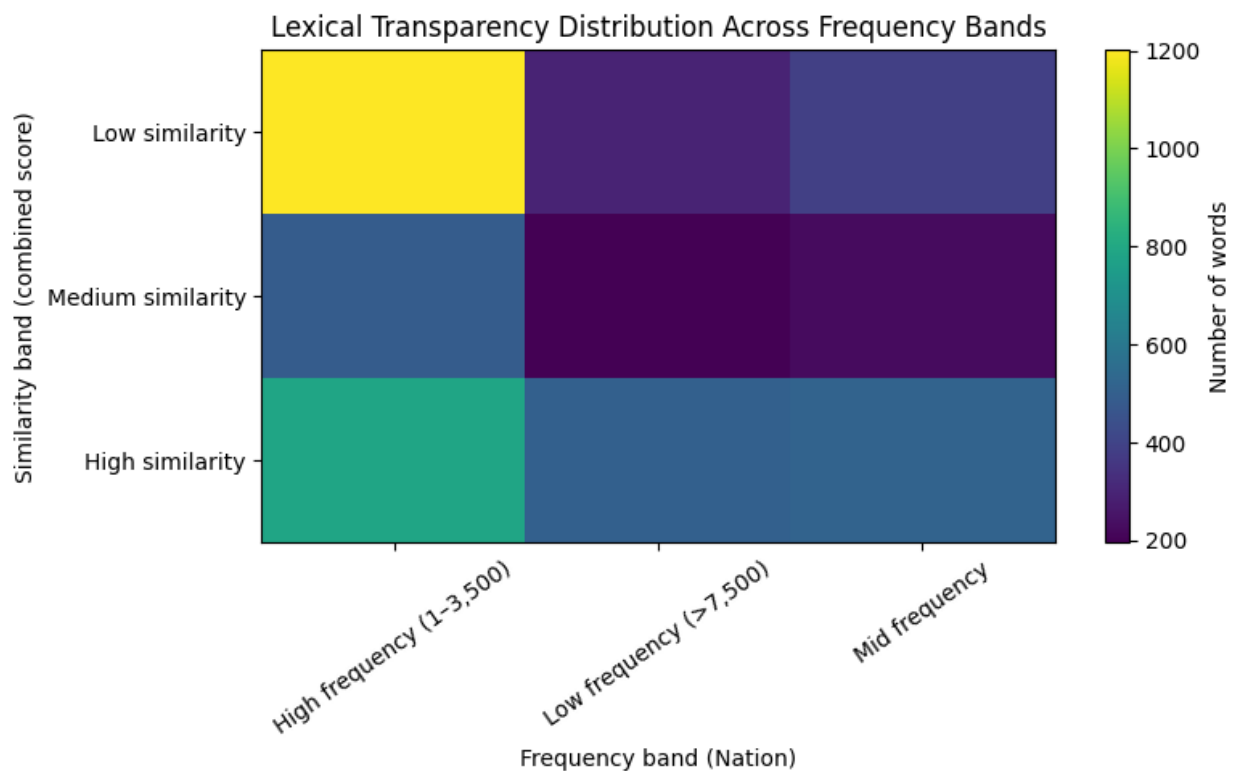


Figure 51 The heatmap reports the distribution of Italian–English word pairs according to their combined similarity score and their frequency band (Nation levels). The frequency bands are divided into high frequency (1–3,500), mid frequency (3,501–7,500), and low frequency (>7,500). Similarity is categorized into low, medium, and high based on the combined transparency score. The results show that high-frequency vocabulary accounts for the largest portion of the dataset, with 1201 items classified as low similarity, 489 as medium similarity, and 786 as high similarity. This indicates that although high-frequency words are abundant, a substantial subset lacks cross-linguistic transparency, which may contribute to processing difficulty and increased reliance on guessing strategies. Mid-frequency vocabulary is less represented, suggesting that learners interact less with items that are typically assumed to support progressive lexical development. In the low-frequency band, the proportion of high-similarity items (506) is comparable to that of low-similarity items (298), suggesting that transparency can compensate for low corpus frequency. Overall, the heatmap highlights that frequency alone is an insufficient predictor of lexical accessibility for Italian learners of English, and that transparency should be integrated into adaptive sequencing and instructional design.

The results show a strong predominance of high-frequency vocabulary (Nation level 1) and, within this band, a surprisingly high proportion of low-similarity items. Conversely, low-frequency vocabulary includes a substantial number of high-similarity items, which suggests that cross-linguistic transparency may compensate for frequency in shaping perceived difficulty and learner engagement. These findings have important implications for both the theoretical understanding of bilingual lexical processing and the practical design of adaptive learning systems.

6.6.3.2 High-frequency vocabulary: quantity does not imply accessibility

The high-frequency band (1–3,500) is the most represented in the dataset, with a total of 2,476 lexical items processed by learners. This confirms the platform's original design assumption that high-frequency words would naturally attract the greatest interaction, as they are typically considered foundational for comprehension and reading (Nation, 2006; Ellis, 2002). High-frequency vocabulary is also widely recognized in second language acquisition research as the most effective entry point for learners because of its repeated occurrence in authentic language use, which facilitates implicit learning and automaticity (Ellis, 2002; Nation, 2013).

However, the distribution within this band challenges the assumption that frequency alone ensures ease of acquisition. The dataset shows that 1,201 items in the high-frequency band fall within the low similarity category. This indicates that a large portion of frequent English vocabulary is not transparent to Italian learners. In other words, the most common words in English do not necessarily provide meaningful cues for learners whose first language is Italian. This result supports the observation from the fieldwork that learners often relied on guessing strategies, rather than straightforward recognition, when engaging with high-frequency words. In a bilingual context, high frequency can therefore be associated with increased difficulty when lexical transparency is limited. This outcome aligns with existing research on cross-linguistic influence and lexical representation. De Groot and Keijzer (2000) demonstrated that cognate status and formal similarity significantly affect vocabulary learning and retention, and that non-cognates, even when frequent, may be more difficult to acquire and recall. The present data suggest that for Italian learners, the most frequent English words may behave as non-transparent items, resulting in a learning experience that is more cognitively demanding than expected. This implies that frequency-driven exposure, while necessary, is not sufficient to guarantee accessibility for all learners.

6.6.3.3 Mid-frequency vocabulary: a gap in progression

The mid-frequency band (Nation level 2) is markedly underrepresented compared to the high-frequency band, with only 1,127 processed items. This pattern suggests that learners do not consistently progress through the vocabulary continuum as expected. In many language learning contexts, mid-frequency vocabulary is considered essential because it represents a transitional zone between foundational lexical knowledge and more specialized vocabulary (Nation, 2006). Mid-frequency words often carry higher semantic specificity and are crucial for moving beyond basic comprehension toward more complex reading and writing.

The relative underrepresentation of this band may be interpreted as evidence of attrition in lexical progression, where learners tend to remain within the initial stages of the learning pathway. This could be due to several factors, including motivational decline, insufficient scaffolding, or a perceived lack of relevance in later modules. However, the data also suggests that the platform's frequency-based structure may unintentionally reinforce this trend. If the initial stages are dominated by frequent but low-transparency words, learners may experience difficulty early on, which can discourage further engagement and reduce progression into mid-frequency vocabulary.

This observation resonates with the broader literature on learning engagement and attrition in digital environments, where learners often disengage after initial exposure if they do not experience rapid success or perceivable progress (Grabe & Stoller, 2002). The present results therefore highlight the need to reconsider the sequencing and scaffolding of lexical content, ensuring that learners are gradually introduced to items that offer clearer cross-linguistic cues.

6.6.3.4 Low-frequency vocabulary: transparency compensates for rarity

One of the most striking findings is that the low-frequency band (>7,500) contains a high number of high-similarity items (506), which is comparable to the mid-frequency band and only slightly lower than the high-frequency band. This suggests that some low-frequency words may be more transparent than many high-frequency items, likely due to shared Latin roots and cognate relationships between Italian and English. This aligns with the concept of lexical transparency as a cross-linguistic phenomenon: words derived from the same etymological source are often recognizable across languages even when they are less common in corpus frequency.

From a theoretical perspective, this supports the idea that lexical difficulty cannot be predicted solely by frequency. While frequency influences exposure and automaticity (Ellis, 2002), cross-linguistic transparency affects recognition and initial comprehension. The presence of transparent low-frequency words implies that learners can access and understand less frequent vocabulary when it has clear morphological and semantic parallels in their L1. This is consistent with research on cognates and cross-linguistic transfer, which demonstrates that formal similarity can accelerate learning and retention even when words are not highly frequent (Jarvis & Pavlenko, 2008; De Groot, 1992).

Pedagogically, this finding suggests that the platform could benefit from integrating transparency-aware metrics into its sequencing algorithm. Rather than relying exclusively on frequency, the system should prioritize words that are both frequent and transparent in early stages, while introducing low-frequency transparent words strategically to build confidence and motivation.

6.6.3.5 Transparency as a central factor in learner behavior

The combined similarity metric reveals that learners are more likely to interact with high-similarity items across all frequency bands. This indicates that transparency plays a significant role in guiding learner choices and comprehension. The high-similarity category contains the highest counts in all bands, reinforcing the idea that learners prefer items that can be recognized through formal or semantic resemblance.

This pattern aligns with research on cross-linguistic influence and bilingual lexical processing. Learners rely on L1 knowledge to make inferences about L2 meaning, particularly when explicit instruction is limited (Jarvis & Pavlenko, 2008). The transparency effect observed in the data is therefore consistent with the notion of transfer: learners use structural similarities to infer meaning and facilitate processing, reducing cognitive load and increasing engagement.

At the same time, the substantial number of low-similarity high-frequency words suggests that learners are frequently confronted with items that offer little cross-linguistic support. This creates a tension between exposure and accessibility, which may explain why the platform's frequency-based loop approach did not produce the expected outcomes in terms of learner progression.

6.6.4 Conclusion

The similarity $\times$ frequency analysis demonstrates that the platform's frequency-driven design, while theoretically grounded, does not fully account for the cross-linguistic realities of Italian learners of English. High-frequency vocabulary is not always accessible due to limited transparency, while low-frequency vocabulary may be more easily processed when it shares morphological or semantic similarities with Italian. These findings support a shift toward a hybrid model that combines frequency exposure with transparency-based sequencing, aligning instructional design with bilingual lexical processing principles and promoting deeper engagement and progression.

These results have direct implications for the future development of the platform and for the broader design of adaptive learning systems to: Rebalance the lexical sequencing algorithm. The system should integrate similarity as a core factor, not only frequency. High-frequency, low-similarity words should be introduced with stronger scaffolding, explicit instruction, or supportive contextual cues to prevent early disengagement and reduce reliance on guessing strategies. Moreover, transparent vocabulary should be introduced earlier in the learning sequence. Low-frequency transparent words can be used strategically to build learner confidence and reduce attrition, since transparency can compensate for low frequency and enable learners to experience early success and sustained motivation. Adaptive pathways can also be designed based on transparency profiles, guiding learners through personalized learning paths that prioritize items with higher transparency

before progressing to more opaque vocabulary. This approach would create a smoother learning curve and better align lexical progression with learner cognitive readiness. Finally, similarity metrics can be used to predict learner difficulty more accurately than frequency alone, allowing for more effective feedback, error correction, and adaptive reinforcement, which are essential for sustained learning and long-term retention.

In conclusion, the overall lesson design exemplifies an integrative approach to language pedagogy, where theory, practice, and technology converge to create a dynamic and learner-centered environment. By structuring instruction around progressive, interrelated modules, the framework ensures continuity between phonetic, lexical, grammatical, and communicative development. Each lesson contributes to a holistic acquisition process grounded in established SLA theories and enriched by interactive, multimodal methodologies. The deliberate use of scaffolding, gamification, and cultural contextualization transforms traditional instruction into an adaptive and motivating experience that supports long-term retention and communicative competence.

Ultimately, the pedagogical architecture presented here demonstrates how principled design, when aligned with cognitive and linguistic research, can produce a sustainable, inclusive, and engaging model for second language learning in digital environments. The evidence from the similarity $\times$ frequency analysis reinforces the need for adaptive systems that balance exposure with cross-linguistic transparency, ensuring that learning pathways are both accessible and challenging in a way that supports real progress and learner autonomy.

In conclusion, the overall lesson design exemplifies an integrative approach to language pedagogy, where theory, practice, and technology converge to create a dynamic and learner-centered environment. By structuring instruction around progressive, interrelated modules, the framework ensures continuity between phonetic, lexical, grammatical, and communicative development. Each lesson contributes to a holistic acquisition process grounded in established SLA theories and enriched by interactive, multimodal methodologies. The deliberate use of scaffolding, gamification, and cultural contextualization transforms traditional instruction into an adaptive and motivating experience that supports long-term retention and communicative competence. Ultimately, the pedagogical architecture presented here demonstrates how principled design, when aligned with cognitive and linguistic research, can produce a sustainable, inclusive, and engaging model for second language learning in digital environments.

7 Conclusion

This final chapter synthesizes the key findings, implications, and reflections emerging from the design, development, and evaluation of the Appil platform. It brings together the theoretical, pedagogical, and technological dimensions explored throughout the study, highlighting both the achievements and limitations of the project.

The discussion that follows examines the outcomes in relation to the research questions, outlines the platform's contributions to the fields of Second Language Acquisition and digital pedagogy, and identifies areas for further research and development. In doing so, it situates the work within the broader discourse on technology-enhanced language learning and emphasizes the ongoing need for collaboration, adaptability, and learner-centered innovation in digital education.

7.1 Study Limitations

Despite the significant advancements and innovations achieved through the development of the platform, it is important to acknowledge several limitations that may affect both its scope and impact. One notable constraint relates to the availability and quality of materials incorporated into the learning environment. In order to respect intellectual property rights and promote open access, the platform has primarily relied on content licensed under Creative Commons. While this approach ensures legal compliance and supports the sharing of resources, it can also result in certain limitations regarding the depth, breadth, and diversity of available learning materials.

The use of freely licensed images, texts, and supplementary resources occasionally necessitates compromises in content quality or relevance, particularly when compared to proprietary, professionally produced educational materials. There may be instances where suitable Creative Commons resources are limited or where higher-quality alternatives are not available within open repositories. This can, in turn, affect the richness and authenticity of the learning experience provided to users, potentially impacting engagement, motivation, and overall satisfaction. For example, while the integration of open-access images and descriptions has enabled the rapid expansion of the noun database and flashcard pool, the platform remains dependent on the quality and specificity of what is publicly available. This may lead to inconsistencies in visual appeal, depth of explanation, or cultural nuance, elements that can be critical for effective and immersive language learning.

To address these challenges and further enhance user engagement, future iterations of the platform may explore partnerships with content creators, the development of original materials, or the inclusion of selectively licensed, higher-quality resources where possible. Ultimately, while open-licensed

content provides an essential foundation for accessible and collaborative learning, ongoing attention to the improvement and curation of educational materials will be vital in sustaining user motivation and supporting deeper learning outcomes.

In addition to material limitations, the platform's technical infrastructure presents its own set of constraints. The deployment and ongoing maintenance of the server, management of the database, and implementation of platform features have all been conducted by a single researcher-developer. This one-person approach introduces bottlenecks in terms of scalability, system robustness, and the ability to implement new features or rapidly respond to technical issues. The provision of active user participation mechanisms and instant feedback, features that typically benefit from team-based development and dedicated support resources, have likewise been managed solely by the project's creator. This situation limits the potential for comprehensive quality assurance, continuous improvement, and the rapid incorporation of user feedback, all of which are crucial for delivering a high-quality, engaging, and reliable learning experience.

Collectively, these material and technical limitations underscore the importance of collaboration, resource diversification, and potential partnerships in future development. While open-licensed content and single-developer ingenuity have allowed the platform to reach its current state, the exclusive responsibility for system deployment, server configuration, and long-term maintenance borne by a single researcher represents a significant constraint. Over an eleven-month operational period, this involved not only the initial deployment of the application in a production-ready format, but also the continuous management of the server environment, including system updates, security considerations, database administration, and service availability.

Managing these tasks independently requires the acquisition and application of specialized technical knowledge typically distributed across roles such as system administration, backend development, and DevOps engineering. As a result, development resources were necessarily divided between pedagogical design, software development, and infrastructural maintenance. This limited the capacity for parallel feature development, extensive testing, and rapid iteration in response to user feedback, and increased reliance on individual expertise for system stability and troubleshooting.

Consequently, scaling the platform to support richer content, higher user volumes, and more advanced functionalities would require a redistribution of these responsibilities within a collaborative development framework. Broader access to technical support and infrastructure expertise, alongside diversified content resources, will be essential for improving robustness, scalability, and long-term sustainability. Addressing these limitations is therefore a prerequisite for maintaining user engagement and for realizing the platform's full potential as a dynamic, user-centered language learning environment.

An often underestimated yet foundational factor in the development of any educational technology platform is the availability of adequate financial resources. The creation of an effective resource, one that offers both high-quality didactic materials and robust technical infrastructure, inevitably depends on sustained investment. Financial support directly influences the platform's ability to access, license, or commission professional-grade educational content, which is essential for maximizing learner engagement, cultural relevance, and pedagogical accuracy. While the use of open-access and Creative Commons-licensed materials has enabled the current platform to reach a broad audience at minimal cost, the pursuit of higher quality, more diverse, and more culturally authentic resources often necessitates expenditure beyond what is available through freely accessible repositories.

Similarly, the development and maintenance of a scalable, secure, and feature-rich technological infrastructure require ongoing financial investment. This includes costs associated with server hosting, database management, user interface design, security updates, and the implementation of advanced functionalities such as AI-driven personalization and real-time feedback systems. Relying on a single developer, while cost-effective in the short term, poses limitations in terms of scalability, innovation, and responsiveness to user needs. Additional funding would not only enable the recruitment of a multidisciplinary team, including educators, developers, designers, and support staff, but also support the platform's evolution toward a more comprehensive, stable, and responsive learning environment.

From an empirical perspective, the evaluation dataset remains limited in scale. Approximately 134 articles were processed by around 20 users, corresponding to an average of six to seven texts per participant. While adequate for exploratory validation, this sample size constrains the statistical robustness of quantitative claims and limits the generalizability of usage patterns observed during the fieldwork.

7.1.1 Ethical considerations

In addition to the challenges associated with technical infrastructure management and pedagogical planning, the development and evaluation of the platform required careful consideration of ethical issues related to data protection, learner profiling, and equitable access to learning resources. These dimensions are closely intertwined: decisions made at the technical and pedagogical levels directly shape how learner data are collected, interpreted, and acted upon.

From a data protection perspective, the processing of learner interaction data was conducted in accordance with the principles of the General Data Protection Regulation (GDPR, (Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and

repealing Directive 95/46/EC (General Data Protection Regulation, 2016)). Personal data were handled in line with the principles of lawfulness, fairness, and transparency, and data collection was restricted to what was strictly necessary for platform functionality and pedagogical personalization. Learner identifiers were pseudonymized, and access to server-side data and logs was limited to the researcher responsible for system deployment and maintenance. Within the empirical study, participation was voluntary, informed consent was obtained, and all evaluation data were anonymized and reported only in aggregated form. No automated decision-making with legal or similarly significant effects, as defined under Article 22 of the GDPR, was implemented.

Beyond regulatory compliance, learner profiling raises broader ethical considerations that intersect with pedagogical design. Longitudinal interaction data may implicitly reveal learning difficulties, engagement patterns, or personal interests. As the platform evolves toward more advanced analytics and AI-driven personalization, ensuring transparency of profiling mechanisms and preserving learner autonomy will become increasingly important. Measures such as clear user-facing explanations, configurable privacy settings, and the possibility to limit or opt out of certain profiling features would help balance personalization benefits with ethical responsibility.

To mitigate these limitations, the platform prioritized the use of resources that are comparatively neutral and linguistically transversal, most notably Wikipedia-based data. Wikipedia offers broad multilingual coverage and a relatively standardized editorial framework, which helps reduce domain-specific and language-specific bias when compared to more specialized or proprietary datasets.

Nevertheless, the use of Linguistic Linked Open Data (LLOD) resources continues to raise ethical considerations related to bias and representational imbalance. Even widely used resources such as Wikipedia may reflect historical, cultural, or linguistic asymmetries, including uneven coverage across languages, registers, or topics. These limitations can influence pedagogical content selection, difficulty progression, and the semantic associations presented to learners.

Accordingly, responsible pedagogical planning requires ongoing monitoring of resource behavior, diversification of data sources where feasible, and critical evaluation of system outputs to identify and mitigate unintended bias. This reflective approach acknowledges the practical benefits of linguistically transversal resources while maintaining awareness of their ethical and pedagogical constraints.

Finally, accessibility and inclusivity constitute an ethical dimension that extends beyond both infrastructure and pedagogy. Variations in linguistic background, digital literacy, cognitive profiles, and access to technology may affect learners' ability to engage with the platform. Future development should therefore integrate accessibility-aware interface design, adaptable learning pathways, and support for diverse learner needs to promote equitable participation.

Taken together, ethical considerations form a third foundational dimension of platform sustainability, complementing technical infrastructure management and pedagogical planning. Addressing these issues explicitly is essential for maintaining user trust and for supporting the responsible, scalable development of personalized digital learning environments.

The implementation and evaluation of the platform involved the processing of learner interaction data, which required careful alignment with the principles set out in the General Data Protection Regulation (GDPR - (Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation, 2016), see Appendix B). Particular attention was given to learner profiling activities, both at the level of server-side system operation and within the empirical research process

On the server side, personal data were processed in accordance with the principles of lawfulness, fairness, and transparency (Article 5(1)(a)). Data collection was limited to what was strictly necessary for platform functionality and pedagogical personalization, reflecting the principle of data minimization (Article 5(1)(c)). Learner interaction data were used exclusively for clearly defined purposes related to learning progression and system operation, consistent with purpose limitation (Article 5(1)(b)). User identifiers were stored in a pseudonymized form where possible, reducing the risk of direct identification in line with Article 4(5).

Access to server-side data and logs was restricted to the researcher responsible for system deployment and maintenance, and appropriate technical and organizational measures were applied to ensure integrity and confidentiality (Article 5(1)(f)). These measures included controlled access to the hosting environment, secure configuration of the server, and regular system updates throughout the eleven-month operational period. No automated decision-making with legal or similarly significant effects, as defined under Article 22, was implemented.

Within the research context, additional safeguards were applied. Participation in the study was voluntary, and learners were informed about the nature of the data collected and its use for academic research purposes, supporting informed consent and transparency requirements (Articles 6 and 13). Data used for analysis were aggregated and anonymized prior to evaluation, and no results were reported at the individual level. This approach ensured that empirical findings could not be traced back to identifiable participants, aligning with GDPR expectations regarding research data processing and storage limitation (Article 5(1)(e)).

Nevertheless, learner profiling, particularly when combined with longitudinal interaction data, raises broader ethical considerations beyond formal compliance. Behavioral data may implicitly reflect learning difficulties, engagement patterns, or personal interests. As future iterations introduce more

advanced analytics or AI-based personalization, particular attention will be required to ensure transparency of profiling mechanisms, prevent unintended bias, and preserve learner autonomy. Measures such as clearer user-facing explanations, configurable privacy options, and the ability to limit or opt out of certain profiling features would further strengthen alignment with GDPR principles and ethical best practices.

Addressing data protection explicitly at both the technical and research levels is essential for maintaining user trust and for ensuring the responsible development of personalized learning environments. Continued attention to GDPR-aligned data governance will therefore remain a prerequisite for the platform's future scalability and integration of more advanced data-driven components.

7.2 Future Research Directions

During the development of the present work, the pace of technological innovation, accelerated by breakthroughs in artificial intelligence, has reshaped educational tools and platforms at an unprecedented rate; yet, even as AI-driven translation engines, conversational agents, and adaptive testing systems become more sophisticated, the core imperative of personalization in second language learning remains paramount. While these advanced tools can automate certain aspects of content delivery and feedback, they cannot replace the nuanced, learner-centered adjustments that arise from understanding individual motivations, interests, and linguistic goals. In fact, the very fluidity of AI technologies underscores the importance of designing systems that can continuously adapt to new developments: by embedding personalized pathways, whether through curated topic modules, dynamically generated exercises, or AI-powered recommendations, educators ensure that learners retain agency over their progress and remain engaged despite the shifting technological landscape. As AI capabilities grow, the challenge shifts from merely implementing new features to thoughtfully integrating them in ways that reinforce, rather than dilute, the learner's control over their own trajectory.

Looking to the future, the ongoing evolution of artificial intelligence presents exciting opportunities for the further development of digital language learning platforms. The integration of AI-driven tools, such as adaptive learning algorithms, natural language understanding, automated feedback systems, and personalized recommendation engines, has the potential to significantly enhance the platform's responsiveness to individual learner needs. For example, AI could be employed to more precisely analyze user progress, tailor content recommendations in real-time, and dynamically adjust the level of difficulty or presentation mode to optimize learning efficiency and engagement. Moreover,

advanced natural language generation models can facilitate the creation of increasingly rich, context-aware explanations, conversation simulations, and culturally authentic materials, thereby broadening the range of experiences available to learners.

It is also important to highlight the transformative potential that artificial intelligence holds for the continued enhancement of the platform. AI technologies, ranging from adaptive learning algorithms and intelligent tutoring systems to advanced natural language processing and data-driven content recommendations, offer unprecedented opportunities to both personalize the learning experience and streamline system operations. Integrating AI into the platform could enable real-time analysis of learner performance, automatic adaptation of didactic materials to individual needs, and the generation of more engaging, contextually relevant content. Additionally, AI-driven engagement strategies, such as conversational agents or gamified feedback mechanisms, could help sustain user motivation and deepen participation. While the platform in its current form has already laid a strong foundation for collaborative and user-centered learning, the thoughtful incorporation of artificial intelligence would further elevate its capacity to respond dynamically to learner input, optimize educational pathways, and create richer, more interactive experiences.

However, despite these technological advances, it is essential to reaffirm that the ultimate success of any AI-enhanced educational system continues to rely fundamentally on the learner's active participation. Artificial intelligence can provide sophisticated scaffolding, anticipate learning obstacles, and make personalized suggestions, but it cannot replace the motivation, curiosity, and sustained effort that characterize effective language learning. The most promising future direction lies in harnessing AI not to automate or supplant the learner's role, but to augment and support the individual's agency, encouraging exploration, promoting self-reflection, and responding to user input in ways that make learning more meaningful and engaging.

It is also important to highlight the transformative potential that artificial intelligence holds for the continued enhancement of the platform. AI technologies, ranging from adaptive learning algorithms and intelligent tutoring systems to advanced natural language processing and data-driven content recommendations, offer unprecedented opportunities to both personalize the learning experience and streamline system operations. Integrating AI into the platform could enable real-time analysis of learner performance, automatic adaptation of didactic materials to individual needs, and the generation of more engaging, contextually relevant content. Additionally, AI-driven engagement strategies, such as conversational agents or gamified feedback mechanisms, could help sustain user motivation and deepen participation. While the platform in its current form has already laid a strong foundation for collaborative and user-centered learning, the thoughtful incorporation of artificial intelligence would further elevate its capacity to respond dynamically to learner input, optimize

educational pathways, and create richer, more interactive experiences. Nevertheless, it remains essential that these technological advancements complement, rather than replace, the crucial role of learner agency and personal motivation, ensuring that AI serves to empower rather than supplant the human elements of language acquisition.

As the platform evolves, ongoing attention must be paid to ensuring that technological advancements serve to empower learners, fostering environments where human engagement and technological innovation work hand in hand. Only by maintaining this focus on the centrality of the learner can the full educational promise of artificial intelligence be realized.

To ensure that the platform evolves in a way that is both pedagogically sound and ethically responsible, the integration of large language models (LLMs) should follow a staged research and implementation plan. The roadmap below outlines the main phases of development:

Phase	Objective	Key Activities	Expected Outcome
1	Frequency-based lexical research and text evaluation	Extend frequency analysis; validate readability and transparency indices; develop text difficulty metrics	A reliable framework for evaluating and selecting texts
2	Controlled integration of LLMs for content generation	Use LLMs to generate exercises and materials only when aligned with established metrics; implement human-in-the-loop review	AI-assisted content generation that remains pedagogically compliant
3	Pedagogical validation and didactic compliance	Validate learning outcomes, level appropriateness, cultural relevance, and	High-quality materials aligned with curriculum and learning goals

Phase	Objective	Key Activities	Expected Outcome
		cognitive load; collaborate with language teachers	
4	Ethical assessment and bias monitoring	Assess model bias, fairness, and inclusivity; ensure privacy compliance (e.g., GDPR); monitor outputs continuously	Safe and responsible AI usage with long-term oversight
5	Adaptive learning and personalized feedback	Deploy AI-driven recommendations, adaptive difficulty, automated feedback, and conversational practice	Personalized learning pathways that respect learner agency and motivation

Table 15 Roadmap for the conscientious integration of large language models (LLMs) into the platform. The roadmap outlines a staged approach to incorporating AI-driven capabilities while maintaining pedagogical validity, ethical compliance, and learner-centered design. Phase 1 focuses on strengthening the existing frequency-based lexical research framework and validating text evaluation metrics. Phase 2 introduces LLMs for content generation under strict constraints and human supervision. Phase 3 emphasizes pedagogical validation through teacher collaboration and curriculum alignment. Phase 4 addresses ethical risks, bias monitoring, and GDPR compliance. Phase 5 proposes the deployment of adaptive, personalized learning features, ensuring that AI supports learner agency and motivation rather than replacing them.

The roadmap presented in Table 15 outlines a phased approach for integrating LLM-based capabilities into the platform in a responsible and pedagogically sound manner. The proposed progression begins with reinforcing the existing frequency-driven lexical framework and validating text complexity metrics, ensuring that any future AI features are grounded in robust linguistic research. Subsequent phases introduce LLMs in a controlled manner, emphasizing human oversight and alignment with established learning goals. Pedagogical validation and teacher collaboration are prioritized to guarantee that generated materials meet didactic standards and curricular relevance. Crucially, ethical assessment and bias monitoring are incorporated before deploying adaptive learning features, reflecting the need for continuous oversight in AI-supported education. Ultimately, the roadmap seeks

to balance innovation with learner agency, ensuring that AI functions as an augmentative tool rather than a replacement for human-centered instruction.

7.3 Concluding Reflections

This study has examined the development, deployment, and usage of a collaborative language learning platform, with specific attention to its technical architecture, patterns of user engagement, and ongoing resource expansion. The platform's design, rooted in the principles of open, shared knowledge-building, demonstrates how collaborative contributions can be integrated to enhance both the depth and quality of learning materials. Features such as the systematic processing of nouns, integration of multimodal resources, and automated generation of personalized flashcards have collectively enriched the learning environment, making it more adaptive and engaging for a diverse user base.

The transition to a fully digital instructional model, while pedagogically and strategically aligned with the goals of the municipality, presented significant logistical and operational challenges, particularly given that the platform development, instructional design, and implementation were carried out by a single individual. This configuration, while agile and responsive in many ways, imposed notable constraints on time, workload, and the iterative refinement of both content and technical infrastructure.

Following consultations with the municipal supervisor, a shared decision was made to pursue a fully online model for the delivery of the language learning program. This approach was selected not only to enhance accessibility and scheduling flexibility for adult learners with varied work commitments, but also to align with the municipality's broader commitment to digital modernization. The digital-first model, however, necessitated a comprehensive overhaul of the original instructional structure, including reformatting pedagogical materials, building platform functionalities from scratch, and developing user support resources suitable for participants with low to moderate digital literacy.

The primary difficulty of operating as a one-person team lay in the simultaneity of roles: the designer was required to function concurrently as curriculum developer, technical integrator, instructional content author, quality assurer, and platform manager. Each of these domains demanded distinct forms of expertise, from knowledge of adult second language acquisition theories to familiarity with user experience (UX) design and learning management system (LMS) technologies. Without departmental or peer support, all decisions, from interface aesthetics to content chunking strategies and error debugging, had to be managed independently, often under tight timelines imposed by institutional coordination needs.

Moreover, the design process was marked by constant iteration. Rather than developing a static course and deploying it at once, the platform was constructed in parallel with ongoing participant feedback and usage data. This meant that instructional materials were updated almost daily, based on real-time observations of learner engagement patterns, platform navigation difficulties, and emergent content preferences. As such, the development process became inherently dialogic, resembling a continuous feedback loop where pedagogical decisions were informed directly by learners' evolving needs and experiences, a dynamic that was pedagogically beneficial but operationally strenuous, as it required ongoing commitment to both content adaptation and technical updates across the program's lifespan. Considerable time was devoted to simplifying the interface, minimizing the number of required steps to complete each task, and embedding scaffolding such as tooltips, video walk-throughs, and clear visual hierarchies. These adjustments were essential for ensuring learner autonomy, especially in the absence of consistent live technical support. Despite these constraints, the one-person design structure also fostered a high degree of pedagogical integrity and continuity. The seamless integration of instructional content with platform functionality allowed for rapid adaptation, minimized misalignment between design intent and implementation, and supported the development of a tightly focused, learner-responsive experience. Nevertheless, this model was only sustainable over a finite project duration and would pose significant challenges for scalability or replication without additional human and technical resources.

In sum, while the fully digital model ultimately delivered a flexible, accessible, and pedagogically coherent learning environment, its development underscored the practical limitations and emotional strain of single-person implementation. It also highlighted the critical importance of institutional recognition of design labor and the need for structured support mechanisms in future iterations or similar undertakings.

Quantitative analysis of web server logs and database records has revealed important behavioral trends. While initial engagement with the platform is strong, especially with entry-level content and onboarding processes, there is a noticeable decline in sustained participation and progression through advanced instructional modules.

A critical aspect to consider is the unique challenge of developing and maintaining such a comprehensive platform as a single individual. Both didactically and functionally, a one-man development approach imposes significant constraints. From a didactic perspective, balancing curriculum design, resource curation, and the integration of effective pedagogical strategies within a rapidly evolving digital environment is particularly demanding. Ensuring instructional coherence and addressing diverse learner needs without the benefit of a multidisciplinary team requires ongoing adaptability and prioritization. Functionally, the responsibility for the full technological stack, from

backend infrastructure and database management to user interface design and feature implementation, falls solely on the developer, limiting both the pace of innovation and the breadth of features that can be introduced or maintained. Inevitably, this also means that scalability, timely troubleshooting, and the regular incorporation of user feedback can be more challenging compared to projects developed by larger teams.

Regarding RQ1: *How can Natural Language Processing (NLP) and personalized content choices enhance the effectiveness of second language learning, allowing learners to have more control over their learning paths and content selection?*, Natural Language Processing (NLP) and personalized content choices significantly enhance second language learning by empowering learners to take greater control of their learning paths. NLP technologies, such as automatic vocabulary extraction, adaptive text simplification, speech recognition, and personalized flashcard generation, enable the creation of materials that align with learners' individual interests, proficiency levels, and goals. By dynamically adjusting input difficulty and providing immediate feedback, NLP fosters an adaptive learning environment where content evolves alongside the learner's progress.

Personalization allows learners to engage with authentic materials, news articles, stories, or dialogues, relevant to their personal or professional contexts, promoting intrinsic motivation and sustained engagement. Moreover, NLP-driven tools can analyze learners' performance data to identify strengths and weaknesses, suggesting targeted activities to address specific gaps. This data-informed approach supports self-regulated learning, as students can track their development and make informed decisions about their next learning steps.

On the other hand, NLP functions as a complementary tool rather than a replacement for traditional pedagogical approaches. While it enhances adaptability, feedback, and personalization, effective language learning still depends on human interaction, communicative practice, and cultural context, all vital for developing fluency and pragmatic competence. When integrated within a blended learning framework, NLP complements the teacher's role by automating repetitive tasks, offering individualized feedback, and freeing instructional time for meaningful communicative activities.

Recent advances in large language models (LLMs) have expanded the capabilities of NLP, introducing sophisticated systems that can generate, analyze, and personalize language input with unprecedented precision. The evolution of NLP through LLM-based architectures has shifted attention toward context-aware, adaptive, and conversational learning experiences. However, these innovations highlight the need for NLP to be integrated into broader educational technologies, ensuring that its potential is harnessed within structured pedagogical frameworks rather than used in isolation. The convergence of NLP, artificial intelligence, and learning analytics points toward a

holistic model of digital language education, one that unites technological intelligence with human-centered pedagogy.

Regarding RQ2: *Is it possible to create a learning environment that follows a traditional learning framework while still allowing learners the freedom to choose their own content, thereby balancing structure with personalization?*, in practice, this balance was achieved through a dual-layered design. The core instructional sequence—comprising eight progressive modules, ensured pedagogical coherence and alignment with A2–B1 proficiency targets. Within this framework, however, the content layer was flexible: learners could select or generate material (e.g., nouns, contexts, or thematic texts) based on personal preference. This design preserved the structural predictability of traditional instruction, allowing teachers or systems to monitor progress, while granting learners freedom of exploration within defined linguistic boundaries. The result was a learning model that combined guided progression with adaptive exploration, encouraging both mastery and autonomy.

Implementing this model in a fully digital environment presented several challenges. Chief among these was maintaining usability and instructional coherence in a platform accessed by users with widely varying digital literacy levels. Ensuring that personalization did not translate into confusion required meticulous attention to interface design, scaffolding, and support mechanisms. Features such as intuitive navigation, clear visual hierarchies, integrated tutorials, and tooltips were essential in reducing cognitive load and enabling learners to focus on content rather than technology. Furthermore, the inclusion of contextual prompts, interactive feedback, and automated content suggestions helped sustain alignment between learner-selected materials and the overall pedagogical trajectory.

Another layer of complexity involved ensuring consistency between pedagogical intent and user experience. In many adaptive systems, algorithmic personalization risks undermining coherence by presenting learners with disjointed or overly individualized pathways. This research sought to counteract that by embedding personalization within a didactic architecture grounded in Second Language Acquisition (SLA) principles, particularly those concerning input comprehensibility, noticing, and output practice (Krashen, 1982; Schmidt, 1990; Swain, 1995). By aligning personalization mechanisms with linguistic progression, the platform avoided the pitfalls of over-personalization, maintaining a careful equilibrium between flexibility and continuity.

Equally important was the social dimension of this model. Even in an asynchronous, digital context, opportunities for collaboration and peer interaction were integrated to support engagement and accountability. Learners could share self-generated content, exchange examples, and discuss topics of mutual interest, thereby reinforcing communal learning within a personalized environment. This

aligns with sociocultural theories of learning (Vygotsky, 1978), which stress the importance of social interaction and scaffolding in cognitive development.

Despite the success of this model, sustaining such a fine balance proved demanding. Continuous monitoring of user behavior, performance, and feedback was necessary to ensure that personalization remained pedagogically meaningful. The system required ongoing updates to maintain alignment between learner-driven content and curricular objectives. The one-person development structure further amplified these challenges, as the designer simultaneously managed technical infrastructure, instructional design, and quality assurance. Nevertheless, this constraint also fostered conceptual unity, each decision, from feature implementation to content adaptation, reflected a coherent pedagogical vision.

Overall, the findings indicate that a hybrid framework, combining the systematic rigor of traditional instruction with the flexibility of learner-driven personalization, can effectively promote both linguistic competence and learner motivation. This model responds to the increasing need in digital education for systems that are adaptive yet pedagogically grounded, ensuring that personalization serves learning objectives rather than fragmenting them.

Looking ahead, further research and development should focus on complementing quantitative usage data with qualitative feedback to better understand user motivations, barriers, and preferences. Surveys, interviews, and focus groups can offer nuanced insights into the learner experience, enabling more responsive and inclusive platform enhancements. By continually integrating user input, technological innovation, and pedagogical best practices, the platform can evolve into a truly dynamic and impactful learning environment, one that not only disseminates knowledge but also actively cultivates a collaborative community of practice.

The findings from this project offer promising insights in response to the research questions posed. Regarding RQ1, the integration of Natural Language Processing (NLP) within the platform has demonstrably enhanced the effectiveness of second language learning by enabling automated processing and enrichment of vocabulary entries. NLP techniques facilitate the generation of personalized learning materials, such as tailored flashcards enriched with context-specific descriptions and images, thereby supporting varied learning preferences and greater autonomy in content selection. This personalized approach empowers learners to chart individualized learning paths, fostering deeper engagement and improved retention. In relation to RQ2, the platform's architecture successfully reconciles the tension between structured instruction and learner autonomy. While the core instructional sequence retains elements of a traditional curriculum, providing a clear progression of lessons and scaffolded activities, the system's design allows users to select and process content of personal interest, such as specific nouns or topics, and integrate these into their individual

study routines. This balance between pedagogical structure and personalization not only accommodates diverse learning goals but also supports sustained motivation, as evidenced by user engagement data. Collectively, these outcomes suggest that technology-driven personalization, when thoughtfully combined with foundational instructional frameworks, can meaningfully expand the possibilities for effective and learner-centered language education.

The achievement of developing a platform that enables genuine personalization and free content choice marks a significant step forward in second language acquisition technology. By leveraging NLP and robust data architecture, the system is able to dynamically adapt to individual learner profiles, preferences, and evolving needs. Each user can select content, such as nouns, images, and descriptions, that aligns with their personal interests, professional goals, or immediate communicative contexts. This level of autonomy fundamentally transforms the learning experience from a rigid, one-size-fits-all progression to a flexible, user-driven journey. At the same time, the platform does not sacrifice the benefits of structured learning; instead, it overlays personalization on top of a coherent pedagogical framework, ensuring that users receive both guidance and freedom. The automatic generation and integration of personalized flashcards, the possibility to revisit and repractice selected material, and the capacity to process and enrich new content all contribute to a learning environment that is both responsive and empowering. Such features not only promote intrinsic motivation but also facilitate deeper cognitive engagement and more meaningful connections with the language being learned. This balance between structure and choice, achieved through thoughtful system design, demonstrates that it is indeed possible to create an environment where learners are both supported by best-practice instructional design and liberated to pursue their own paths, addressing the core aspirations set out in both research questions.

The strengths of the platform are closely aligned with several well-established theories from educational psychology and second language acquisition research. Central to this is the principle of learner autonomy, highlighted in self-determination theory (Deci & Ryan, 1985), which posits that individuals are most motivated and effective in their learning when they have meaningful control over their choices and activities. By allowing users to select the content most relevant or interesting to them and integrating personalized resources such as custom-generated flashcards and context-specific imagery, the platform creates an environment where learners can experience a high degree of autonomy and competence, two pillars that fuel intrinsic motivation. Furthermore, the platform's emphasis on user-driven content selection resonates with constructivist approaches (Vygotsky, 1978; Bruner, 1966), which emphasize the importance of active engagement, contextualization, and the construction of personal meaning in the learning process. The ability for users to process, enrich, and revisit content in accordance with their own preferences mirrors best practices in adaptive learning

and mastery-based progression, offering ample opportunities for repetition, elaboration, and deeper cognitive processing.

Despite these strengths, it is crucial to recognize that technological affordances alone are not sufficient to guarantee successful or sustained language learning outcomes. Decades of research, including insights from motivational psychology and self-regulated learning (Zimmerman, 2002) make clear that the most effective learning environments are those that not only provide scaffolding and resources, but also inspire and sustain the learner's own commitment and investment. In this sense, motivation and personal involvement remain absolutely central to the platform's effectiveness. The system can facilitate choice, tailor content, and offer repeated practice opportunities, but the transformative impact ultimately depends on the user's willingness to engage, explore, and persevere. As learners move through personalized materials, their sense of ownership and connection to the language grows, but only if these experiences are met with genuine curiosity and active participation. This underscores the ongoing challenge for platform design: to not only provide flexible, adaptive pathways, but also to continually nurture user motivation, through elements such as meaningful feedback, progress tracking, community support, and the clear demonstration of relevance to individual goals. Ultimately, the platform's technological innovations must be coupled with strategies to foster self-reflection, intrinsic interest, and long-term commitment if its full potential for effective, lasting second language learning is to be realized.

While the platform employs advanced computational resources and automated processes to support and enrich language learning experience, it is important to recognize the inherent limitations of technology-driven solutions in the absence of meaningful learner participation. Computer-based systems can generate personalized content, scaffold learning sequences, and provide immediate feedback on a scale, yet they ultimately serve as facilitators rather than substitutes for genuine engagement. No matter how sophisticated the underlying algorithms or user interfaces, the effectiveness of such tools is fundamentally contingent upon the learner's willingness to actively engage, practice, and reflect. Research in educational psychology repeatedly demonstrates that technology alone cannot motivate or sustain deep learning without the learner's own initiative, self-regulation, and persistence. As such, the value of computational resources lies not in their capacity to automate learning, but in their ability to create environments that invite, support, and reward active learner participation. It is through this partnership between technological affordances and personal commitment that the most meaningful and lasting language learning outcomes can be achieved.

This doctoral research has explored the field of Intelligent Computer-Assisted Language Learning (ICALL), examining both its theoretical foundations and its practical implementation through the design and deployment of a digital English language learning platform. From the earliest stages of

conceptualization, it was clear that the landscape of second language learning technologies was already populated by a vast array of applications and digital tools. Many of these systems claimed to support personalized learning experiences, yet closer analysis revealed a persistent limitation: while they offered adaptive pathways or individualized feedback at a superficial level, few were genuinely grounded in the principles of Second Language Acquisition (SLA) or capable of dynamically adapting to learners' evolving linguistic needs.

The motivation behind this research was therefore to bridge this gap, to design and test an instrument that could embody SLA-informed pedagogy within a technologically mediated environment. The project sought not merely to digitize language instruction, but to create a system in which personalization arises from pedagogical intelligence rather than mechanical adaptation. In other words, the goal was to build a digital environment that reflects the interactional, incremental, and context-sensitive nature of language learning itself.

Through successive stages of theoretical review, system design, implementation, and evaluation, the research has demonstrated how core SLA principles, such as input comprehensibility, output production, feedback responsiveness, and lexical depth, can be operationalized within an ICALL framework. The platform developed during this project served both as an experimental testbed and as a functional learning environment, integrating modules that encouraged learner autonomy while supporting guided interaction with linguistic input. The deployment of the system at the Municipality of Fisciano provided a real-world context in which adult learners could engage with the platform as part of their professional development, allowing the research to examine authentic usage behaviors, motivational patterns, and pedagogical outcomes.

The empirical findings revealed strong initial engagement and positive attitudes toward flexibility, modularity, and contextual relevance. At the same time, the data highlighted familiar challenges in sustaining motivation and maintaining consistent participation over time, challenges that mirror those widely documented in the literature on digital and distance learning. Nevertheless, the integration of learner feedback, iterative updates, and the researcher's continuous involvement ensured that the platform evolved in a responsive and pedagogically sound manner. The project's dual emphasis on technical reliability and pedagogical validity has shown that effective digital learning design must combine both dimensions from the outset.

Ultimately, this research contributes to the growing body of ICALL scholarship by demonstrating that adaptive, learner-centered systems need not rely solely on complex artificial intelligence algorithms to achieve personalization. Instead, meaningful adaptation can emerge from the thoughtful alignment of SLA theory, instructional design, and responsive digital infrastructure. The prototype

developed here represents an initial but significant step toward such an integrative model, one in which technology serves pedagogy, and not the reverse.

Looking ahead, future work may further refine the platform's adaptive mechanisms through the integration of natural language processing and data-driven learner modeling, enabling real-time personalization of tasks and feedback. Equally, longitudinal studies with larger and more diverse learner populations would allow for a deeper understanding of how ICALL environments can support sustained language development across different contexts and proficiency levels.

In conclusion, this thesis has demonstrated that it is possible to design a digital learning environment that remains faithful to the principles of second language acquisition while embracing the flexibility and scalability of digital technology. By combining pedagogical insight with technical innovation, the research has contributed both a conceptual framework and a practical implementation that together point toward a more intelligent, adaptive, and human-centered future for computer-assisted language learning.

When this research project was first conceived in 2022, the global conversation around artificial intelligence in education was only beginning to take shape. At that stage, AI (as the expression of Large Language Models) tools were still largely confined to experimental or specialist domains, and their integration into everyday teaching practice remained limited. However, as the project unfolded, we found ourselves amid an accelerating AI revolution. By the end of the research period, artificial intelligence had become a pervasive and transformative force, reshaping not only the technological landscape but also the public imagination of what digital learning could become.

This rapid diffusion of AI inevitably reframed the context of our work. What began as an exploration of adaptive, theory-driven language learning tools gradually came to be surrounded by a new ecosystem of AI-driven applications promising instant personalization, automated assessment, and conversational tutoring. Against this backdrop, the system developed in this research might appear, at least on the surface, more traditional than innovative, emphasizing linguistic structure, learner reflection, and gradual skill development rather than automation or generative interaction. Yet, we argue that this apparent traditionalism is, in fact, a deliberate and meaningful design choice.

The philosophy underpinning this project has been to preserve the pedagogical depth and methodological rigor of established Second Language Acquisition (SLA) traditions while translating them into a digital, data-informed form. The intention was not to replace teachers or to delegate learning entirely to algorithms, but rather to extend the reach of sound pedagogical principles through technology. In this sense, the project stands as a counterpoint to the current wave of AI enthusiasm: it demonstrates that true innovation in educational technology lies not in novelty for its own sake, but in the thoughtful integration of proven methodologies within new technological frameworks.

We remain convinced that this approach, anchoring digital tools in human-centered, pedagogically informed design, will yield more sustainable benefits in the long term. While artificial intelligence offers immense potential for personalization and efficiency, language learning remains an intrinsically human process, grounded in communication, reflection, and cultural understanding. By embedding traditional methods within adaptive digital environments, we preserve these human dimensions while still leveraging technological affordances.

In retrospect, this temporal journey, from the pre-AI boom of 2022 to the fully digitized, AI-saturated context of 2025, has itself been an instructive lesson. It has shown that amid rapid technological change, the core pedagogical questions remain the same: How do learners engage? What motivates them to persist? How can digital environments foster genuine understanding rather than superficial performance? This research suggests that the most meaningful answers to these questions are not found in algorithms alone, but in the dialogue between technology and pedagogy, between innovation and tradition.

References

- Agustin-Llach, M. (2022). How age and L2 proficiency affect the L2 lexicon. *System*, 104, 102697. <https://doi.org/10.1016/j.system.2021.102697>
- Albared, M., Gallofre Ocana, M., Ghareb, A., & Al-Moslmi, T. (2019). Recent Progress of Named Entity Recognition over the Most Popular Datasets. *2019 First International Conference of Intelligent Computing and Engineering (ICOICE)*, 1–9. <https://doi.org/10.1109/ICOICE48418.2019.9035170>
- Aldabe, I., Chiarcos, C., Gracia, J., Roeder, C., & Villegas, M. (2012). Towards a linguistic linked open data cloud: The open linguistics working group. *Linked Open Data—Creating Knowledge Out of Interlinked Data*, 948, 19–25.
- Aldabe, I., Chiarcos, C., Gracia, J., Roeder, C., & Villegas, M. (2014a). The open linguistics working group. *Linked Open Data—Creating Knowledge Out of Interlinked Data*, 948, 19–25.
- Aldabe, I., Chiarcos, C., Gracia, J., Roeder, C., & Villegas, M. (2014b). The open linguistics working group: Developing the linguistic linked open data cloud. *Linked Open Data—Creating Knowledge Out of Interlinked Data*, 948, 19–25.
- Alderson, J. C. (1984). *Reading in a Foreign Language*. Longman.
- AlMousa, M., Benlamri, R., & Khoury, R. (2022). A novel word sense disambiguation approach using WordNet knowledge graph. *Computer Speech & Language*, 74, 101337. <https://doi.org/10.1016/j.csl.2021.101337>
- Altarriba, J., & Mathis, K. M. (1997). Conceptual and Lexical Development in Second Language Acquisition. *Journal of Memory and Language*, 36(4), 550–568. <https://doi.org/10.1006/jmla.1997.2493>
- Alzahrani, S. (2017). Linking Multimedia Vocabulary CALL Research to SLA Cognitive Theories. *American Journal of Educational Research*, 5(7), 821–827. <https://doi.org/10.12691/education-5-7-19>

- Amaral, L. A., & Meurers, D. (2011). On using intelligent computer-assisted language learning in real-life foreign language teaching and learning. *ReCALL*, 23(1), 4–24. <https://doi.org/10.1017/S0958344010000261>
- Amiruddin, A. Z., Ghani, M. T. A., Daud, W. A. A. W., Hussein, N. H., Yamin, N. H. M., & Abdullah, Y. S. (2025). Empowering language learners: Innovations in CEFR-based language learning applications. *Edelweiss Applied Science and Technology*, 9(2), 352–359.
- Amiruddin, M. Z. B., Samsudin, A., Suhandi, A., Coştu, B., & Prahani, B. K. (2025). Scientific mapping and trend of conceptual change: A bibliometric analysis. *Social Sciences & Humanities Open*, 11, 101208. <https://doi.org/10.1016/j.ssaho.2024.101208>
- Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge and reading comprehension. *Journal of Reading Behavior*, 13(3), 233–255.
- Anthony, L. (s.d.). *AntConc: A Learner and Classroom Friendly, Multi-Platform Corpus Analysis Toolkit*.
- Apache Software Foundation. (2025). *Apache JMeter*. <https://jmeter.apache.org/>
- Arani, M. R. T., & Yazdanimoghaddam, M. (2016). The Effects of Input Flooding and Textual Enhancement on the Acquisition of English Passive Voice by Iranian EFL Learners. *Journal of Language Teaching and Research*, 7(4), 802–808.
- Arani, S. G., & Yazdanimoghaddam, M. (2016). The Impact of Input Flooding and Textual Enhancement on Iranian EFL Learners' Syntactic Development. *Studies in Applied Linguistics and TESOL*, 16(1). <https://doi.org/10.7916/SALT.V16I1.1262>
- Armaselu, F., Apostol, E.-S., Khan, A. F., Liebeskind, C., McGillivray, B., Truică, C.-O., Utkă, A., Valūnaitė Oleškevičienė, G., & van Erp, M. (2022). LL(O)D and NLP perspectives on semantic change for humanities research. *IOS Press*.
- Arnold, K. E., & Pistilli, M. D. (2012). Course Signals at Purdue: Using Learning Analytics to Increase Student Success. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 267–270. <https://doi.org/10.1145/2330601.2330666>

- Arnold, N., & Ducate, L. (2006). Integrating technology into language learning. *CALICO Journal*, 23(2), 309–317.
- Arnon, I., & Snider, N. (2010). More than Words: Frequency Effects for Multi-word Phrases. *Journal of Memory and Language*, 62(1), 67–82. <https://doi.org/10.1016/j.jml.2009.09.005>
- Aronoff, M., & Rees-Miller, J. (2001). *The Handbook of Linguistics*. Blackwell.
- Aronoff, M., & Rees-Miller, J. (2003). *The Handbook of Linguistics*. Wiley. <https://books.google.it/books?id=r18LDiR11nYC>
- Auralog. (2002). *TELL Me More: Computer-Assisted Language Learning Program*.
- Austin, J. L. (1962). *How to do things with words*. Oxford University Press. http://scholar.google.de/scholar.bib?q=info:xI2JvixH8_QJ:scholar.google.com/&output=citation&hl=de&as_sdt=0,5&ct=citation&cd=1
- Azab, M., Salama, A., Oflazer, K., Shima, H., Araki, J., & Mitamura, T. (2013). An NLP-based Reading Tool for Aiding Non-native English Readers. *Proceedings of RANLP 2013*, 41–48. <https://aclanthology.org/R13-1006>
- Baccianella, S., Esuli, A., & Sebastiani, F. (2010). SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, & D. Tapias (A c. Di), *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2010/pdf/769_Paper.pdf
- Baddeley, A. (2003). Working Memory: Looking Back and Looking Forward. *Nature Reviews Neuroscience*, 4(10), 829–839.
- Baddeley, A. D. (1992). *Working Memory*. Clarendon Press.
- Baddeley, A. D. (2003). Working Memory and Language: An Overview. *Journal of Communication Disorders*, 36(3), 189–208.
- Bakkalbasi, N., Bauer, K., Glover, J., & Wang, L. (2006). Three options for citation tracking: Google Scholar, Scopus and Web of Science. *Biomedical digital libraries*, 3, 1–8. <https://doi.org/10.1186/1742-5581-3-7>
- Pasquale Esposito

- Bartolini, R., Del Gratta, R., & Frontini, F. (2013). Towards the establishment of a linguistic linked data network for Italian. *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language data*, 76–81. <https://aclanthology.org/W13-5512>
- Bartsch, L. M., & Shepherdson, P. (2023). Chunking, boosting, or offloading? Using serial position to investigate long-term memory's enhancement of verbal working memory performance. *Attention, Perception, & Psychophysics*, 85(5), 1566–1581. <https://doi.org/10.3758/s13414-022-02625-w>
- Basile, P., Caputo, A., & Semeraro, G. (2014). *An Enhanced Lesk Word Sense Disambiguation Algorithm through a Distributional Semantic Model*.
- Batsuren, K., Bella, G., & Giunchiglia, F. (2019). CogNet: A Large-Scale Cognate Database. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3136–3145. <https://doi.org/10.18653/v1/P19-1302>
- Bauer, G. (2005). Hot Potatoes: A Suite of Authoring Tools for Language Teachers. *ReCALL*, 17(1), 139–142.
- Bax, S. (2003). CALL—Past, Present and Future. *System*, 31(1), 13–28.
- Beatty, K. (2010). *Teaching and Researching Computer-Assisted Language Learning* (2nd ed.). Routledge.
- Bechhofer, S., Horrocks, I., & Patel-Schneider, P. F. (2004). *OWL Web Ontology Language Reference*. W3C. <https://www.w3.org/TR/owl-ref/>
- Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing Words to Life: Robust Vocabulary Instruction*. Guilford Press.
- Beckage, N. M., Smith, L. B., & Hills, T. T. (2011). Small worlds and semantic network growth in typical and late talkers. *PLoS ONE*, 6(5), e19348. <https://doi.org/10.1371/journal.pone.0019348>
- Becker, M., Korfhage, K., & Frank, A. (2021). COCO-EX: A Tool for Linking Concepts from Texts to ConceptNet. *Proceedings of the 16th Conference of the European Chapter of the Association*

- for *Computational Linguistics: System Demonstrations*, 119–126.
<https://doi.org/10.18653/v1/2021.eacl-demos.15>
- Béguin, M. (2016). Object Pragmatics and Language Development. *Integrative Psychological and Behavioral Science*, 50(4), 603–620. <https://doi.org/10.1007/s12124-016-9361-7>
- Bell, M. J., & Schäfer, M. (2016). Modelling Semantic Transparency. *Morphology*, 26(2), 157–199. <https://doi.org/10.1007/s11525-016-9286-3>
- Bella, G., Batsuren, K., & Giunchiglia, F. (2021). A Database and Visualization of the Similarity of Contemporary Lexicons. In K. Ekšteín, F. Pártl, & M. Konopík (A c. Di), *Text, Speech, and Dialogue* (Vol. 12848, pp. 95–104). Springer International Publishing. https://doi.org/10.1007/978-3-030-83527-9_8
- Bellezza, F. S. (1981). Mnemonic devices and memory: A review. *Journal of Verbal Learning and Verbal Behavior*, 20(3), 201–216.
- Bergman, M., & Neumark, D. (2017). *Online Dictionaries: Their Role in Vocabulary Learning*. Oxford University Press.
- Bernacki, M. L., Greene, J. A., & Lobczowski, N. G. (2021). A Systematic Review of Research on Personalized Learning: Personalized by Whom, to What, How, and for What Purpose(s)? *Educational Psychology Review*, 33(4), 1675–1715. <https://doi.org/10.1007/s10648-021-09615-8>
- Berners-Lee, T., Hendler, J., & Lassila, O. (2021). The Semantic Web: A New Form of Web Content that is Meaningful to Computers will Unleash a Revolution of New Possibilities. In O. Seneviratne & J. Hendler (A c. Di), *Linking the World's Information* (pp. 91–103). ACM. <https://dl.acm.org/doi/10.1145/3591366.3591376>
- Biber, D. (1988). *Variation across Speech and Writing*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511621024>
- Biber, D. (2012). Register as a predictor of linguistic variation. *Corpus Linguistics and Linguistic Theory*, 8(1), 9–37. <https://doi.org/doi:10.1515/cllt-2012-0002>
- Biber, D., & Conrad, S. (2009). *Register, Genre, and Style*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511814358>

- Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45(1), 5–35. <https://doi.org/10.5054/tq.2011.244483>
- Bilson, S., Yoshida, H., Tran, C. D., Woods, E. A., Hills, T. T., & Smith, L. B. (2015). Semantic network structure in early lexical development: The case of bilingualism. *Frontiers in Psychology*, 6, 1187. <https://doi.org/10.3389/fpsyg.2015.01187>
- Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python*. O'Reilly Media.
- Birdsong, D. (1994a). Decision Making in Second Language Acquisition. *Studies in Second Language Acquisition*, 16(2), 169–182. <https://doi.org/10.1017/S0272263100012869>
- Birdsong, D. (1994b). Input Processing and the Assessment of Second Language Intuitions. *Second Language Research*, 10(1), 65–79.
- Birdsong, D. (2018). Plasticity, Variability and Age in Second Language Acquisition and Bilingualism. *Frontiers in Psychology*, 9, 81. <https://doi.org/10.3389/fpsyg.2018.00081>
- Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., & Becker, C. (2009). DBpedia – A Crystallization Point for the Web of Data. *Journal of Web Semantics*, 7(3), 154–165. <https://doi.org/10.1016/j.websem.2009.01.002>
- Black, P. E. (2021, gennaio). *Ratcliff/Obershelp Pattern Recognition*. <https://www.nist.gov/dads/HTML/ratcliffObershelp.html>
- Blake, R. J. (2000). Computer Mediated Communication: A Window on L2 Spanish Interlanguage. *Language Learning & Technology*, 4(1), 120–136.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Blum-Kulka, S., House, J., & Kasper, G. (1989). *Cross-Cultural Pragmatics: Requests and Apologies*. Ablex.
- Boleda, G. (2020). Distributional Semantics and Linguistic Theory. *Annual Review of Linguistics*, 6(1), 213–234. <https://doi.org/10.1146/annurev-linguistics-011619-030303>

- Bonial, C., Stowe, K., & Palmer, M. (2013). Renewing and Revising SemLink. *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language data*, 9–17. <https://aclanthology.org/W13-5503>
- Bonnefon, J.-F., & Rahwan, I. (2020). Thinking Fast and Slow in Artificial Intelligence. *Nature Machine Intelligence*, 2(6), 292–294.
- Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J., & Yakhnenko, O. (2013). Translating embeddings for modeling multi-relational data. *Advances in neural information processing systems*, 26.
- Borodkin, K. & Others. (2025). Lexical networks in second language learners: A comparison with native speakers. *To Appear*.
- Bosch, A. J. T., Rohm, T. V., AlAsfoor, S., Low, A. J. Y., Baumann, Z., Parayil, N., Noreen, F., Roux, J., Meier, D. T., & Cavelti-Weder, C. (2023). Diesel Exhaust Particle (DEP)-induced Glucose Intolerance is Driven by an Intestinal Innate Immune Response and NLRP3 Activation in Mice. *Particle and Fibre Toxicology*, 20(1), 25. <https://doi.org/10.1186/s12989-023-00536-8>
- Bosque-Gil, J., Gracia, J., & Montiel-Ponsoda, E. (2017). *Towards a Module for Lexicography in OntoLex*.
- Bosque-Gil, J., Gracia, J., Montiel-Ponsoda, E., & Gómez-Pérez, A. (2018). Models to represent linguistic linked data. *Natural Language Engineering*, 24(6), 811–859. <https://doi.org/10.1017/S1351324918000347>
- Branco, A., Rodrigues, J. A., Salawa, M., Branco, R., & Saedi, C. (2020). Comparative Probing of Lexical Semantics Theories for Cognitive Plausibility and Technological Usefulness. *Proceedings of the 28th International Conference on Computational Linguistics*, 4004–4019. <https://doi.org/10.18653/v1/2020.coling-main.354>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brooke, J., Wang, T., & Hirst, G. (2010). Automatic Acquisition of Lexical Formality. In C.-R. Huang & D. Jurafsky (A c. Di), *Coling 2010: Posters* (pp. 90–98). Coling 2010 Organizing Committee. <https://aclanthology.org/C10-2011>

- Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism: Language and Cognition*, 13(3), 359–371. <https://doi.org/10.1017/S1366728909990344>
- Buitelaar, P., Arcan, M., Iglesias, C., Sánchez-Rada, F., & Strapparava, C. (2013). Linguistic Linked Data for Sentiment Analysis. *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language data*, 1–8. <https://aclanthology.org/W13-5502>
- Buitelaar, P., & Chiarcos, C. (2011). Multilingual semantic representation of smart connected world data. *Proceedings of the Workshop on Multilingual Linked Open Data at the 10th International Semantic Web Conference (ISWC 2011)*.
- Buitelaar, P., Declerck, T., Sintek, M., & Racioppa, S. (2010). Leveraging the crowdsourcing of lexical resources for bootstrapping a linguistic data cloud. *International Conference on Language Resources and Evaluation (LREC 2010)*.
- Burns, G. A., Hermjakob, U., & Ambite, J. L. (2016). Abstract Meaning Representations as Linked Data. In P. Groth, E. Simperl, A. Gray, M. Sabou, M. Krötzsch, F. Lecue, F. Flöck, & Y. Gil (A c. Di), *The Semantic Web – ISWC 2016* (pp. 12–20). Springer International Publishing.
- Busuu Ltd. (2008). *Busuu Online Language Learning Platform*.
- Butzkamm, W. (2009). The Language Acquisition Mystique: Tried and Found Wanting. *Forum Sprache*, 2, 83–93.
- Byram, M. (2020). *Teaching and Assessing Intercultural Communicative Competence: Revisited* (2^a ed.). Channel View Publications. <https://doi.org/10.2307/jj.22730614>
- Byram, M. (2021). *Teaching and assessing intercultural communicative competence: Revisited* (2nd edition). Multilingual Matters.
- Calzolari, N. (A c. Di). (2018). *LREC 2018: Eleventh International Conference on Language Resources and Evaluation*. European Language Resources Association.
- Cambridge University Press & Assessment. (2025). *English Vocabulary Profile (A1–C2)*. <https://www.englishprofile.org/wordlists/evp>

- Canale, M., & Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics*, 1(1), 1–47.
- Carlo, G. S. di. (2013). Lexical differences between American and British English: A survey study. *Language design: journal of theoretical and experimental linguistics*, 15, 0061–0075.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation hardback with audio CDs (2): A course book and reference guide*. Cambridge University Press.
- Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching Pronunciation: A Course Book and Reference Guide* (2nd ed.). Cambridge University Press.
- Celikyilmaz, A., Hakkani-Tür, D., Pasupat, P., & Sarikaya, R. (2015, gennaio). *Enriching Word Embeddings Using Knowledge Graph for Semantic Tagging in Conversational Dialog Systems*. <https://www.microsoft.com/en-us/research/publication/enriching-word-embeddings-using-knowledge-graph-for-semantic-tagging-in-conversational-dialog-systems/>
- Cepeda, N. J., Pashler, H., Vul, E., Rohrer, D., Wixted, J. T., & Carpenter, S. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3), 354–380.
- Chang, H.-S., Kuo, T.-L., Chen, Y.-Y., & Ku, L.-W. (2017). A deep learning approach to readability prediction. *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, 165–170. <https://doi.org/10.18653/v1/I17-2030>
- Chapelle, C. A. (2001a). *Computer Applications in Second Language Acquisition*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524681>
- Chapelle, C. A. (2001b). *Computer Applications in Second Language Acquisition: Foundations for Teaching, Testing and Research*. Cambridge University Press.
- Chapelle, C. A. (2005). Computer-assisted language learning. In *Handbook of research in second language teaching and learning* (pp. 743–755). Routledge.
- Chapelle, C. A. (2009). The Relationship Between Second Language Acquisition Theory and Computer-Assisted Language Learning. *The Modern Language Journal*, 93(s1), 741–753. <https://doi.org/10.1111/j.1540-4781.2009.00970.x>

- Cheirmpas, G., Tabatabaei, S. A., Kanoulas, E., & Tsatsaronis, G. (2024). Benchmarking Named Entity Recognition Approaches for Extracting Research Infrastructure Information from Text. In G. Nicosia, V. Ojha, E. La Malfa, G. La Malfa, P. M. Pardalos, & R. Umeton (A c. Di), *Machine Learning, Optimization, and Data Science* (Vol. 14505, pp. 131–141). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-53969-5_11
- Chen, H.-C. (2021). Word usage in context: Alignment effects in L2 written production tasks. *Language Testing*, 38(2), 195–217. <https://doi.org/10.1177/0265532220969483>
- Chen, H.-R., Hsieh, P.-H., & Kinshuk. (2018). The effects of short-term memory and content representation type on mobile language learning. *Journal of Educational Technology & Society*, 21(2), 186–198.
- Chen, H.-R., Wang, S.-C., & Lin, H.-I. (2019). The effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education*, 136, 104–115. <https://doi.org/10.1016/j.compedu.2019.04.005>
- Chen, M., Shi, W., Zhou, B., & Roth, D. (2021). Cross-lingual Entity Alignment with Incidental Supervision. In P. Merlo, J. Tiedemann, & R. Tsarfaty (A c. Di), *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume* (pp. 645–658). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.eacl-main.53>
- Cheng, F., & He, X. (2023). A Review of Input Enhancement and Vocabulary Acquisition Studies (2009–2023). *Journal of Second Language Studies*, 6(2), 203–229.
- Chiarcos, C. (2012a). LL(O)D and NLP perspectives on semantic change for humanities research. *Journal of Language Technology and Computational Linguistics*, 27(1), 21–36.
- Chiarcos, C. (2012b). Ontologies of Linguistic Annotation: Survey and perspectives. *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC)*, 303–310. http://www.lrec-conf.org/proceedings/lrec2012/pdf/911_Paper.pdf
- Chiarcos, C. (2013a). Models to represent linguistic linked data. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 67–72.
- Chiarcos, C. (2013b). Terminology and ontologies. *Data Science Journal*, 12, 150–161.

- Chiarcos, C. (2014a). Language resources and linked data: A practical perspective. *Proceedings of the Language Resources and Evaluation Conference (LREC 2014)*, 1094–1100.
- Chiarcos, C. (2014b). Linguistic linked open data for speech processing. *Journal of the International Phonetic Association*, 44(1), 103–120.
- Chiarcos, C. (2015a). Recent developments for the linguistic linked open data infrastructure. *Journal of Data Mining and Digital Humanities*, 3(1), 51–63.
- Chiarcos, C. (2015b). Review of research on synonym equivalence relation mining. *Journal of Knowledge Engineering Review*, 30(3), 148–159.
- Chiarcos, C. (2015c). When linguistics meets web technologies. Recent advances in modelling linguistic linked data. *Journal of Web Semantics*, 32, 3–8.
- Chiarcos, C., Cimiano, P., Declerck, T., & McCrae, J. P. (2013a). Linguistic Linked Open Data (LLOD). Introduction and Overview. *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL): Representing and linking lexicons, terminologies and other language data*, 1-. <https://aclanthology.org/W13-5501>
- Chiarcos, C., Cimiano, P., Declerck, T., & McCrae, J. P. (A c. Di). (2013b). *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013)*. Association for Computational Linguistics. <https://aclanthology.org/W13-5500>
- Chiarcos, C., Cimiano, P., Declerck, T., & McCrae, J. P. (A c. Di). (2013c). *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language data*. Association for Computational Linguistics. <https://aclanthology.org/W13-5500>
- Chiarcos, C., Fäth, C., Renner-Westermann, H., Abromeit, F., & Dimitrova, V. (2016). Lin|gu|is|tik: Building the linguist's pathway to bibliographies, libraries, language resources and linked open data. *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC)*, 4463–4471. <https://aclanthology.org/L16-1707>
- Chiarcos, C., Gurevych, I., & Rölz, R. (2015). Wikidata and DBpedia: A Comparative Study. *Linked Open Data—Creating Knowledge Out of Interlinked Data*, 948, 11–18.

- Chiarcos, C., Hellmann, S., & Nordhoff, S. (2011). Get! Mimetypes! Right! *Proceedings of the 10th International Semantic Web Conference (ISWC 2011), Bonn, Germany, October*.
- Chiarcos, C., Hellmann, S., & Nordhoff, S. (2012). Linking Linguistic Resources: Examples from the Open Linguistics Working Group. In *Linked Data in Linguistics* (pp. 201–216). https://doi.org/10.1007/978-3-642-28249-2_19
- Chiarcos, C., Hellmann, S., Nordhoff, S., Moran, S., Littauer, R., Eckle-Kohler, J., Gurevych, I., Hartmann, S., Matuschek, M., & Meyer, C. M. (2012). The Open Linguistics Working Group. In N. Calzolari, K. Choukri, T. Declerck, M. U. Doğan, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, & S. Piperidis (A c. Di), *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)* (pp. 3603–3610). European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2012/pdf/912_Paper.pdf
- Chiarcos, C., & Henrich, V. (2012). TermitUp: Generation and enrichment of linked terminologies. *Proceedings of the 24th International Conference on Computational Linguistics (COLING 2012)*, 513–522.
- Chiarcos, C., & Nordhoff, S. (2012a). Linghub: A linked data based portal supporting the discovery of language resources. *Proceedings of the ACL 2012 System Demonstrations*, 81–86.
- Chiarcos, C., & Nordhoff, S. (2012b). Linguistic linked data: Paving the way towards maximising (Re)usability of linguistic resources. *Proceedings of the 2012 Joint Workshop on Multilingual OCR and Analytics for Noisy Unstructured Text Data*, 19–26.
- Chiarcos, C., & Nordhoff, S. (2013). Linguistic Linked Data: Representation, Generation and Applications. *Data Science Journal*, 12, 126–149.
- Chiarcos, C., & Nordhoff, S. (2015a). Editorial of the Special Issue on Latest Advancements in Linguistic Linked Data. *Journal of Web Semantics*, 32, 1–2.
- Chiarcos, C., & Nordhoff, S. (2015b). Observing LOD: Its Knowledge Domains and the Varying Behavior of Ontologies Across Them. *Journal of Web Semantics*, 32, 18–29.
- Chiarcos, C., Nordhoff, S., & Hellmann, S. (2010). Addressing structural and linguistic heterogeneity in the Web. *Proceedings of the 8th International Conference on Semantic Systems (I-Semantics 2012)*.

- Chinkina, M., & Meurers, D. (2016a). Linguistically Aware Information Retrieval for Language Learners. *Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications (BEA)*, 7–16.
- Chinkina, M., & Meurers, D. (2016b). Linguistically Aware Information Retrieval: Providing Input Enrichment for Second Language Learners. *Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications*, 188–198. <https://doi.org/10.18653/v1/W16-0521>
- Chomsky, N. (1959). A Review of B. F. Skinner's Verbal Behavior. *Language*, 35(1), 26–58. <https://doi.org/10.2307/411334>
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press.
- Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa Lectures*. Foris Publications.
- Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin, and Use*. Praeger.
- Chomsky, N. (1995). *The Minimalist Program*. MIT Press.
- Christian M. Meyer, I. G. (2012). Wiktionary: A new rival for expert-built lexicons? Exploring the possibilities of collaborative lexicography. In *Electronic Lexicography*. Oxford: Oxford University Press., <https://doi.org/10.1093/acprof:oso/9780199654864.003.0013>
- Chrome Developers. (2025). *Workbox: Libraries for Progressive Web Apps*. <https://developer.chrome.com/docs/workbox/>
- Cimiano, P., Gracia, J., Villegas, M., Gómez-Pérez, A., & Bel, N. (2018). The Apertium Bilingual Dictionaries on the Web of Data. *Semant. Web*, 9(2), 231–240. <https://doi.org/10.3233/SW-170258>
- Cinganotto, L., Sbardella, T., & Montanucci, G. (2024). *The Journal of Language and Teaching Technology*.
- Clark, E. V. (1999). *The Lexicon in Acquisition*. Cambridge University Press.
- Cobb, T. (2007). Computing the Vocabulary Demands of L2 Reading. *Language Learning & Technology*, 11(3), 38–63.

- Coleman, M., & Liao, T. L. (1975). A Computer Readability Formula Designed for Machine Scoring. *Journal of Applied Psychology*, 60(2), 283–284. <https://doi.org/10.1037/h0076540>
- Collins, A. M., & Quillian, M. R. (1969). Retrieval Time from Semantic Memory. *Journal of Verbal Learning and Verbal Behavior*, 8(2), 240–247.
- Co-operation, O. for E., & Development (OECD). (2019). *Getting Skills Right: Future-Ready Adult Learning Systems*. OECD Publishing. <https://doi.org/10.1787/9789264311756-en>
- Corbett, A. T., & Anderson, J. R. (1995). Knowledge Tracing: Modeling the Acquisition of Procedural Knowledge. *User Modeling and User-Adapted Interaction*, 4(4), 253–278. <https://doi.org/10.1007/BF01099821>
- Coronel, J. C., & Federmeier, K. D. (2016). The N400 Reveals How Personal Semantics Is Special. *Journal of Cognitive Neuroscience*, 28(11), 1837–1854.
- COST Action IS1305. (2013). *COST Action «European network for Web-centred linguistic data science» (NexusLinguarum); [Acción COST «Red europea para la ciencia de datos lingüísticos centrada en la web» (NexusLinguarum)]*.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press. <https://www.coe.int/en/web/common-european-framework-reference-languages>
- Council of Europe. (2020a). *Common European Framework of Reference for Languages: Companion Volume*. <https://www.coe.int/en/web/common-european-framework-reference-languages/companion-volume>
- Council of Europe. (2020b). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment – Companion Volume*. Council of Europe Publishing. <https://www.coe.int/en/web/common-european-framework-reference-languages/companion-volume>
- Council of Europe. (2020c). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume*. Council of Europe Publishing.

- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34(2), 213.
<https://doi.org/10.2307/3587951>
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of Processing: A Framework for Memory Research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684.
- Crossley, S. A., Salsbury, T., & McNamara, D. S. (2014). Assessing lexical proficiency using analytic ratings: A case for collocation accuracy. *Applied Linguistics*, 36(5), 570–590.
<https://doi.org/10.1093/applin/amt056>
- Cruse, A. (2004). *Meaning in Language: An Introduction to Semantics and Pragmatics*. Oxford University Press UK.
- Cruse, D. A. (2000). *Meaning in Language: An Introduction to Semantics and Pragmatics*. Oxford University Press.
- Cruse, D. A. (2004). *Meaning in Language: An Introduction to Semantics and Pragmatics*. Oxford University Press. <https://books.google.it/books?id=K6uZIXQYoowC>
- Cypress.io. (2025). *Cypress End-to-End Testing: Guides*. <https://docs.cypress.io/>
- Dabrowska, E. (2019). Experience, Aptitude, and Individual Differences in Native Language Ultimate Attainment. *Language Learning*, 69(Suppl 1), 133–152.
- Danet, B., & Herring, S. C. (2007). *The Multilingual Internet: Language, Culture, and Communication Online*. Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780195304794.001.0001>
- Davis, M. H., & Gaskell, G. M. (2009). A Complementary Systems Account of Word Learning: Neural and Behavioural Evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1536), 3773–3800.
- De Bot, K., Lowie, W., & Verspoor, M. (2005). *Second Language Acquisition: An Advanced Resource Book*. Routledge.
- De Bot, K., Lowie, W., & Verspoor, M. (2007). A Dynamic Systems Theory Approach to Second Language Acquisition. *Bilingualism: Language and Cognition*, 10(1), 7–21.

- de Bot, K., Lowie, W., & Verspoor, M. (2007). *Second language acquisition: An advanced resource book*. Routledge.
- De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2021). The “Small World of Words” English Word Association Norms for Over 12,000 Cue Words. *Behavior Research Methods*, 53(2), 733–746.
- De Groot, A. M. B., & Keijzer, R. (2000). What Is Hard to Learn Is Easy to Forget: The Roles of Word Concreteness, Cognate Status, and Word Frequency in Foreign-Language Vocabulary Learning and Forgetting. *Language Learning*, 50(1), 1–56. <https://doi.org/10.1111/0023-8333.00110>
- De Mauro, T. (1999). *GRADIT: Grande dizionario italiano dell’uso*. UTET.
- De Mauro, T. (2016). *Il Nuovo Vocabolario di Base della Lingua Italiana*. <https://www.internazionale.it/opinione/tullio-de-mauro/2016/12/23/il-nuovo-vocabolario-di-base-della-lingua-italiana>
- De Mauro, T. (2019). *Guida all’uso delle parole: Parlare e scrivere semplice e preciso per capire e farsi capire* (1^a ed.). GLF Editori Laterza.
- De Mauro, T., & Chiari, I. (2016). *Il Nuovo Vocabolario di Base della Lingua Italiana*. Internazionale / Internazionale Edizioni.
- Debattista, J., Auer, Sö., & Lange, C. (2016). Luzzu-A Methodology and Framework for Linked Data Quality Assessment. *Journal of Data and Information Quality*, 8(1). <https://doi.org/10.1145/2992786>
- Deci, E. L., & Ryan, R. M. (2014). *Intrinsic motivation and self-determination in human behavior*. Springer Science+Business Media.
- DeKeyser, R. (2007). Skill Acquisition Theory. In B. VanPatten & J. Williams (A c. Di), *Theories in Second Language Acquisition* (pp. 97–113). Lawrence Erlbaum Associates.
- DeKeyser, R. M. (2000). The Robustness of Critical Period Effects in Second Language Acquisition. In *Studies in Second Language Acquisition* (Vol. 22, pp. 499–533). Cambridge University Press.

- DeKeyser, R. M. (2003). Implicit and Explicit Learning. In C. J. Doughty & M. H. Long (A c. Di), *The Handbook of Second Language Acquisition* (pp. 313–348). Blackwell.
- Dell’Orletta, F., Montemagni, S., & Venturi, G. (2011). READ–IT: Assessing Readability of Italian Texts with a View to Text Simplification. *Proceedings of the Second Workshop on Speech and Language Processing for Assistive Technologies*, 73–83. <https://aclanthology.org/W11-2308/>
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation Fundamentals: Evidence-based Perspectives for L2 Teaching and Research*. John Benjamins.
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining «Gamification». *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of NAACL-HLT*, 4171–4186. <https://doi.org/10.48550/arXiv.1810.04805>
- di Carlo, G. S. (2013). Lexical differences between American and British English: A survey study. *Language Design: Journal of Theoretical and Experimental Linguistics*, 15, 61–75.
- Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5(3), 175–197. Cambridge Core. <https://doi.org/10.1017/S1366728902003012>
- Doughty, C., & Williams, J. (A c. Di). (1998). *Focus on Form in Classroom Second Language Acquisition*. Cambridge University Press.
- Du, J., Way, A., & Zydron, A. (2016). Using BabelNet to Improve OOV Coverage in SMT. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC)*, 9–15. <https://aclanthology.org/L16-1002>
- Dulay, H. C., & Burt, M. K. (1974). You Can’t Learn Without Goofing: An Analysis of Children’s Second Language Errors. *TESOL Quarterly*, 8(2), 129–136. <https://doi.org/10.2307/3585536>
- Dybkjær, L., Bernsen, N. O., & Dybkjær, H. (2005). Evaluation of Spoken Dialogue Systems. In *Spoken Multimodal Human-Computer Dialogue in Mobile Environments* (pp. 237–274). Springer.

- Eder, E., Krieg-Holz, U., & Hahn, U. (2021). Acquiring a Formality-Informed Lexical Resource for Style Analysis. In P. Merlo, J. Tiedemann, & R. Tsarfaty (A c. Di), *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume* (pp. 2028–2041). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.eacl-main.174>
- Egbert, J. (2005). *CALL Essentials: Principles and Practice in CALL Classrooms*. TESOL Publications.
- Egbert, J., Chao, C.-C., & Hanson-Smith, E. (1999). Computer-Enhanced Language Learning Environments: An Overview. In J. Egbert & E. Hanson-Smith (A c. Di), *CALL Environments: Research, Practice, and Critical Issues* (pp. 1–13). TESOL Publications.
- Ehrmann, M., Hamdi, A., Pontes, E. L., Romanello, M., & Doucet, A. (2024). Named Entity Recognition and Classification in Historical Documents: A Survey. *ACM Computing Surveys*, 56(2), 1–47. <https://doi.org/10.1145/3604931>
- Elbedweihi, K. (2013, ottobre). *Using BabelNet in Bridging the Gap Between Natural Language Queries and Linked Data Concepts*.
- Elliot, J., & Mitkov, R. (1997). ELLIPSIS: An Intelligent Tutoring System for Second Language Acquisition. *Proceedings of the 1997 International Conference on Artificial Intelligence in Education*, 432–439.
- Ellis, N. C. (2002). Frequency Effects in Language Processing: A Review with Implications for Theories of Implicit and Explicit Language Acquisition. *Studies in Second Language Acquisition*, 24(2), 143–188.
- Ellis, N. C. (2005). At the Interface: Dynamic Interactions of Explicit and Implicit Language Knowledge. In *Studies in Second Language Acquisition* (Vol. 27, pp. 305–352). Cambridge University Press.
- Ellis, N. C. (2006). Language Acquisition as Rational Contingency Learning. *Applied Linguistics*, 27(1), 1–24.

- Ellis, N. C., & Larsen-Freeman, D. (2006). Language Emergence: Implications for Applied Linguistics--Introduction to the Special Issue. *Applied Linguistics*, 27(4), 558–589. <https://doi.org/10.1093/applin/aml028>
- Ellis, N. C., & Schmidt, R. (1997). Morphology and Longer Distance Dependencies: Laboratory Research Illuminating the A in SLA. *Studies in Second Language Acquisition*, 19(2), 145–171.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford University Press. <https://books.google.it/books?id=3KglibyrZ5sC>
- Ellis, R. (2001). Investigating Form-Focused Instruction. In C. N. Candlin & N. Mercer (A c. Di), *Learning and Teaching Languages: A Lifelong Learning Perspective* (pp. 142–169). Multilingual Matters.
- Ellis, R. (2003). *Task-Based Language Learning and Teaching*. Oxford University Press.
- Ellis, R. (2005a). Planning and Task-Based Performance: Theory and Research. In R. Ellis (Ed.), *Planning and Task Performance in a Second Language*, 3–34.
- Ellis, R. (2005b). Principles of Instructed Language Learning. *System*, 33(2), 209–224.
- Ellis, R. (2006). *The Study of Second Language Acquisition*. Oxford University Press.
- Ellis, R. (2008). *The Study of Second Language Acquisition* (2nd ed.). Oxford University Press.
- Ellis, R. (2009). *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching*. Multilingual Matters.
- Ellis, R. (2010). Second Language Acquisition, Teacher Education and Language Pedagogy. *Language Teaching*, 43(2), 182–201.
- Ellis, R. (2017). Task-Based Language Teaching: Sorting Out the Misunderstandings. *International Journal of Applied Linguistics*, 27(2), 386–396.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and Explicit Corrective Feedback and the Acquisition of L2 Grammar. *Studies in Second Language Acquisition*, 28(2), 339–368. <https://doi.org/10.1017/S0272263106060141>

- Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. Psychology Press.
- Erling, O., Alexander, K., & Coetzee, S. (2012). Characterizing RDF graphs through graph-based measures—Framework and assessment. *Web Semantics: Science, Services and Agents on the World Wide Web*, 12, 27–43.
- Eskenazi, M. A., & Folk, J. R. (2015). Reading Skill and Word Skipping: Implications for Visual and Linguistic Accounts of Word Skipping. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(6), 1923–1928. <https://doi.org/10.1037/xlm0000156>
- Esposito, P., Pellegrino, M. A., Scarano, V., & TuoZZo, G. (2024, agosto). *The Linguistic Linked Open Data Cloud: Phenomenal Cosmic Powers... Itty Bitty Quality Space!* (Versione 1.0). Zenodo. <https://doi.org/10.5281/zenodo.13449868>
- European Parliament and Council. (2016). *Regulation (EU) 2016/679 (GDPR)*. <https://eur-lex.europa.eu/eli/reg/2016/679/oj>
- Evans, J. S. B. T. (2008). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology*, 59, 255–278.
- Fang, F., & Jiang, L. (2020). Textual Enhancement and the Acquisition of English Collocations by Chinese EFL Learners. *Language Teaching Research*, 24(6), 757–776.
- Fazlali, B., & Shahini, G. (2019). Textual Enhancement: Effects on Iranian EFL Learners' Grammatical and Lexical Collocation Knowledge. *Journal of Language and Linguistic Studies*, 15(3), 1121–1136.
- Fellbaum, C. (A. c. Di). (1998). *WordNet: An Electronic Lexical Database*. MIT Press. <https://doi.org/10.7551/mitpress/7287.001.0001>
- Fiorelli, M., Pazienza, M. T., & Stellato, A. (2013). LIME: Towards a Metadata Module for Ontolex. *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language data*, 8–27. <https://aclanthology.org/W13-5504>
- Flege, J. E. (1995). Second Language Speech Learning: Theory, Findings, and Problems. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 233–277.

- Fleming, N., Motamedinia, K., & Reynolds, D. (2017). Learning Analytics for Language Learning and Teaching. *CALICO Journal*, 34(1), 1–15.
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32(3), 221–233. <https://doi.org/10.1037/h0057532>
- Freeman, D. E. L. (1975). The Acquisition of Grammatical Morphemes by Adult ESL Students. *TESOL Quarterly*, 9(4), 409. <https://doi.org/10.2307/3585625>
- Futia, G., & Vetrò, A. (2020). On the Integration of Knowledge Graphs into Deep Learning Models for a More Comprehensible AI—Three Challenges for Future Research. *Information*, 11(2). <https://doi.org/10.3390/info11020122>
- Gamper, J., & Knapp, P. (2002). A Review of Intelligent CALL Systems. *Computer Assisted Language Learning*, 15(4), 329–342.
- Gangemi, A., Alam, M., & Presutti, V. (2016). Word Frame Disambiguation: Evaluating Linguistic Linked Data on Frame Detection. *LD4IE@ISWC*. <https://api.semanticscholar.org/CorpusID:11945337>
- Gass, S. M., Behney, J., Plonsky, L., & Selinker, L. (2008). *Second Language Acquisition* (0 ed.). Routledge. <https://doi.org/10.4324/9780203932841>
- Gass, S. M., & Mackey, A. (2000). *Stimulated Recall Methodology in Second Language Research*. Lawrence Erlbaum Associates.
- Gass, S. M., & Mackey, A. (2006). Input, Interaction and Output: An Overview. *AILA Review*, 19(1), 3–17.
- Gass, S. M., & Mackey, A. (2015). Input, Interaction, and Output in Second Language Acquisition. In B. VanPatten & J. Williams (A c. Di), *Theories in Second Language Acquisition: An Introduction* (pp. 180–206). Routledge.
- Gass, S. M., & Varonis, E. M. (1985). Task Variation and Non-Native/Nonnative Negotiation of Meaning. *Second Language Research*, 1(2), 94–113. <https://doi.org/10.1177/026765838500100202>

- Ghafouri, F., & Masoomi, E. (2016). The Effect of Textual Enhancement on Vocabulary Acquisition of Iranian EFL Learners. *International Journal of Applied Linguistics & English Literature*, 5(4), 199–205.
- Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *Academy of Management Annals*, 14(2), 627–660.
- Godwin-Jones, R. (2008). Emerging Technologies: Mobile Computing Trends: Lighter, Faster, Smarter. *Language Learning & Technology*, 12(3), 3–9.
- Godwin-Jones, R. (2016). Emerging Technologies: Augmented Reality and Language Learning. *Language Learning & Technology*, 20(3), 9–19.
- Godwin-Jones, R. (2018). Emerging Technologies: Second Language Vocabulary Learning and Mobile Devices. *Language Learning & Technology*, 22(3), 2–11.
- Goldschneider, J. M., & DeKeyser, R. M. (2001). Explaining the “Natural Order of L2 Morpheme Acquisition” in English: A Meta-analysis of Multiple Determinants. *Language Learning*, 51(1), 1–50. <https://doi.org/10.1111/1467-9922.00147>
- Google Design. (2025). *Material Design 3*. <https://m3.material.io/>
- Goossens, T. (2014). Complementary Learning Systems and Word Learning: A Developmental Perspective. *Frontiers in Psychology*, 5, 1260.
- Grabe, W., & Stoller, F. L. (2020). *Teaching and researching reading* (Third edition). Routledge, Taylor & Francis Group.
- Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher*, 40(5), 223–234. <https://doi.org/10.3102/0013189X11413260>
- Grainger, J., & Dijkstra, T. (1992). On the Representation and Use of Language Information in Bilinguals. In R. J. Harris (A c. Di), *Advances in Psychology* (Vol. 83, pp. 207–220). North-Holland. [https://doi.org/10.1016/S0166-4115\(08\)61496-X](https://doi.org/10.1016/S0166-4115(08)61496-X)
- Gromann, D., Apostol, E. S., Chiarcos, C., Cremaschi, M., Gracia, J., Gkirtzou, K., Liebeskind, C., Mockiene, L., Rosner, M., Schuurman, I., Sérasset, G., Silvano, P., Spahiu, B., Truică, C.-O., Utku, Pasquale Esposito

- A., & Valunaite Oleskeviciene, G. (2024). Multilinguality and LLOD: A survey across linguistic description levels. *Semantic Web*, 1–44. <https://doi.org/10.3233/SW-243591>
- Gromann, D., Apostol, E.-S., Chiarcos, C., Cremaschi, M., Gracia, J., Gkirtzou, K., Liebeskind, C., Mititelu, V., Mockiene, L., Rosner, M., Schuurman, I., Sérasset, G., Silvano, P., Spahiu, B., Truic, C.-O., Utko, A., & Oleskeviciene, G. V. (2022). Multilinguality and LLOD: A Survey Across Linguistic Description Levels. *Semantic Web 1 (0)* 1–39, IOS Press.
- Gunning, R. (1952). *The Technique of Clear Writing*. McGraw-Hill.
- Guo, M., Chen, Y., Xu, J., & Zhang, Y. (2022). Dynamic Knowledge Integration for Natural Language Inference. *2022 4th International Conference on Natural Language Processing (ICNLP)*, 360–364. <https://doi.org/10.1109/ICNLP55136.2022.00066>
- Guo, T. (2023). A Review of the Effects of Textual Enhancement on Second Language Learning. *International Journal of Linguistics, Literature and Translation*, 6(1), 107–111. <https://doi.org/10.32996/ijllt.2023.6.1.15>
- Gurevych, I., Hartmann, S., & Matuschek, M. (2011). Harmonizing Lexical Data for their Linking to Knowledge Objects in the Linked Data Framework. *Proceedings of the 5th Global Wordnet Conference*.
- Half-Baked Software, Inc., & Centre, U. of V. H. C. & M. (2022). *Hot Potatoes Home Page*. <https://hotpot.uvic.ca/>
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.
- Hampel, R., & Stickler, U. (2005). New Skills for New Classrooms: Training Tutors to Teach Languages Online. *Computer Assisted Language Learning*, 18(4), 311–326.
- Hampel, R., & Stickler, U. (2012). The Use of Videoconferencing to Support Multimodal Interaction in an Online Language Classroom. *ReCALL*, 24(2), 116–137.
- Han, Z., Park, E. S., & Combs, C. (2008). Textual Enhancement of Input: Issues and Possibilities. *Applied Linguistics*, 29(4), 597–618.
- Harris, Z. S. (1954). *Distributional Structure* (Vol. 10). Word.

- Hawkins, R. (2001). *Second Language Syntax: A Generative Introduction*. Blackwell.
- Hayashi, Y. (2013). Migrating Psycholinguistic Semantic Feature Norms into Linked Data in Linguistics. *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language data*, 70–75. <https://aclanthology.org/W13-5511>
- Hayet, L. (1994). The French Learner's Workbench: Multimedia Resources for Language Learning. *CALICO Journal*, 11(4), 75–90.
- Heift, T., & Schulze, M. (2007). *Errors and intelligence in computer-assisted language learning: Parsers and pedagogues*. Routledge, Taylor & Francis Group. <https://doi.org/10.4324/9780203012215>
- Heift, T., & Schulze, M. (2015). *Errors and Intelligence in Computer-Assisted Language Learning: Parsers and Pedagogues*. Routledge.
- Hellmann, S., Chiarcos, C., & Nordhoff, S. (2012). Word frame disambiguation: Evaluating linguistic linked data on frame detection. *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2012)*, 3647–3653.
- Henderson, L. M., Weighall, A. R., Brown, H., & Gaskell, G. M. (2013). Consolidation of Vocabulary Is Associated with Sleep in Children. *Developmental Science*, 16(5), 791–801.
- Henrich, V., Chiarcos, C., Beck, K., Witt, A., & Thomsen, T. (2014). Knowledge driven intelligent survey systems for linguists. *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*.
- Hernandez, T. A. (2008). The Effect of Explicit Instruction and Input Flood on the Acquisition of Spanish Discourse Markers. *Hispania*, 91(3), 665–675.
- Heuss, T. (2013). Faust.rdf—Taking RDF literally. *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language data*, 59–63. <https://aclanthology.org/W13-5509>
- Heylighen, F., Dewaele, J.-M., & Apostel, L. (1999). *Formality of Language: Definition, measurement and behavioral determinants*. <https://api.semanticscholar.org/CorpusID:16450928>

- Hou, I., Katinskaia, A., & Yangarber, R. (2022). Gamification in Intelligent Computer-Assisted Language Learning: Evaluating User Engagement in the Revita System. *Proceedings of the 13th Language Resources and Evaluation Conference (LREC 2022)*, 5261–5270.
- Hu, M.-T., & Nation, P. (2000). Unknown Vocabulary Density and Reading Comprehension. *Reading in a Foreign Language*, 13.
- Huang, L., Sun, C., Qiu, X., & Huang, X. (2020). *GlossBERT: BERT for Word Sense Disambiguation with Gloss Knowledge* (arXiv:1908.07245). arXiv. <https://doi.org/10.48550/arXiv.1908.07245>
- Hulstijn, J. H. (1992). Retention of Inferred and Given Word Meanings: Experiments in Incidental Vocabulary Learning. In P. J. L. Arnaud & H. Béjoint (A c. Di), *Vocabulary and Applied Linguistics* (pp. 113–125). Palgrave Macmillan UK. https://doi.org/10.1007/978-1-349-12396-4_11
- Hulstijn, J. H. (1993). When Do Foreign-Language Readers Look up the Meaning of Unfamiliar Words? The Influence of Task and Learner Variables. *The Modern Language Journal*, 77(2), 139–147.
- Hulstijn, J. H. (2001a). Intentional and Incidental Second Language Vocabulary Learning: A Reappraisal of Elaboration and Processing. In P. Robinson (A c. Di), *Cognition and Second Language Instruction* (pp. 258–286). Cambridge University Press.
- Hulstijn, J. H. (2001b). Intentional and Incidental Second-Language Vocabulary Learning: A Reappraisal of Elaboration, Rehearsal and Automaticity. In P. Robinson (A c. Di), *Cognition and Second Language Instruction* (pp. 258–286). Cambridge University Press.
- Hutchinson, T., & Waters, A. (2008). *English for Specific Purposes: A Learning-Centered Approach* (23rd printing). Cambridge University Press.
- Hyland, K. (2011). *English for Academic Purposes: An Advanced Resource Book*. Routledge.
- Ishihara, N., & Cohen, A. D. (2010). *Teaching and Learning Pragmatics: Where Language and Culture Meet*. Pearson Education.
- Izumi, S. (2002). Output, Input Enhancement, and the Noticing Hypothesis: An Experimental Study on ESL Relativization. *Studies in Second Language Acquisition*, 24(4), 541–577.

- James, E., Gaskell, M. G., Murphy, G., Tulip, J., & Henderson, L. M. (2023). Word learning in the context of semantic prior knowledge: Evidence of interference from feature-based neighbours in children and adults. *Language, Cognition and Neuroscience*, 38(2), 157–174. <https://doi.org/10.1080/23273798.2022.2102198>
- James, E. L. M., Gaskell, G. M., & Henderson, L. M. (2019). Consolidation of Word Learning through Sleep Depends on Task Demands. *Language Learning*, 69, 47–65.
- James, E. L. M., Henderson, L. M., & Gaskell, G. M. (2023). Encoding and Consolidation in Word Learning: Differential Effects of Repeated Exposure on Word-Form and Meaning Knowledge. *Journal of Memory and Language*, 132, 104–118.
- James, L. M., Gaskell, G. M., & Henderson, L. M. (2019). The impact of sleep on vocabulary acquisition in adults. *Cognitive Psychology*, 108, 1–20.
- James, L. M., Gaskell, G. M., & Henderson, L. M. (2020). Word learning and consolidation: What can we learn from complementary learning systems theory? *Psychonomic Bulletin & Review*, 27(1), 1–15.
- Jaro, M. A. (1989). Advances in Record-Linkage Methodology as Applied to Matching the 1985 Census of Tampa, Florida. *Journal of the American Statistical Association*, 84(406), 414–420. <https://doi.org/10.1080/01621459.1989.10478785>
- Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic Influence in Language and Cognition* (0 ed.). Routledge. <https://doi.org/10.4324/9780203935927>
- Jehangir, B., Radhakrishnan, S., & Agarwal, R. (2023). A survey on Named Entity Recognition—Datasets, tools, and methodologies. *Natural Language Processing Journal*, 3, 100017. <https://doi.org/10.1016/j.nlp.2023.100017>
- Jiang, H., Zhang, D., Cao, T., Yin, B., & Zhao, T. (2021). Named Entity Recognition with Small Strongly Labeled and Large Weakly Labeled Data. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 1775–1789. <https://doi.org/10.18653/v1/2021.acl-long.140>

- John P. McCrae, C. D. F., & Cimiano, P. (2014). *Publishing and Linking WordNet using lemon and RDF*. <https://api.semanticscholar.org/CorpusID:60275443>
- Johnson, J. S., & Newport, E. L. (1989a). Critical Period Effects in Second Language Learning. *Cognitive Psychology*, 21(1), 60–99.
- Johnson, J. S., & Newport, E. L. (1989b). Critical Period Effects in Second Language Learning: The Influence of Maturational State on the Acquisition of English as a Second Language. *Cognitive Psychology*, 21(1), 60–99. [https://doi.org/10.1016/0010-0285\(89\)90003-0](https://doi.org/10.1016/0010-0285(89)90003-0)
- Jones, C., & Waller, D. (2017). *Understanding Fluency: Textual and Auditory Enhancement in Second Language Vocabulary Learning*. Routledge.
- Junczys-Dowmunt, M., Grundkiewicz, R., Dwojak, T., Hoang, H., Heafield, K., Neckermann, T., Seide, F., Hermann, U., Aji, A. F., Bogoychev, N., Fankem, A., & Birch, A. (2018). Marian: Fast Neural Machine Translation in C++. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 116–121. <https://doi.org/10.18653/v1/P18-4020>
- Jurafsky, D., & Martin, J. H. (2024). *Speech and Language Processing* (3rd ed., draft). Prentice Hall. <https://web.stanford.edu/~jurafsky/slp3/>
- Kahneman, D. (1973). *Attention and Effort*. Prentice-Hall.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
- Karipke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, 319(5865), 966–968.
- Katinskaia, A., & Yangarber, R. (2018). Revita: A Language-Learning Platform at the Intersection of ITS and CALL. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 3229–3236.
- Katinskaia, A., & Yangarber, R. (2021a). Assessing Grammatical Correctness in Language Learning. *Proceedings of the 16th Workshop on Innovative Use of NLP for Building Educational Applications*, 135–146. <https://doi.org/10.18653/v1/2021.bea-1.15>

- Katinskaia, A., & Yangarber, R. (2021b). Assessing Grammatical Correctness in Language Learning: Evaluating Alternative-Correct Answers in Cloze Exercises. *Proceedings of the 16th Workshop on Innovative Use of NLP for Building Educational Applications (BEA)*, 243–252.
- Keefe, J. W. (1979). *Learning Style: An Overview*. National Association of Secondary School Principals.
- Kenett, Y. N., Anaki, D., & Faust, M. (2013). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience*, 7, 787. <https://doi.org/10.3389/fnhum.2013.00787>
- Keraghel, I., Morbieu, S., & Nadif, M. (2024, dicembre). *Recent Advances in Named Entity Recognition: A Comprehensive Survey and Comparative Study*. arXiv. <https://doi.org/10.48550/arXiv.2401.10825>
- Kerr, B. (2014). *Personalized Learning: A Guide for Educators*. Jossey-Bass.
- Khan, A. F., Díaz-Vera, J. E., & Monachini, M. (2016). Representing Polysemy and Diachronic Lexico-Semantic Data on the Semantic Web? *SWASH@ESWC*. <https://api.semanticscholar.org/CorpusID:16564118>
- Khan, F., Boschetti, F., & Frontini, F. (2014, maggio). *Using lemon to Model Lexical Semantic Shift in Diachronic Lexical Resources*.
- Kincaid, J. P., Fishburne Jr., R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy enlisted personnel* (Research Branch Report 8-75). Naval Technical Training Command, Millington, TN.
- Kılıçkaya, F. (2007). Website review: WordChamp: Learn Language Faster. *Educational Technology & Society*, 10(4), 298–299.
- Knowles, M. S., Holton Iii, E. F., & Swanson, R. A. (2014). *The Adult Learner* (0 ed.). Routledge. <https://doi.org/10.4324/9781315816951>
- Koide, S., & Takeda, H. (2013). RDFization of Japanese Electronic Dictionaries and LOD. *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and*

- linking lexicons, terminologies and other language data*, 64–69. <https://aclanthology.org/W13-5510>
- Konstantinidis, S. (2007). Computing the Edit Distance of a Regular Language. *Information and Computation*, 205(9), 1307–1316. <https://doi.org/10.1016/j.ic.2007.06.001>
- Koprowski, M. (2005). Investigating the Usefulness of Lexical Phrases in Contemporary Coursebooks. *ELT Journal*, 59(4), 322–332.
- Krashen, S. D. (1981). *Second Language Acquisition and Second Language Learning*. Pergamon Press. http://www.sdkrashen.com/content/books/sl_acquisition_and_learning.pdf
- Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition*. Pergamon Press. http://www.sdkrashen.com/content/books/principles_and_practice.pdf
- Krashen, S. D. (1985). *The Input Hypothesis: Issues and Implications*. Longman.
- Krashen, S. D. (1989a). *Language Acquisition and Language Learning*. Prentice Hall.
- Krashen, S. D. (1989b). We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis. *The Modern Language Journal*, 73, 440–464.
- Krashen, S. D. (1994). The Input Hypothesis and Its Rivals. In N. C. Ellis (A c. Di), *Implicit and Explicit Learning of Languages* (pp. 45–77). Academic Press.
- Krashen, S. D. (2002). *Explorations in Language Acquisition and Use: The Taipei Lectures*. Heinemann.
- Krashen, S. D. (2003). *Explorations in Language Acquisition and Use*. Heinemann.
- Kroll, J. F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections between Bilingual Memory Representations. *Journal of Memory and Language*, 33(2), 149–174.
- Kroll, J. F., & Tokowicz, N. (2005a). Models of Bilingual Representation and Processing: Looking Back and Moving Forward. In J. F. Kroll & A. M. B. De Groot (A c. Di), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 531–553). Oxford University Press.

- Kroll, J. F., & Tokowicz, N. (2005b). Models of Bilingual Representation and Processing: Looking Back and to the Future. In *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 531–553). Oxford University Press.
- Ladefoged, P., & Johnson, K. (2014). *A Course in Phonetics*. Cengage Learning. <https://books.google.it/books?id=U4XaAgAAQBAJ>
- Lantolf, J. P. (2000). *Sociocultural Theory and Second Language Learning*. Oxford University Press.
- Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural Theory and the Genesis of Second Language Development*. Oxford University Press.
- Larsen-Freeman, D. (1997). Chaos/Complexity Science and Second Language Acquisition. *Applied Linguistics*, 18(2), 141–165.
- Larsson, T., & Biber, D. (2024). On the Perils of Linguistically Opaque Measures and Methods. In P. Crosthwaite (A c. Di), *Corpora for Language Learning: Bridging the Research-Practice Divide* (pp. 131–141). Routledge.
- Laufer, B. (1989). What Percentage of Text Lexis is Necessary for Comprehension? *Proceedings of the European ACL Conference*, 316–323.
- Laufer, B. (1992). Reading comprehension in second languages: The role of lexical knowledge. In P. J. L. Arnaud & H. Béjoint (A c. Di), *Vocabulary and Applied Linguistics* (pp. 126–132). Macmillan.
- Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, 54(3), 399–436.
- Laufer, B., & Hulstijn, J. H. (2001). Incidental Vocabulary Acquisition in a Second Language: The Construct of Task-Induced Involvement. *Applied Linguistics*, 22(1), 1–26.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307–322. <https://doi.org/10.1093/applin/16.3.307>
- Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation* (1^a ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511815355>

- Lee, H., Chang, A., Peirsman, Y., Chambers, N., Surdeanu, M., & Jurafsky, D. (2013). Deterministic coreference resolution based on entity-centric, precision-ranked rules. *Computational Linguistics*, 39(4), 885–916.
- Lee, H., Peirsman, Y., Chang, A., Chambers, N., Surdeanu, M., & Jurafsky, D. (2011). Stanford's Multi-Pass Sieve Coreference Resolution System at the CoNLL-2011 Shared Task. *Proceedings of the Fifteenth Conference on Computational Natural Language Learning: Shared Task*, 28–34.
- Lee, L. (2008). Focus-on-Form through Collaborative Scaffolding in Expert-to-Novice Online Interaction. *Language Learning & Technology*, 12(3), 53–72.
- Lee, S.-K., & Huang, H.-T. (2008). Visual Input Enhancement and Grammar Learning: A Meta-Analytic Review. *Studies in Second Language Acquisition*, 30(3), 307–331.
- Lee, S.-M. (2023). Factors affecting incidental L2 vocabulary acquisition and retention in a game-enhanced learning environment. *ReCALL*, 35(3), 274–289. <https://doi.org/10.1017/S0958344022000209>
- Lenci, A., & Littell, J. S. (2008). Distributional semantics in linguistic and cognitive research. *The Italian Journal of Linguistics*, 20, 1–32.
- Lenneberg, E. H. (1967). *Biological Foundations of Language*. Wiley.
- Leow, R. P. (2009). Input Enhancement and L2 Grammatical Development: What the Research Reveals. *Studies in Second Language Acquisition*, 31(2), 269–288.
- Leow, R. P. (2015). *Explicit Learning in the L2 Classroom: A Student-Centered Approach*. Routledge.
- Lesk, M. (1986). Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from an ice cream cone. *Proceedings of the 5th Annual International Conference on Systems Documentation*, 24–26. <https://doi.org/10.1145/318723.318728>
- Lesson Nine GmbH. (2007). *Babbel Language Learning Platform*.
- Levy, M. (1997a). *CALL: Context and Conceptualization*. Oxford University Press.
- Levy, M. (1997b). *Computer-Assisted Language Learning: Context and Conceptualization*. Clarendon Press.

- Levy, M., & Hubbard, P. (2005). Why call call “CALL”? *Computer Assisted Language Learning*, 18(3), 143–149.
- Levy, M., & Stockwell, G. (2006). *CALL Dimensions: Options and Issues in Computer-Assisted Language Learning*. Routledge.
- Lewis, M. (1993a). *The Lexical Approach*. Language Teaching Publications.
- Lewis, M. (1993b). *The Lexical Approach: The State of ELT and a Way Forward*. Language Teaching Publications.
- Lewis, M. (1997). *Implementing the Lexical Approach: Putting Theory into Practice*. Language Teaching Publications.
- Li, H., & Xu, F. (2016). Question Answering with DBpedia Based on the Dependency Parser and Entity-centric Index. *2016 International Conference on Computational Intelligence and Applications (ICCI)*, 41–45. <https://doi.org/10.1109/ICCI.2016.10>
- Li, J., Sun, A., Han, J., & Li, C. (2022). A Survey on Deep Learning for Named Entity Recognition. *IEEE Transactions on Knowledge and Data Engineering*, 34(1), 50–70. <https://doi.org/10.1109/TKDE.2020.2981314>
- Li, X., Tur, G., Hakkani-Tür, D., & Li, Q. (2014). Personal knowledge graph population from user utterances in conversational understanding. *2014 IEEE Spoken Language Technology Workshop (SLT)*, 224–229. <https://doi.org/10.1109/SLT.2014.7078578>
- Li, Z., Link, S., & Hegelheimer, V. (2015). Rethinking the Role of Automated Corrective Feedback in Writing Instruction: New Perspectives. *CALICO Journal*, 32(3), 512–537.
- Lightbown, P. M., & Spada, N. (2006). *How Languages are Learned* (3rd ed.). Oxford University Press.
- Lightbown, P. M., & Spada, N. (2013). *How Languages Are Learned*. Oxford University Press.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22 140, 55–55.

- Lin, C.-C., & Lin, H.-Y. (2019). Mobile-Assisted ESL/EFL Vocabulary Learning: A Systematic Review and Meta-Analysis. *Computer Assisted Language Learning*, 32(8), 878–919.
- Littauer, R., Villazon-Terrazas, B., & Moran, S. (2013). Linguistic Resources Enhanced with Geospatial Information. *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language data*, 53–58. <https://aclanthology.org/W13-5508>
- Loewen, S. (2005). Incidental Focus on Form and Second Language Learning. *Studies in Second Language Acquisition*, 27(3), 361–386.
- Loewen, S., Isbell, D. R., & Sporn, Z. (2019). The effectiveness of app-based language instruction for developing receptive linguistic knowledge and oral communicative ability. *Foreign Language Annals*, 52(4), 779–795.
- Long, M. H. (1983a). Linguistic and Conversational Adjustments to Non-Native Speakers. In N. Wolfson & E. Judd (A c. Di), *Conversational Routine: Explorations in Standardized Communication Situations and Prepatterned Speech* (pp. 177–193). Newbury House.
- Long, M. H. (1983b). Native Speaker/Non-Native Speaker Conversation and the Negotiation of Comprehensible Input. *Applied Linguistics*, 4(2), 126–141.
- Long, M. H. (1985). Input and Second Language Acquisition Theory. In S. M. Gass & C. G. Madden (A c. Di), *Input in Second Language Acquisition* (pp. 377–393). Newbury House.
- Long, M. H. (1991). Focus on Form: A Design Feature in Language Teaching Methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (A c. Di), *Foreign Language Research in Cross-Cultural Perspective* (pp. 39–52). John Benjamins.
- Long, M. H. (1996). The Role of the Linguistic Environment in Second Language Acquisition. In W. C. Ritchie & T. K. Bhatia (A c. Di), *Handbook of Second Language Acquisition* (pp. 413–468). Academic Press.
- Long, M. H. (1998). Focus on Form in Task-Based Language Teaching. In C. Doughty & J. Williams (A c. Di), *Focus on Form in Classroom Second Language Acquisition* (pp. 15–41). Cambridge University Press.

- Long, M. H. (2015). *Second Language Acquisition and Task-Based Language Teaching*. Wiley-Blackwell.
- Lyster, R., & Ranta, L. (1997a). Corrective Feedback and Learner Uptake. *Studies in Second Language Acquisition*, 19(1), 37–66.
- Lyster, R., & Ranta, L. (1997b). Corrective Feedback and Learner Uptake: Negotiation of Form in Communicative Classrooms. *Studies in Second Language Acquisition*, 19(1), 37–66. <https://doi.org/10.1017/S0272263197001034>
- Macalister, J., & Nation, I. S. P. (2019). *Language Curriculum Design* (2^a ed.). Routledge. <https://doi.org/10.4324/9780429203763>
- MacIntyre, P. D. (1995). How does anxiety affect second language learning? A reply to Sparks and Ganschow. *The Modern Language Journal*, 79(1), 90–99.
- Mackey, A. (1999a). Input, Interaction, and Second Language Development. *Studies in Second Language Acquisition*, 21(4), 557–587.
- Mackey, A. (1999b). Input, Interaction, and Second Language Development: An Empirical Study of Question Formation in ESL. *Studies in Second Language Acquisition*, 21(4), 557–587. <https://doi.org/10.1017/S0272263199004027>
- Mackey, A. (2006). Feedback, Noticing and Instructed Second Language Learning. *Applied Linguistics*, 27(3), 405–430. <https://doi.org/10.1093/applin/ami051>
- Mackey, A., & Gass, S. M. (2015). *Second Language Research: Methodology and Design* (2^a ed.). Routledge.
- Mackey, A., & Goo, J. (2007a). Interaction Research in SLA: A Meta-Analysis. In A. Mackey (A c. Di), *Conversational Interaction in Second Language Acquisition* (pp. 407–452). Oxford University Press.
- Mackey, A., & Goo, J. (2007b). Interaction Research in SLA: A Meta-Analysis and Research Synthesis. In A. Mackey (A c. Di), *Conversational Interaction in Second Language Acquisition: A Series of Empirical Studies* (pp. 407–452). Oxford University Press.

- Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual Differences in Working Memory, Noticing of Interactional Feedback and L2 Development. In P. Robinson (A c. Di), *Individual Differences and Instructed Language Learning* (pp. 181–209). John Benjamins.
- Maistro, M., & Chiarcos, C. (2012). Linked data for language-learning applications. *Proceedings of the 2012 Joint Workshop on Multilingual OCR and Analytics for Noisy Unstructured Text Data*, 59–66.
- Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & López-Cózar, E. D. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of informetrics*, 12(4), 1160–1177. <https://doi.org/10.1016/j.joi.2018.09.002>
- McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why There Are Complementary Learning Systems in the Hippocampus and Neocortex: Insights from the Successes and Failures of Connectionist Models of Learning and Memory. *Psychological Review*, 102(3), 419–457.
- McCrae, J., & Cimiano, P. (2011). Multilingual linked data. *Proceedings of the 10th International Semantic Web Conference (ISWC 2011), Bonn, Germany, October*.
- McCrae, J. P., Buitelaar, P., Gracia, J., Kernerman, I., McCrae, M., Montiel-Ponsoda, E., Ordan, N., & Villegas, M. (2019). The ELEXIS Interface for Interoperable Lexical Resources. *Proceedings of the 6th Biennial Conference on Electronic Lexicography (eLex 2019)*, 339–352. <https://aclanthology.org/2019.elex-1.28>
- McCrae, J. P., Buitelaar, P., Montiel-Ponsoda, E., Villegas, M., & Gracia, J. (2021). Linked Open Data for Lexicography: The Case of ELEXIS. *Proceedings of the 19th International Conference on Artificial Intelligence: Methodology, Systems, and Applications (AIMSA 2021)*, 3–12. https://doi.org/10.1007/978-3-030-85713-4_1
- McCrae, J. P., Chiarcos, C., Bond, F., Cimiano, P., Declerck, T., de Melo, G., Gracia, J., Hellmann, S., Klimek, B., Moran, S., Osenova, P., Pareja-Lora, A., & Pool, J. (2016). The Open Linguistics Working Group: Developing the Linguistic Linked Open Data Cloud. *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC)*, 2435–2441. <https://aclanthology.org/L16-1386>
- McCrae, J. P., & Fellbaum, C. (2016). The Semantic Web and WordNet. *Proceedings of the 8th Global WordNet Conference*, 387–394. <https://aclanthology.org/W16-2029>

- McCrae, J. P., Gil, J. B., Gràcia, J., Bitelaar, P., & Cimiano, P. (2017). *The OntoLex-Lemon Model: Development and Applications*. <https://api.semanticscholar.org/CorpusID:36965997>
- McCrae, J. P., Rademaker, A., Bond, F., Rudnicka, E., & Fellbaum, C. (2019). English WordNet 2019 – An Open-Source WordNet for English. *Proceedings of the 10th Global Wordnet Conference*, 245–252. <https://aclanthology.org/2019.gwc-1.31>
- McCrae, J. P., Rademaker, A., Rudnicka, E., & Bond, F. (2020). English WordNet 2020: Improving and Extending a WordNet for English using an Open-Source Methodology. *Proceedings of the LREC Workshop on Multimodal Wordnets (MMW)*, 14–19. <https://aclanthology.org/2020.mmw-1.3>
- McLaughlin, B. (1987). *Theories of Second-Language Learning*. Edward Arnold.
- McQuillan, J., & Krashen, S. D. (2007). Extensive Reading and Vocabulary Acquisition in Second Language Acquisition: A Critique. *Reading in a Foreign Language*, 19(1), 1–13.
- Meara, P. (1982). Word Associations in a Foreign Language: A Report on the Birkbeck Vocabulary Project. *Nottingham Linguistic Circular*, 11, 29–37.
- Meara, P. (1996). The Dimensions of Lexical Competence. *Performance and Competence in Second Language Acquisition*, 35–53.
- Meara, P. (2002). Connected Words: Word Associations and Second Language Vocabulary Acquisition. In N. Schmitt & M. McCarthy (A c. Di), *The Vocabulary Dimension* (pp. 57–67). Cambridge University Press.
- Melchers, G., & Shaw, P. (2011). *World Englishes*. Hodder Education. <https://books.google.it/books?id=r4GInQEACAAJ>
- Menke, P., McCrae, J. P., & Cimiano, P. (2013). Releasing multimodal data as Linguistic Linked Open Data: An experience report. *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language data*, 44–52. <https://aclanthology.org/W13-5507>
- Merriam-Webster.com dictionary*. (2023). <https://www.merriam-webster.com/>

- Messick, S. (1984). The Nature of Cognitive Styles: Problems and Promise in Educational Practice. In *Educational Testing Service Research Report Series* (pp. 59–74). Educational Testing Service.
- Meurers, D. (2012). Natural Language Processing and Language Learning. In C. A. Chapelle (A c. Di), *The Encyclopedia of Applied Linguistics* (pp. 4193–4205). Wiley-Blackwell.
- Meurers, D. (2015). Integrating Advanced NLP Technology for Supporting Language Learning. *Proceedings of the 2015 Conference on Computer Assisted Language Learning (CALL)*, 14–23.
- Meurers, D. R. (2010). ICALL and NLP: Exploring the Role of Natural Language Processing in CALL. *Speech and Language Technology in Education (SLaTE 2010)*, 25–32. <https://aclanthology.org/2010.slate-1.5>
- Meurers, D., Ziai, R., Amaral, L., & Boyd, A. (2010). Enhancing Authentic Web Pages for Language Learners. *Proceedings of the NAACL HLT 2010 Workshop on NLP and Linguistics: Finding the Common Ground*, 1–9.
- Meurers, D., Ziai, R., Amaral, L., & Boyd, A. (2011). WERTi: Web-Based Exercises for Language Learning. *ReCALL*, 23(2), 135–153.
- Meyer, C. M., & Gurevych, I. (2012). Wiktionary: A new rival for expert-built lexicons? Exploring the possibilities of collaborative lexicography. In *Electronic Lexicography*. Oxford University Press.
- Miao, Q., Lu, H., Zhang, S., & Meng, Y. (2013). Cross-lingual link discovery between Chinese and english wiki knowledge bases. *27th Pacific Asia Conference on Language, Information, and Computation, PACLIC 27*, 374–381.
- Microsoft Edge Dev. (2025). *Progressive Web Apps on Microsoft Edge*. <https://learn.microsoft.com/microsoft-edge/progressive-web-apps-chromium/>
- Microsoft Playwright Team. (2025). *Playwright Documentation*. <https://playwright.dev/docs/intro>
- Miller, E. (2011). Libraries, languages of description, and linked data: A Dublin Core perspective. *Information Technology and Libraries*, 30(3), 148–155.
- Miller, G. A. (1990). Nouns in WordNet: A Lexical Inheritance System. *International Journal of Lexicography*, 3(4), 245–264. <https://doi.org/10.1093/ijl/3.4.245>
- Pasquale Esposito

- Miller, G. A. (1995). WordNet: A Lexical Database for English. *Communications of the ACM*, 38(11), 39–41. <https://doi.org/10.1145/219717.219748>
- Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. J. (1990). Introduction to WordNet: An On-line Lexical Database. *International Journal of Lexicography*, 3(4), 235–244.
- Miller, G. A., & Fellbaum, C. (2007). WordNet Then and Now. *Language Resources and Evaluation*, 41(2), 209–214.
- Milroy, L., & Gordon, M. K. (2003). *Sociolinguistics: Method and interpretation*. Blackwell Pub. <https://doi.org/10.1002/9780470758359>
- Milton, J. (2009). Measuring Second Language Vocabulary Acquisition. *International Journal of Applied Linguistics*, 19(1), 97–112. <https://doi.org/10.1111/j.1473-4192.2009.00221.x>
- Mitkov, R., & Ha, L. A. (2014). Automatic Generation of Multiple-Choice Questions for Language Learning and Assessment. *Proceedings of the Ninth Workshop on Innovative Use of NLP for Building Educational Applications*, 56–65. <https://doi.org/10.3115/v1/W14-1807>
- Molnár, A., Nagy, Z., & Yangarber, R. (2022). Large Language Models in Second Language Learning: Opportunities and Challenges. *ReCALL*, 34(3), 239–257. <https://doi.org/10.1017/S095834402200009X>
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106, 213–228. <https://doi.org/10.1007/s11192-015-1765-5>
- Montemagni, S., Barsotti, F., Battista, M., Calzolari, N., Lenci, A., Monachini, M., Pirrelli, V., & Soria, C. (2000). The Italian Treebank: Building a Treebank for Italian. *Proceedings of the 2nd International Conference on Language Resources and Evaluation (LREC)*, 1899–1904. <https://aclanthology.org/L00-1185>
- Moran, S., & Brümmer, M. (2013). Lemon-aid: Using Lemon to aid quantitative historical linguistic analysis. *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language data*, 28–33. <https://aclanthology.org/W13-5505>

- Moran, S., & McCloy, D. (A c. Di). (2019). *PHOIBLE 2.0*. Max Planck Institute for the Science of Human History. <https://phoible.org>
- Morgan, J. L. (2013). Pragmatics and Language Teaching. *Annual Review of Applied Linguistics*, 33, 113–129. <https://doi.org/10.1017/S0267190513000043>
- Morris, R. D., Lovett, M. W., & Wolf, M. (2008). The role of morphological awareness in children's reading development. *Reading Research Quarterly*, 43(4), 362–381. <https://doi.org/10.1598/RRQ.43.4.2>
- Mozilla MDN Contributors. (2025). *Progressive Web Apps (PWA)—MDN Web Docs*. https://developer.mozilla.org/en-US/docs/Web/Progressive_web_apps
- MUI. (2025). *MUI: React UI Components*. <https://mui.com/>
- Mukherjee, J. (2009). *Introducing Second Language Acquisition*. Cambridge University Press.
- Nacke, P. L. (2011). *Skimming Strategy in Reading as a Function of Familiarity with Content and Redundancy Reduction in Printed Discourse* [PhD Thesis, University of British Columbia]. <https://doi.org/10.14288/1.0104082>
- Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning Words from Context. *Reading Research Quarterly*, 20, 233–253.
- Nahavandi, N., & Mukundan, J. (2014). The Impact of Textual Enhancement on Vocabulary Intake of Iranian Elementary EFL Learners. *English Language Teaching*, 7(4), 28–35.
- Nakata, T. (2011). Computer-Assisted Second Language Vocabulary Learning in a Paired-Associate Paradigm: A Critical Investigation of Flashcard Software. *Computer Assisted Language Learning*, 24(1), 17–38.
- Narducci, F., Palmonari, M., & Semeraro, G. (2013). Cross-language semantic matching for discovering links to e-gov services in the LOD cloud. *CEUR Workshop Proceedings*, 992, 21–32.
- Nation, I. S. P. (1987). *Teaching & Learning Vocabulary*. Heinle & Heinle.
- Nation, I. S. P. (1990). *Teaching Vocabulary: Strategies and Techniques*. Heinle & Heinle.

- Nation, I. S. P. (1994). *New Ways in Teaching Vocabulary*. TESOL.
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge University Press.
- Nation, I. S. P. (2006). How Large a Vocabulary is Needed for Reading and Listening? *The Canadian Modern Language Review / La revue canadienne des langues vivantes*, 63(1), 59–82.
- Nation, I. S. P. (2007a). The Four Strands. *Innovation in Language Learning and Teaching*, 1(1), 2–13.
- Nation, I. S. P. (2007b). *What Should Every EFL Teacher Know?* Compass Publishing.
- Nation, I. S. P. (2008). *Teaching ESL/EFL Reading and Writing*. Routledge.
- Nation, I. S. P. (2009). *Language Curriculum Design*. Routledge.
- Nation, I. S. P. (2013a). *Learning Vocabulary in Another Language (2nd ed.)*. Cambridge University Press.
- Nation, I. S. P. (2013b). *What Should Every EFL Teacher Know?* Compass Publishing.
- Nation, I. S. P. (2013c). *What Should Every EFL Teacher Know? (Second Revised Edition)*. Compass Publishing.
- Nation, I. S. P. (2017). *The BNC/COCA Word Family Lists (Version 1.0.0)*. <http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>
- Nation, I. S. P. (2019). *Teaching Extensive Reading in Another Language*. Routledge.
- Nation, I. S. P., & Crabbe, D. (1991). A Survival Language Learning Syllabus for Foreign Travel. *System*, 19(3), 191–201. [https://doi.org/10.1016/0346-251X\(91\)90044-P](https://doi.org/10.1016/0346-251X(91)90044-P)
- Nation, I. S. P., & Macalister, J. (2020). *Language Curriculum Design (Second Edition)*. Routledge.
- Nation, P. (2001). Learning Vocabulary in Another Language. *Cambridge Applied Linguistics*. <https://doi.org/10.1017/CBO9781139524759>
- Nation, P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63(1), 59–82.

- Nation, P. (2017). *The BNC/COCA word family lists*. <http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical Phrases and Language Teaching*. Oxford University Press.
- Navigli, R., Bevilacqua, M., Conia, S., Montagnini, D., & Cecconi, F. (2021). Ten Years of BabelNet: A Survey. *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, 4559–4567. <https://doi.org/10.24963/ijcai.2021/620>
- Navigli, R., & Ponzetto, S. P. (2012). BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence*, 193, 217–250. <https://doi.org/10.1016/j.artint.2012.07.001>
- Odlin, T. (1989). *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge University Press.
- Ohta, A. S. (2001). *Second Language Acquisition Processes in the Classroom: Learning Japanese*. Lawrence Erlbaum Associates.
- OpenTelemetry Authors. (2025). *OpenTelemetry Documentation*. <https://opentelemetry.io/docs/>
- Osterhout, L., & Holcomb, P. J. (1992). Event-Related Brain Potentials Elicited by Syntactic Anomaly. *Journal of Memory and Language*, 31(6), 785–806.
- Oxford English Dictionary*. (2023). <https://www.oed.com/>
- Paivio, A. (1990). *Mental Representations: A dual coding approach*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195066661.001.0001>
- Pareja-Lora, A., Blume, M., & Lust, B. (2013). Transforming the Data Transcription and Analysis Tool Metadata and Labels into a Linguistic Linked Open Data Cloud Resource. *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language data*, 34–43. <https://aclanthology.org/W13-5506>
- Pasquale Esposito. (2025). *pasesposito/Appil-an-Interface-for-Lexical-and-Linguistic-Competence-Improvement: Appil: an Interface for Lexical and Linguistic Competence Improvement* (Versione v.1.0.0) [Software]. Zenodo. <https://doi.org/10.5281/ZENODO.17490980>
- Pasquale Esposito

- Patterson, M. B., & Mellard, D. (2007). Program Characteristics that Predict Improved Learner Outcomes. *Adult Basic Education and Literacy Journal*, 1, 84–95.
- Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and Cognition*, 11(2), 147–164. <https://doi.org/10.1017/S1366728908003283>
- Philp, J. (2003a). Constraints on “Noticing the Gap”: Nonnative Speakers’ Noticing of Recasts. *Studies in Second Language Acquisition*, 25(1), 99–126.
- Philp, J. (2003b). Constraints on “Noticing the Gap”: Nonnative Speakers’ Noticing of Recasts in NS–NNS Interaction. *Studies in Second Language Acquisition*, 25(1), 99–126. <https://doi.org/10.1017/S0272263103000044>
- Pica, T. (1994a). Research on Negotiation: What Does It Reveal about Second-Language Learning? *Language Learning*, 44(3), 493–527.
- Pica, T. (1994b). Research on Negotiation: What Does It Reveal about Second-Language Learning Conditions, Processes, and Outcomes? *Language Learning*, 44(3), 493–527. <https://doi.org/10.1111/j.1467-1770.1994.tb01115.x>
- Pinker, S. (1994). *The language instinct: How the mind creates language* (2. [print.]). Morrow.
- Popov, A., Simov, K., & Osenova, P. (2019). Know Your Graph. State-of-the-Art Knowledge-Based WSD. *Proceedings - Natural Language Processing in a Deep Learning World*, 949–958. https://doi.org/10.26615/978-954-452-056-4_110
- Porter, M. F. (2006). An Algorithm for Suffix Stripping. *Program*, 40(3), 211–218. <https://doi.org/10.1108/00330330610681286>
- Python Software Foundation. (2022a). *Python (Version 3.11)*. <https://www.python.org/>
- Python Software Foundation. (2022b). *Python (Version 3.11) [Software]*. <https://www.python.org/>
- Quecedo, J., Katinskaia, A., & Yangarber, R. (2020). Word-Sense Disambiguation for Intelligent Computer-Assisted Language Learning in Morphologically Rich Languages. *Proceedings of the 12th Language Resources and Evaluation Conference (LREC 2020)*, 7120–7129.

- RabbitMQ Team. (2025). *RabbitMQ Documentation and Tutorials*. <https://www.rabbitmq.com/getstarted.html>
- Raghunathan, K., Lee, H., Rangarajan, S., Chambers, N., Surdeanu, M., Jurafsky, D., & Manning, C. (2010). A Multi-Pass Sieve for Coreference Resolution. *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 492–501.
- Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A., ... Wellman, M. (2019). Machine Behaviour. *Nature*, 568(7753), 477–486.
- Ratcliff, J. W., & Metzener, D. E. (1988). Pattern Matching: The Gestalt Approach. *Dr. Dobb's Journal*, 46.
- React Team. (2025). *React: Learn*. <https://react.dev/learn>
- Recasens, M., Marneffe, M.-C. de, & Potts, C. (2013). The Life and Death of Discourse Entities: Identifying Singleton Mentions. *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*.
- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation, Legislazione No. 2016/679, 32016R0679 (2016). <http://data.europa.eu/eli/reg/2016/679/oj>
- Reinders, H., & Ellis, R. (2009). The Effects of Two Types of Input on Intake and the Acquisition of Implicit and Explicit Knowledge. *Studies in Second Language Acquisition*, 31(2), 277–305.
- Reinders, H., & White, C. (2016). 20 Years of Autonomy and Technology: How Far Have We Come and Where to Next? *Language Learning & Technology*, 20(2), 143–154.
- Reiners, T., & Wood, L. C. (2015). *Gamification in Education and Business*. Springer.
- Reinhardt, J., & Sykes, J. M. (2014). Digital Game and Play Activity in L2 Teaching and Learning. *Language Learning & Technology*, 18(2), 2–8.

- Richards, J. C. (2001). *Curriculum Development in Language Teaching* (1^a ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667220>
- Richards, J. C., & Rodgers, T. S. (2001). *Approaches and Methods in Language Teaching* (2^a ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667305>
- Riding, R., & Rayner, S. (2013). *Cognitive Styles and Learning Strategies: Understanding Style Differences in Learning and Behavior*. Routledge.
- Rodd, J. (2025). *Settling into Semantic Space: An Ambiguity-Focused Account of Word-Meaning Access*. Oxford University Press.
- Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15(1), 20–27. <https://doi.org/10.1016/j.tics.2010.09.003>
- Rosetta Stone Ltd. (1992). *Rosetta Stone Language Learning Software*.
- Rott, S. (1999). The Effect of Exposure Frequency on Intermediate Language Learners' Incidental Vocabulary Acquisition and Retention through Reading. *Studies in Second Language Acquisition*, 21(4), 589–619. <https://doi.org/10.1017/S0272263199004039>
- Rott, S. (2007a). The Effect of Frequency of Input Enhancement and Type of Enhancement on L2 Text Comprehension and Vocabulary Acquisition. *Language Learning*, 57(2), 165–199.
- Rott, S. (2007b). The Effect of Frequency of Input-Enhancements on Word Learning and Text Comprehension. *Language Learning*, 57(2), 165–199.
- Rott, S., & Williams, J. (2003). Input Enhancement and L2 Text Comprehension. *Studies in Second Language Acquisition*, 25(4), 539–566.
- Roy, S. D., & Zeng, W. (2014). Cognitive canonicalization of natural language queries using semantic strata. *ACM Trans. Speech Lang. Process.*, 10(4). <https://doi.org/10.1145/2539053>
- Rugg, M. D. (1985). The Effects of Semantic Priming and Word Repetition on Event-Related Potentials. *Psychophysiology*, 22(6), 642–647.

- Sauer, L. (2017). Depth of Processing in Vocabulary Learning: Revisiting Input Enhancement. In S. Loewen & M. Sato (A c. Di), *The Routledge Handbook of Instructed Second Language Acquisition* (pp. 162–176). Routledge.
- Sauro, S. (2009). Computer-Mediated Corrective Feedback and the Development of L2 Grammar. *Language Learning & Technology*, 13(1), 96–120.
- Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., & Monfardini, G. (2009). The Graph Neural Network Model. *IEEE Transactions on Neural Networks*, 20(1), 61–80. <https://doi.org/10.1109/TNN.2008.2005605>
- Schmidt, R. (2001). Attention. In P. Robinson (A c. Di), *Cognitive Psychology and Second Language Learning: The Role of Attention and Awareness in Experience* (pp. 3–32). Cambridge University Press.
- Schmidt, R. W. (1990). The Role of Consciousness in Second Language Learning. *Applied Linguistics*, 11(2), 129–158. <https://doi.org/10.1093/applin/11.2.129>
- Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge University Press.
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329–363. <https://doi.org/10.1177/1362168808089921>
- Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Palgrave Macmillan.
- Schramm, S., Wehner, C., & Schmid, U. (2023). Comprehensible Artificial Intelligence on Knowledge Graphs: A survey. *Journal of Web Semantics*, 79, 100806. <https://doi.org/10.1016/j.websem.2023.100806>
- Schreuder, R., & Baayen, R. H. (1997). How Complex Simplex Words Can Be. *Journal of Memory and Language*, 37(1), 118–139. <https://doi.org/10.1006/jmla.1997.2510>
- Schulze, M. (2008a). AI in CALL: Artificially Intelligent CALL—An Overview. *CALICO Journal*, 26(3), 464–475.
- Schulze, M. (2008b). Interfaces in Intelligent CALL. *Computer Assisted Language Learning*, 21(4), 301–303. <https://doi.org/10.1080/09588220802393608>

- Schwartz, B. D., & Sprouse, R. A. (1996). L2 Cognitive States and the Full Transfer/Full Access Model. *Second Language Research*, 12(1), 40–72. <https://doi.org/10.1177/026765839601200103>
- Seargeant, P., & Tagg, C. (2011). English on the internet and a ‘post-varieties’ approach to language. *World Englishes*, 30(4), 496–514. <https://doi.org/10.1111/j.1467-971X.2011.01730.x>
- Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press.
- Seedhouse, P. (2004). *The Interactional Architecture of the Language Classroom: A Conversation Analysis Perspective*. Blackwell.
- Selenium Project. (2025). *Selenium WebDriver Documentation*. <https://www.selenium.dev/documentation/>
- Sérasset, G. (2015). DBnary: Wiktionary as a Lemon-based multilingual lexical resource in RDF. *Semantic Web*, 6(4), 355–361. <https://doi.org/10.3233/SW-140147>
- Sérasset, G., & Tchechmedjiev, A. (2014). DBnary: Wiktionary as Linked Data for 12 Language Editions with Enhanced Translation Relations. *3rd Workshop on Linked Data in Linguistics: Multilingual Knowledge Resources and Natural Language Processing*, 67–71. <https://hal.science/hal-00990876>
- Sharwood Smith, M. (1993). Input Enhancement in Instructed SLA. In S. M. Gass & L. Selinker (A c. Di), *Second Language Acquisition in Context* (pp. 165–179). Lawrence Erlbaum.
- Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*, 57(10), 1380–1400. <https://doi.org/10.1177/0002764213498851>
- Simov, K., Popov, A., & Osenova, P. (2015). Improving word sense disambiguation with linguistic knowledge from a sense annotated treebank. *Proceedings of the International Conference Recent Advances in Natural Language Processing*, 596–603. <https://aclanthology.org/R15-1077>
- Skandan, S., Kanungo, S., Devaraj, S., Gupta, S., & Narayan, S. (2023). Question Answering System using Knowledge Graphs. *International Conference on Inventive Computation Technologies (ICICT)*, 656–661. <https://doi.org/10.1109/ICICT57646.2023.10134047>
- Skinner, B. F. (1957). *Verbal Behavior*. Appleton-Century-Crofts.
- Pasquale Esposito

- Slobin, D. I. (1985). Crosslinguistic Evidence for the Language-Making Capacity. In *The Crosslinguistic Study of Language Acquisition, Vol. 1: The Data; Vol. 2: Theoretical Issues* (pp. 1157–1256). Lawrence Erlbaum Associates.
- Smith, B. (2003). Computer-Mediated Negotiated Interaction: An Expanded Model. *The Modern Language Journal*, 87(1), 38–57.
- Smith, E. A., & Senter, R. J. (1967). *Automated Readability Index* (AMRL-TR; pp. 1–14). Aerospace Medical Research Laboratories (U.S.).
- Smith, M. S. (1991). Speaking to Many Minds: On the Relevance of Different Types of Language Information for the L2 Learner. *Second Language Research*, 7(2), 118–132.
- Sparck Jones, K., & Willett, P. (A c. Di). (1997). *Readings in Information Retrieval*. Morgan Kaufmann.
- Speer, R., Chin, J., & Havasi, C. (2017). ConceptNet 5.5: An Open Multilingual Graph of General Knowledge. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 4444–4451. <https://doi.org/10.5555/3298023.3298212>
- Speer, R., Chin, J., & Havasi, C. (2018). *ConceptNet 5.5: An Open Multilingual Graph of General Knowledge* (arXiv:1612.03975). arXiv. <https://doi.org/10.48550/arXiv.1612.03975>
- Steyvers, M., & Tenenbaum, J. B. (2005). The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth. *Cognitive Science*, 29(1), 41–78.
- Stockwell, G. (2010a). Using Mobile Phones for Vocabulary Activities: Examining the Effect of Platform. *Language Learning & Technology*, 14(2), 95–110.
- Stockwell, G. (2010b). Using Mobile Phones for Vocabulary Activities: Examining the Effect of the Platform. *Language Learning & Technology*, 14(2), 95–110.
- Swain, M. (1985a). Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in Its Development. In S. Gass & C. Madden (A c. Di), *Input in Second Language Acquisition* (pp. 235–253). Newbury House.

- Swain, M. (1985b). Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in Its Development. In S. Gass & C. Madden (A c. Di), *Input in Second Language Acquisition* (pp. 235–253). Newbury House.
- Swain, M. (1985c). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. *Input in Second Language Acquisition*, 235–253.
- Swain, M. (1995). Three Functions of Output in Second Language Learning. In G. Cook & B. Seidlhofer (A c. Di), *Principles and Practice in Applied Linguistics* (pp. 125–144). Oxford University Press.
- Swain, M. (2005). The Output Hypothesis: Theory and Research. In E. Hinkel (A c. Di), *Handbook of Research in Second Language Teaching and Learning* (pp. 471–483). Lawrence Erlbaum Associates.
- Swain, M., & Lapkin, S. (1995). Problems in Output and the Cognitive Processes They Generate: A Step Towards Second Language Learning. *Applied Linguistics*, 16, 371–391.
- Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Sweller, J. (1994). Cognitive Load Theory, Learning Difficulty, and Instructional Design. *Learning and Instruction*, 4(4), 295–312.
- Szudarski, P., & Carter, R. (2016). The Role of Input Flood and Textual Enhancement in EFL Learners' Acquisition of Collocations. *International Journal of Applied Linguistics*, 166(2), 251–276.
- Taguchi, N. (2011). Teaching Pragmatics: Trends and Issues. *Annual Review of Applied Linguistics*, 31, 289–310.
- Taguchi, N. (2019). *Pragmatic Development in a Second Language*. Routledge.
- Tailwind Labs. (2025). *Tailwind CSS Documentation*. <https://tailwindcss.com/docs>
- Talamas, A., Kroll, J. F., & Dufour, R. (1999). From Form to Meaning: Stages in the Acquisition of Second-Language Vocabulary. *Bilingualism: Language and Cognition*, 2(1), 45–58.

- Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, G. M. (2010). Sleep Spindle Activity Is Associated with the Integration of New Memories and Existing Knowledge. *Journal of Neuroscience*, 30(43), 14356–14360.
- Tarski, A. (1995). *Introduction to Logic and to the Methodology of Deductive Sciences*. Dover Publications. <https://books.google.it/books?id=5MeNCgAAQBAJ>
- Te Molder, H., & Potter, J. (2005). *Conversation and Cognition*. Cambridge University Press.
- Thomas, J. (1983). Cross-Cultural Pragmatic Failure. *Applied Linguistics*, 4(2), 91–112. <https://doi.org/10.1093/applin/4.2.91>
- Thomas, M. (2008). CALL dimensions – By Mike Levy & Glenn Stockwell. *British Journal of Educational Technology*, 39(3), 563–565. https://doi.org/10.1111/j.1467-8535.2008.00855_6.x
- Tiedemann, J., & Thottingal, S. (2020). OPUS-MT: Building open translation services for the World. *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*, 479–480. <https://doi.org/10.5281/zenodo.3766943>
- Tokowicz, N., & Kroll, J. F. (2007). Number of Meanings and Concreteness: Consequences of Ambiguity within and across Languages. *Language and Cognitive Processes*, 22(5), 727–779.
- Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.
- Trahey, M., & White, L. (1993). Positive Evidence and Preemption in the Second Language Classroom. *Studies in Second Language Acquisition*, 15(2), 181–204.
- Ullman, M. T. (2016). The Declarative/Procedural Model: A Neurobiological Model of Language Learning, Knowledge, and Use. *Neurobiology of Language*, 953–968.
- U.S. Department of Education. (2025). *Family Educational Rights and Privacy Act (FERPA)*. <https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html>
- Vaculíková, J., Kalenda, J., & Kočvarová, I. (2024). Participation in non-formal adult education within the European context: Examining multilayer approach. *Frontiers in Education*, 9, 1380865. <https://doi.org/10.3389/feduc.2024.1380865>

- van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, 39(3), 458–483. <https://doi.org/10.1006/jmla.1998.2584>
- van Rijsbergen, C. J., Robertson, S. E., & Porter, M. F. (1980). *New Models in Probabilistic Information Retrieval*. British Library.
- Vankrunkelsven, H., Verheyen, S., Storms, G., & De Deyne, S. (2018). Predictive Power of Word Associations for Lexical Processing. *Behavior Research Methods*, 50(4), 1553–1567.
- VanPatten, B. (1996a). *Input Processing and Grammar Instruction in Second Language Acquisition*. Ablex.
- VanPatten, B. (1996b). *Input Processing and Grammar Instruction: Theory and Research*. Ablex.
- VanPatten, B. (2004). *Processing Instruction: Theory, Research, and Commentary*. Lawrence Erlbaum.
- VanPatten, B. (2015). *While We're on the Topic: BVP on Language, Acquisition, and Classroom Practice*. ACTFL.
- VanPatten, B., & Williams, J. (A c. Di). (2014). *Theories in Second Language Acquisition* (0 ed.). Routledge. <https://doi.org/10.4324/9780203628942>
- VMware Spring. (2025). *Spring Boot Actuator: Production-Ready Features*. <https://docs.spring.io/spring-boot/docs/current/reference/html/actuator.html>
- Voghera, M. (2022). *From Speaking to Grammar*. Peter Lang Verlag. <https://doi.org/10.3726/b19221>
- von Ahn, L. (2013). Duolingo: Learn a Language for Free While Helping to Translate the Web. *Proceedings of the 2013 International Conference on Intelligent User Interfaces*, 1–2.
- Vue.js Core Team. (2025). *Vue.js Guide*. <https://vuejs.org/guide/introduction.html>
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, A c. Di). Harvard University Press.
- Vygotsky, L. S. (1986). *Thought and Language* (A. Kozulin, A c. Di). MIT Press.

- Wallace, R. S. (2009). *ALICE: Artificial Linguistic Internet Computer Entity*.
- Wang, M., Mak, M. Y., Chen, B., & Huang, H. (2022). Core Vocabulary in the Mental Lexicon: Evidence from a Word-Guessing Game. *Frontiers in Psychology, 13*, 861024.
- Wang, W., & Wang, W. (2009). Teaching Vocabulary by Combining Output Tasks, Glosses, and Textual Enhancement. *Foreign Language Annals, 42*(3), 703–720.
- Warschauer, M. (1996). Computer-Assisted Language Learning: An Introduction. In S. Fotos (A c. Di), *Multimedia Language Teaching* (pp. 3–20). Logos International.
- Warschauer, M. (2002). A Developmental Perspective on Technology in Language Education. *TESOL Quarterly, 36*(3), 453–475.
- Warschauer, M., & Healey, D. (1998). Computers and Language Learning: An Overview. *Language Teaching, 31*(2), 57–71.
- Warschauer, M., & Kern, R. (2000). *Network-Based Language Teaching: Concepts and Practice*. Cambridge University Press.
- Watanabe, Y. (1997). Input, Intake, and Retention: Effects of Increased Processing on Incidental Learning of Foreign Language Vocabulary. *Studies in Second Language Acquisition, 19*(3), 287–307.
- Webb, S., & Nation, I. S. P. (2017). *How Vocabulary is Learned*. Oxford University Press.
- White, L. (1989). *Universal Grammar and Second Language Acquisition*. John Benjamins.
- White, L. (2003). *Second Language Acquisition and Universal Grammar*. Cambridge University Press.
- Wible, D., & Tsao, N.-L. (2010). Incorporating NLP Techniques into Computer-Assisted Language Learning: Principles and Practices. *Computer Assisted Language Learning, 23*(4), 277–294.
- Wierzbicka, A. (1996). *Semantics: Primes and Universals*. Oxford University Press.

- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., & others. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3(1), 1–9.
- Willis, D., & Willis, J. R. (2007). *Doing task-based teaching*. Oxford University Press.
- Winke, P., Gass, S., & Sydorenko, T. (2010). The Effects of Captioning Videos Used for Foreign Language Listening Activities. *Language Learning & Technology*, 14(1), 65–86.
- Winkler, W. E. (1990). String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage. *Proceedings of the Section on Survey Research Methods*, 354–359.
- Witkin, H. A., Moore, C. A., Goodenough, D. R., & Cox, P. W. (1977). *Field-Dependent and Field-Independent Cognitive Styles and Their Educational Implications*. ERIC.
- Wolter, B. (2001). Comparative Lexical Networks of Japanese L2 English Learners and Native Speakers. *Studies in Second Language Acquisition*, 23(1), 41–69.
- Wong, W. (2001). Modality and Attention to Meaning and Form in the Input. *Studies in Second Language Acquisition*, 23(3), 345–368.
- Woolley, D. R. (1994). PLATO: The Emergence of Online Community. In L. Harasim (A c. Di), *Computers, Communications, and Mental Models* (pp. 91–108). MIT Press.
- World Wide Web Consortium. (2018). *Web Content Accessibility Guidelines (WCAG) 2.1*. <https://www.w3.org/TR/WCAG21/>
- World Wide Web Consortium. (2023). *Web Content Accessibility Guidelines (WCAG) 2.2*. <https://www.w3.org/TR/WCAG22/>
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge University Press.
- Yangarber, R., Katinskaia, A., Kisselev, O., & Németh, R. (2022). Using Revita for Language Learning: From Learner Modeling to Proficiency Tracking. *Proceedings of the 14th International Conference on Computer-Supported Education (CSEDU 2022)*, 195–206.

- Year, J., & Gordon, P. (2009). Korean Speakers' Acquisition of the English Ditransitive Construction: The Role of Verb Prototype, Input Distribution, and Frequency. *The Modern Language Journal*, 93(3), 399–417. <https://doi.org/10.1111/j.1540-4781.2009.00898.x>
- Young, D. J. (1990). An Investigation of Students' Perspective on Anxiety and Speaking. *Foreign Language Annals*, 23, 539–553.
- Zaveri, A., Rula, A., Maurino, A., Pietrobon, R., Lehmann, J., & Auer, S. (2016). Quality assessment for Linked Data: A Survey. *Semantic Web*, 7(1), 63–93. <https://doi.org/10.3233/SW-150175>
- Zhao, Y., Guo, Q., & Dynia, J. M. (2017). SmartReader: Designing a Mobile Reading Platform to Support Second Language Vocabulary Learning. *Journal of Educational Technology Development and Exchange*, 10(1), 53–68.
- Zhou, A., & Lu, X. (2010). The Effect of Textual Enhancement on the Learning of English Collocations. *System*, 38(4), 581–591.
- Zilio, L., Wilkens, R., & Fairon, C. (2017a). SMILLE: A System for Monitoring and Identifying Language Learning Errors. *Proceedings of the 12th Workshop on Innovative Use of NLP for Building Educational Applications (BEA)*, 170–179.
- Zilio, L., Wilkens, R., & Fairon, C. (2017b). Using NLP for Enhancing Second Language Acquisition. *RANLP 2017 - Recent Advances in Natural Language Processing Meet Deep Learning*, 839–846. https://doi.org/10.26615/978-954-452-049-6_107
- Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory Into Practice*, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2
- Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Addison-Wesley.

Appendix

A. Profiling questionnaire

Comune di Fisciano

Corso di Inglese – Questionario

Questo breve questionario servirà ad acquisire alcune informazioni utili all'organizzazione del corso di inglese che si terrà presso il Comune di Fisciano. I dati raccolti saranno trattati in conformità al Regolamento Generale sulla Protezione dei Dati (GDPR) e utilizzati esclusivamente per le finalità sopra indicate.

Informazioni personali

1. Email*

.....

2. Nome*

.....

3. Cognome*

.....

4. Età*

.....

5. Sesso*

☐ Uomo

☐ Donna

☐ Altro:

6. Email*

.....

7. Nazionalità*

- ☐ Italiana
- ☐ Altro:

8. Lingua parlata in famiglia*

- ☐ Italiano
- ☐ Italiano / Dialetto
- ☐ Dialetto
- ☐ Italiano / altra lingua
- ☐ Altra lingua
- ☐ Altro:

9. Se altra, quale/i?

.....

10. Qual è stata l'ultima occasione in cui hai studiato l'inglese?*

(Es. scuola superiore, corso di inglese, corso universitario, apprendimento autonomo...)

.....

11. Quanto tempo fa?

.....

12. Attraverso quale lingua ti riesce più facile comunicare?*

.....

13. In quale abilità pensi di dover migliorare il tuo inglese?*

- ☐ Scritto
- ☐ Parlato
- ☐ Comprensione
- ☐ Tutte le precedenti

14. Apprendere da contenuti di tuo interesse e in lingua inglese sarebbe di tuo gradimento?*☐ Sì☐ No

Lingua inglese**15. Capire l'inglese – Come definisci la tua capacità?**

- ☐ Non ho difficoltà a capire dal vivo o trasmissioni anche se veloci
- ☐ Riesco a capire discorsi lunghi, film e trasmissioni senza sottotitoli
- ☐ Capisco discorsi e argomentazioni complesse se il tema mi è familiare
- ☐ Capisco l'essenziale di trasmissioni lente e chiare
- ☐ Capisco facilmente ciò che mi riguarda direttamente, messaggi e annunci brevi
- ☐ Riconosco parole familiari ed espressioni semplici

16. Scrivere in inglese – Come definisci la tua capacità?

- ☐ So scrivere brevi frasi e compilare moduli
- ☐ Riesco a scrivere brevi messaggi o appunti
- ☐ Riesco a scrivere testi semplici per raccontare esperienze
- ☐ Riesco a scrivere saggi e relazioni articolate
- ☐ Riesco a scrivere testi complessi evidenziando i punti chiave
- ☐ Riesco a scrivere testi lunghi e riassumere contenuti complessi

17. Leggere in inglese – Come definisci la tua capacità di lettura?

- ☐ Capisco frasi semplici (cartelloni, cataloghi, annunci)
- ☐ Comprendo testi molto brevi e semplici
- ☐ Capisco testi di uso corrente (lavoro, avvenimenti, sentimenti)
- ☐ Riesco a leggere articoli e relazioni su questioni d'attualità

- ☐ Capisco testi lunghi e complessi, incluse istruzioni tecniche
 - ☐ Capisco praticamente tutte le forme di lingua scritta
-

Test linguistico**18. They have a lot of *time***

- ☐ money
- ☐ food
- ☐ hours
- ☐ friends

19. Where did you go?

- ☐ at what time
- ☐ for what reason
- ☐ in which way
- ☐ to what place

20. I have *many*

- ☐ none
- ☐ enough
- ☐ one
- ☐ a lot

21. All the children *grew*

- ☐ drew pictures
- ☐ spoke
- ☐ became bigger
- ☐ cried a lot

22. He is in the *camp*

- ☐ sea
- ☐ place outside where people enjoy nature

☐ hospital☐ building where people sleep**23. I am *upset***☐ tired☐ famous☐ rich☐ unhappy**24. I can *deal with it***☐ fix☐ remember☐ find☐ see**25. Is this the right *figure*?**☐ answer☐ place☐ number☐ time**26. This was used as the *basis***☐ answer☐ place to take a rest☐ next step☐ main part**27. Where is your *shoe*?**☐ the person who looks after you☐ the thing you keep your money in☐ the thing you use for writing☐ the thing you wear on your foot

Esercizio: Completa i verbi**28. Albert Einstein**

(1) _____ (be) a famous physicist.

He (2) _____ (develop) the theory of relativity.

Einstein (3) _____ (be) born in Ulm, a city in Germany.

As a young boy, he (4) _____ (be) curious about the nature of light and (5) _____ (spend) hours experimenting.

Throughout his career, Einstein (6) _____ (write) many scientific papers and (7) _____ (give) lectures around the world.

His ideas (8) _____ (change) how we understand time and space.

Einstein (9) _____ (believe) in using science for the betterment of humanity.

He (10) _____ (pass away) in 1955, but his legacy (11) _____ (continue) to influence modern physics.

☐ was

☐ developed

☐ was

☐ was

☐ spent

☐ wrote

☐ gave

☐ changed

☐ believed

☐ passed away

☐ continues

Table 16 Profiling assessment questionnaire - This questionnaire has been created solely for organizational and educational purposes related to the English course promoted by the Municipality of Fisciano. The information collected will be used exclusively for educational purposes and will be processed in compliance with the General Data Protection Regulation (GDPR – EU 2016/679).

B. Art. 5 GDPR Principles relating to processing of personal data

- 1) Personal data shall be:
 - a) processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency');
 - b) collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes ('purpose limitation');
 - c) adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation');
 - d) accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay ('accuracy');
 - e) kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject ('storage limitation');
 - f) processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures ('integrity and confidentiality').
- 2) The controller shall be responsible for, and be able to demonstrate compliance with, paragraph 1 ('accountability').

C. Licensing

As mentioned in the main text, the content generated by the web platform adopts an open-access licensing framework to ensure transparency, accessibility, and the free exchange of knowledge. All materials presented in, or generated through, this platform are released under the Creative Commons

Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0 - <https://creativecommons.org/licenses/by-sa/4.0/>).

This license permits users to copy, redistribute, remix, transform, and build upon the material for any purpose, including commercial use, provided that appropriate credit is given to the original authors, a link to the license is included, and any derivative works are distributed under the same terms.

As stated above, we hereby report the full version of the legal code attributed to the materials presented and produced within this platform:

Attribution-ShareAlike 4.0 International

=====

Creative Commons Corporation ("Creative Commons") is not a law firm and does not provide legal services or legal advice. Distribution of Creative Commons public licenses does not create a lawyer-client or other relationship. Creative Commons makes its licenses and related information available on an "as-is" basis. Creative Commons gives no warranties regarding its licenses, any material licensed under their terms and conditions, or any related information. Creative Commons disclaims all liability for damages resulting from their use to the fullest extent possible.

Using Creative Commons Public Licenses

Creative Commons public licenses provide a standard set of terms and conditions that creators and other rights holders may use to share original works of authorship and other material subject to copyright and certain other rights specified in the public license below. The following considerations are for informational purposes only, are not exhaustive, and do not form part of our licenses.

Considerations for licensors: Our public licenses are intended for use by those authorized to give the public permission to use material in ways otherwise restricted by copyright and certain other rights. Our licenses are irrevocable. Licensors should read and understand the terms and conditions of the license they choose before applying it. Licensors should also secure all rights necessary before applying our licenses so that the public can reuse the material as expected. Licensors should clearly mark any material not subject to the license. This includes other CC-licensed material, or material used under an exception or limitation to copyright. More considerations for licensors: wiki.creativecommons.org/Considerations_for_licensors

Considerations for the public: By using one of our public licenses, a licensor grants the public permission to use the licensed material under specified terms and conditions. If the licensor's permission is not necessary for any reason--for example, because of any applicable exception or limitation to copyright--then that use is not regulated by the license. Our licenses grant only permissions under copyright and certain other rights that a licensor has authority to grant. Use of the licensed material may still be restricted for other reasons, including because others have copyright or other

rights in the material. A licensor may make special requests, such as asking that all changes be marked or described. Although not required by our licenses, you are encouraged to respect those requests where reasonable. More considerations for the public:

wiki.creativecommons.org/Considerations_for_licensees

=====

Creative Commons Attribution-ShareAlike 4.0 International Public License

By exercising the Licensed Rights (defined below), You accept and agree to be bound by the terms and conditions of this Creative Commons Attribution-ShareAlike 4.0 International Public License ("Public License"). To the extent this Public License may be interpreted as a contract, You are granted the Licensed Rights in consideration of Your acceptance of these terms and conditions, and the Licensor grants You such rights in consideration of benefits the Licensor receives from making the Licensed Material available under these terms and conditions.

Section 1 -- Definitions.

- a. Adapted Material means material subject to Copyright and Similar Rights that is derived from or based upon the Licensed Material and in which the Licensed Material is translated, altered, arranged, transformed, or otherwise modified in a manner requiring permission under the Copyright and Similar Rights held by the Licensor. For purposes of this Public License, where the Licensed Material is a musical work, performance, or sound recording, Adapted Material is always produced where the Licensed Material is synched in timed relation with a moving image.
- b. Adapter's License means the license You apply to Your Copyright and Similar Rights in Your contributions to Adapted Material in accordance with the terms and conditions of this Public License.
- c. BY-SA Compatible License means a license listed at creativecommons.org/compatiblelicenses, approved by Creative Commons as essentially the equivalent of this Public License.
- d. Copyright and Similar Rights means copyright and/or similar rights closely related to copyright including, without limitation, performance, broadcast, sound recording, and Sui Generis Database Rights, without regard to how the rights are labeled or categorized. For purposes of this Public License, the rights specified in Section 2(b)(1)-(2) are not Copyright and Similar Rights.
- e. Effective Technological Measures means those measures that, in the absence of proper authority, may not be circumvented under laws fulfilling obligations under Article 11 of the WIPO Copyright Treaty adopted on December 20, 1996, and/or similar international agreements.
- f. Exceptions and Limitations means fair use, fair dealing, and/or any other exception or limitation to Copyright and Similar Rights that applies to Your use of the Licensed Material.
- g. License Elements means the license attributes listed in the name

of a Creative Commons Public License. The License Elements of this Public License are Attribution and ShareAlike.

- h. Licensed Material means the artistic or literary work, database, or other material to which the Licensor applied this Public License.
- i. Licensed Rights means the rights granted to You subject to the terms and conditions of this Public License, which are limited to all Copyright and Similar Rights that apply to Your use of the Licensed Material and that the Licensor has authority to license.
- j. Licensor means the individual(s) or entity(ies) granting rights under this Public License.
- k. Share means to provide material to the public by any means or process that requires permission under the Licensed Rights, such as reproduction, public display, public performance, distribution, dissemination, communication, or importation, and to make material available to the public including in ways that members of the public may access the material from a place and at a time individually chosen by them.
- l. Sui Generis Database Rights means rights other than copyright resulting from Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, as amended and/or succeeded, as well as other essentially equivalent rights anywhere in the world.
- m. You means the individual or entity exercising the Licensed Rights under this Public License. Your has a corresponding meaning.

Section 2 -- Scope.

a. License grant.

1. Subject to the terms and conditions of this Public License, the Licensor hereby grants You a worldwide, royalty-free, non-sublicensable, non-exclusive, irrevocable license to exercise the Licensed Rights in the Licensed Material to:
 - a. reproduce and Share the Licensed Material, in whole or in part; and
 - b. produce, reproduce, and Share Adapted Material.
2. Exceptions and Limitations. For the avoidance of doubt, where Exceptions and Limitations apply to Your use, this Public License does not apply, and You do not need to comply with its terms and conditions.
3. Term. The term of this Public License is specified in Section 6(a).
4. Media and formats; technical modifications allowed. The Licensor authorizes You to exercise the Licensed Rights in all media and formats whether now known or hereafter created, and to make technical modifications necessary to do so. The Licensor waives and/or agrees not to assert any right or authority to forbid You from making technical modifications necessary to exercise the Licensed Rights, including

technical modifications necessary to circumvent Effective Technological Measures. For purposes of this Public License, simply making modifications authorized by this Section 2(a)(4) never produces Adapted Material.

5. Downstream recipients.

- a. Offer from the Licensor -- Licensed Material. Every recipient of the Licensed Material automatically receives an offer from the Licensor to exercise the Licensed Rights under the terms and conditions of this Public License.
- b. Additional offer from the Licensor -- Adapted Material. Every recipient of Adapted Material from You automatically receives an offer from the Licensor to exercise the Licensed Rights in the Adapted Material under the conditions of the Adapter's License You apply.
- c. No downstream restrictions. You may not offer or impose any additional or different terms or conditions on, or apply any Effective Technological Measures to, the Licensed Material if doing so restricts exercise of the Licensed Rights by any recipient of the Licensed Material.

6. No endorsement. Nothing in this Public License constitutes or may be construed as permission to assert or imply that You are, or that Your use of the Licensed Material is, connected with, or sponsored, endorsed, or granted official status by, the Licensor or others designated to receive attribution as provided in Section 3(a)(1)(A)(i).

b. Other rights.

1. Moral rights, such as the right of integrity, are not licensed under this Public License, nor are publicity, privacy, and/or other similar personality rights; however, to the extent possible, the Licensor waives and/or agrees not to assert any such rights held by the Licensor to the limited extent necessary to allow You to exercise the Licensed Rights, but not otherwise.
2. Patent and trademark rights are not licensed under this Public License.
3. To the extent possible, the Licensor waives any right to collect royalties from You for the exercise of the Licensed Rights, whether directly or through a collecting society under any voluntary or waivable statutory or compulsory licensing scheme. In all other cases the Licensor expressly reserves any right to collect such royalties.

Section 3 -- License Conditions.

Your exercise of the Licensed Rights is expressly made subject to the following conditions.

a. Attribution.

1. If You Share the Licensed Material (including in modified

form), You must:

- a. retain the following if it is supplied by the Licensor with the Licensed Material:
 - i. identification of the creator(s) of the Licensed Material and any others designated to receive attribution, in any reasonable manner requested by the Licensor (including by pseudonym if designated);
 - ii. a copyright notice;
 - iii. a notice that refers to this Public License;
 - iv. a notice that refers to the disclaimer of warranties;
 - v. a URI or hyperlink to the Licensed Material to the extent reasonably practicable;
 - b. indicate if You modified the Licensed Material and retain an indication of any previous modifications; and
 - c. indicate the Licensed Material is licensed under this Public License, and include the text of, or the URI or hyperlink to, this Public License.
2. You may satisfy the conditions in Section 3(a)(1) in any reasonable manner based on the medium, means, and context in which You Share the Licensed Material. For example, it may be reasonable to satisfy the conditions by providing a URI or hyperlink to a resource that includes the required information.
 3. If requested by the Licensor, You must remove any of the information required by Section 3(a)(1)(A) to the extent reasonably practicable.

b. ShareAlike.

In addition to the conditions in Section 3(a), if You Share Adapted Material You produce, the following conditions also apply.

1. The Adapter's License You apply must be a Creative Commons license with the same License Elements, this version or later, or a BY-SA Compatible License.
2. You must include the text of, or the URI or hyperlink to, the Adapter's License You apply. You may satisfy this condition in any reasonable manner based on the medium, means, and context in which You Share Adapted Material.
3. You may not offer or impose any additional or different terms or conditions on, or apply any Effective Technological Measures to, Adapted Material that restrict exercise of the rights granted under the Adapter's License You apply.

Section 4 -- Sui Generis Database Rights.

Where the Licensed Rights include Sui Generis Database Rights that

apply to Your use of the Licensed Material:

- a. for the avoidance of doubt, Section 2(a)(1) grants You the right to extract, reuse, reproduce, and Share all or a substantial portion of the contents of the database;
- b. if You include all or a substantial portion of the database contents in a database in which You have Sui Generis Database Rights, then the database in which You have Sui Generis Database Rights (but not its individual contents) is Adapted Material, including for purposes of Section 3(b); and
- c. You must comply with the conditions in Section 3(a) if You Share all or a substantial portion of the contents of the database.

For the avoidance of doubt, this Section 4 supplements and does not replace Your obligations under this Public License where the Licensed Rights include other Copyright and Similar Rights.

Section 5 -- Disclaimer of Warranties and Limitation of Liability.

- a. UNLESS OTHERWISE SEPARATELY UNDERTAKEN BY THE LICENSOR, TO THE EXTENT POSSIBLE, THE LICENSOR OFFERS THE LICENSED MATERIAL AS-IS AND AS-AVAILABLE, AND MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND CONCERNING THE LICENSED MATERIAL, WHETHER EXPRESS, IMPLIED, STATUTORY, OR OTHER. THIS INCLUDES, WITHOUT LIMITATION, WARRANTIES OF TITLE, MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, NON-INFRINGEMENT, ABSENCE OF LATENT OR OTHER DEFECTS, ACCURACY, OR THE PRESENCE OR ABSENCE OF ERRORS, WHETHER OR NOT KNOWN OR DISCOVERABLE. WHERE DISCLAIMERS OF WARRANTIES ARE NOT ALLOWED IN FULL OR IN PART, THIS DISCLAIMER MAY NOT APPLY TO YOU.
- b. TO THE EXTENT POSSIBLE, IN NO EVENT WILL THE LICENSOR BE LIABLE TO YOU ON ANY LEGAL THEORY (INCLUDING, WITHOUT LIMITATION, NEGLIGENCE) OR OTHERWISE FOR ANY DIRECT, SPECIAL, INDIRECT, INCIDENTAL, CONSEQUENTIAL, PUNITIVE, EXEMPLARY, OR OTHER LOSSES, COSTS, EXPENSES, OR DAMAGES ARISING OUT OF THIS PUBLIC LICENSE OR USE OF THE LICENSED MATERIAL, EVEN IF THE LICENSOR HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH LOSSES, COSTS, EXPENSES, OR DAMAGES. WHERE A LIMITATION OF LIABILITY IS NOT ALLOWED IN FULL OR IN PART, THIS LIMITATION MAY NOT APPLY TO YOU.
- c. The disclaimer of warranties and limitation of liability provided above shall be interpreted in a manner that, to the extent possible, most closely approximates an absolute disclaimer and waiver of all liability.

Section 6 -- Term and Termination.

- a. This Public License applies for the term of the Copyright and Similar Rights licensed here. However, if You fail to comply with this Public License, then Your rights under this Public License terminate automatically.
- b. Where Your right to use the Licensed Material has terminated under Section 6(a), it reinstates:
 1. automatically as of the date the violation is cured, provided it is cured within 30 days of Your discovery of the violation; or

2. upon express reinstatement by the Licensor.

For the avoidance of doubt, this Section 6(b) does not affect any right the Licensor may have to seek remedies for Your violations of this Public License.

- c. For the avoidance of doubt, the Licensor may also offer the Licensed Material under separate terms or conditions or stop distributing the Licensed Material at any time; however, doing so will not terminate this Public License.
- d. Sections 1, 5, 6, 7, and 8 survive termination of this Public License.

Section 7 -- Other Terms and Conditions.

- a. The Licensor shall not be bound by any additional or different terms or conditions communicated by You unless expressly agreed.
- b. Any arrangements, understandings, or agreements regarding the Licensed Material not stated herein are separate from and independent of the terms and conditions of this Public License.

Section 8 -- Interpretation.

- a. For the avoidance of doubt, this Public License does not, and shall not be interpreted to, reduce, limit, restrict, or impose conditions on any use of the Licensed Material that could lawfully be made without permission under this Public License.
- b. To the extent possible, if any provision of this Public License is deemed unenforceable, it shall be automatically reformed to the minimum extent necessary to make it enforceable. If the provision cannot be reformed, it shall be severed from this Public License without affecting the enforceability of the remaining terms and conditions.
- c. No term or condition of this Public License will be waived and no failure to comply consented to unless expressly agreed to by the Licensor.
- d. Nothing in this Public License constitutes or may be interpreted as a limitation upon, or waiver of, any privileges and immunities that apply to the Licensor or You, including from the legal processes of any jurisdiction or authority.

=====

Creative Commons is not a party to its public licenses. Notwithstanding, Creative Commons may elect to apply one of its public licenses to material it publishes and in those instances will be considered the "Licensor." The text of the Creative Commons public licenses is dedicated to the public domain under the CC0 Public Domain Dedication. Except for the limited purpose of indicating that material is shared under a Creative Commons public license or as otherwise permitted by the Creative Commons policies published at creativecommons.org/policies, Creative Commons does not authorize the use of the trademark "Creative Commons" or any other trademark or logo

of Creative Commons without its prior written consent including, without limitation, in connection with any unauthorized modifications to any of its public licenses or any other arrangements, understandings, or agreements concerning use of licensed material. For the avoidance of doubt, this paragraph does not form part of the public licenses.

Creative Commons may be contacted at creativecommons.org.

D. Didactic plan overview

This didactic plan presents a structured and interactive approach to English language learning, designed to progressively develop students' linguistic competence. Each of the eight lessons follows a coherent framework that integrates listening, speaking, reading, and writing activities, while promoting active engagement and learner autonomy. The methodological orientation emphasizes communication, interaction, and independent knowledge construction through the practical and meaningful use of language. Grounded in established pedagogical theories, the plan combines traditional and digital tools to create a dynamic and inclusive learning environment that accommodates diverse learning styles.

Designed for learners at the A2–B1 proficiency levels, as assessed through the profiling questionnaire, the plan ensures that linguistic objectives and tasks are appropriately calibrated to the learners' abilities and communicative needs. As a complementary lesson plan, it was implemented with the informants to support and enhance autonomous learning, offering additional opportunities for self-directed practice and reflection.

It is important to note that the content of each activity across the lessons was not predetermined, but rather automatically generated in real time based on the main theme or the text selected from those proposed by the platform. This adaptive design ensures flexibility, contextual relevance, and responsiveness to learners' individual needs and progression.

Overall, the lessons aim to foster linguistic confidence, motivation, and intercultural awareness, establishing a solid foundation for continued development in English language proficiency.

1. Lesson 1: Introduction to English Phonetics and the Alphabet

a. Lesson Plan Overview

The primary objectives of this lesson are to introduce the English alphabet and its pronunciation, as well as to teach basic vocabulary. This vocabulary includes numbers from 1 to 20, days of the week, countries and nationalities, personal information, and common objects. By the end of the lesson, students will have a solid foundation in these essential areas of language learning, which will serve as the basis for more complex English concepts in future lessons.

b. Methodology

The lesson follows a multi-sensory approach to ensure effective learning. By incorporating both auditory and visual components, it facilitates the acquisition of phonetic and vocabulary skills. Students will be engaged through interactive tools that combine sound and visual cues, helping them associate letters and words with their corresponding sounds. For instance, the lesson utilizes search engines to highlight contextual sounds, offering a way to reinforce correct pronunciation with visual references.

Active repetition plays a crucial role in this methodology, with students encouraged to practice sounds and words repeatedly to aid memorization. Moreover, the lesson integrates an online platform featuring gamified interactive activities, which promote autonomous learning. These activities provide immediate feedback, allowing students to self-correct and enhance their understanding. This feedback loop supports the learning process by giving students an opportunity to refine their skills and gain confidence in using English.

c. Content Breakdown

The lesson begins with an Introduction to the Course, welcoming students and encouraging them to see language as a way to connect with and describe the world. This helps build motivation and a sense of belonging. In Alphabet and Pronunciation, students learn to recognize and pronounce each letter of the English alphabet, with examples to illustrate sounds. Basic vocabulary is also introduced, including numbers (1–20), days of the week, countries and nationalities, personal information, and common objects.

The Alphabet Overview deepens understanding by comparing the English alphabet to others, such as Italian, helping students grasp spelling and writing fundamentals.

During Interactive Alphabet Practice, students complete matching, fill-in-the-blank, and pronunciation recording activities. Immediate feedback allows them to self-correct and strengthen their skills.

In Numbers and Practice, students learn to recognize, pronounce, and associate numbers with images to reinforce comprehension and fluency.

The Days of the Week section introduces each day's name and order, supported by visual and routine-based activities.

In Countries and Nationalities, students explore world cultures through interactive map and flag-matching exercises, linking language learning with cultural awareness.

The lesson concludes with a Lesson Summary, reviewing key topics and encouraging reflection, continued practice, and confidence in using English in everyday situations.

d. Interactive Phase

During the Interactive Phase, students will engage in a series of exercises designed to reinforce their learning through active participation. One of the first activities is the Sound-Letter Matching Exercise, where students will match sounds with their corresponding letters. This will help students solidify their understanding of the phonetic structure of English.

Next, the Letter-Image Matching Exercise will require students to match letters with images representing words that start with those letters. This exercise reinforces vocabulary acquisition in a hands-on, interactive way.

Another key activity is the Fill-in-the-Blank Exercise, where students will complete words by filling in the initial letter of a word depicted in an image. This exercise emphasizes the connection between sounds and letters, reinforcing spelling and pronunciation skills.

Lastly, Pronunciation Recording and Comparison will allow students to record their own pronunciation and compare it with the correct pronunciation provided by the platform. This will give them a chance to assess their own progress and make adjustments as needed.

e. Consolidation Phase

In the Consolidation Phase, students will take part in engaging activities to review and strengthen what they've learned. A Phonetic Crossword Puzzle will help them recall letters and sounds, while a Memory Game will reinforce their ability to match letters with their corresponding sounds.

Each vocabulary section includes targeted practice to support understanding. In Numbers (1–20), students will complete Number-Image Matching Exercises to connect numbers with visuals. In Countries and Nationalities, they will explore an Interactive Map and a Country-Flag Matching Game to link places, nationalities, and flags. For Personal Information, students will create Information Cards with their name, age, and city, promoting self-expression. Finally, in Common Objects, a Memory Game will help them pair images with words, reinforcing everyday vocabulary in a fun way.

f. Didactic choices

The lesson plan uses a variety of teaching methods designed to make language learning engaging, effective, and inclusive. A multi-sensory approach, combining visual, auditory, and kinesthetic elements, helps students connect sounds, letters, and words through different channels, improving understanding and memory. Activities such as sound-letter matching, alphabet overviews, and contextual listening tasks build phonemic awareness and expose learners to real examples of language use.

Interactive and game-based exercises like crosswords, memory games, and pronunciation practice keep students motivated and provide instant feedback, encouraging active participation and self-correction. Repetition and association activities help reinforce vocabulary and sound patterns, making it easier for students to recall and apply them later.

The lesson also promotes independent learning by allowing students to practice at their own pace and receive feedback online, building confidence and responsibility for their progress. Cultural elements, like exploring countries, nationalities, and flags, broaden students' understanding of the world and foster appreciation for diversity.

Finally, the lesson is structured with clear stages and gradual support, guiding learners from simple recognition to independent use of new language skills. This step-by-step approach ensures that students feel supported as they build their confidence and language competence.

2. Lesson 2: The Verb *To Be*

a. Lesson Plan Overview

This lesson is designed to enhance students' ability to use pronouns and the Present Simple tense correctly and confidently. Through structured exposure and practical application, students will learn to construct grammatically accurate sentences and expand their vocabulary related to daily routines and familiar objects. By the end of the session, learners should be able to apply subject and object pronouns in context, form both affirmative and negative Present Simple sentences, and incorporate relevant vocabulary into meaningful communication.

b. Methodology

The lesson adopts an inductive teaching approach, where students encounter patterns and examples before deriving grammatical rules. This method encourages active participation and supports deeper conceptual understanding. Rather than being presented with explicit rules upfront, students discover linguistic structures through guided exploration.

Instruction is interactive and student-centered. Pronoun identification tasks allow learners to see how pronouns function in real sentences, replacing nouns to reduce redundancy. Sentence construction exercises provide practice in forming complete statements using pronouns and Present Simple verb forms. Flashcards are used to reinforce recognition and understanding of correct pronoun-verb combinations. Vocabulary is introduced thematically, with a focus on daily activities and familiar objects, ensuring the language is functional and relevant. These components together support retention and the practical application of language skills.

c. Content Breakdown

The lesson starts with an introduction to pronouns, covering subject pronouns (I, you, he, she, it, we, they) and object pronouns (me, you, him, her, it, us, them). Students learn how pronouns replace nouns for clearer, more natural communication, for example, replacing Mary with she in She is reading.

Next, students study Present Simple sentence structure. They practice forming affirmative sentences (e.g., They play soccer) and negative ones using do not or does not (e.g., He does not like apples). Controlled and guided exercises help reinforce these patterns.

A brief review follows, focusing on recognizing and producing correct Present Simple forms, such as She works and They do not work.

Finally, students expand their vocabulary in two areas: daily activities (eat, go, study, play; breakfast, school, game) and common objects (pen, book, computer, chair). They use new words in simple sentences like I eat breakfast or She reads a book, combining grammar and vocabulary practice in meaningful contexts.

d. Interactive Phase

This phase of the lesson emphasizes practical application through structured, hands-on activities. Students begin with a pronoun identification task, where they analyze sentences to determine and apply the correct pronoun. This activity ensures they understand both the function and appropriate context for each pronoun.

Following this, learners complete sentence construction exercises that require them to fill in the correct pronoun or verb form. This reinforces their grasp of grammatical structure and encourages accurate sentence

formation. Flashcards are used interactively, allowing students to match pronouns with appropriate sentence parts, thus enhancing visual learning and reinforcing grammatical understanding.

To consolidate learning, students write original sentences using the Present Simple tense and vocabulary introduced earlier. They also participate in role-playing scenarios that simulate real-life conversations, giving them opportunities to apply grammar and vocabulary in context. The lesson concludes with interactive games that review and reinforce pronoun usage and sentence formation. These activities provide repetition in a dynamic format, supporting engagement and long-term retention.

e. Didactic Choices

The lesson is designed around key teaching principles that promote effective language learning. Using an inductive approach, students discover grammar rules by observing patterns and using language in context, encouraging active participation and deeper understanding. Interactive exercises and repetition reinforce learning and build fluency.

Flashcards and role-play activities help students use vocabulary and grammar naturally, focusing on meaningful language in real situations. The selected vocabulary emphasizes practical, everyday communication, keeping lessons relevant and motivating.

Multisensory activities like matching, identifying, and classifying words combine visual and verbal input, supporting different learning styles and helping students retain information more easily.

The lesson follows a scaffolded structure, starting with simple concepts and moving to more complex tasks. This gradual progression builds confidence and independence. Collaborative and interactive activities provide repeated opportunities to use new language in meaningful contexts.

Overall, the lesson aims to develop both accuracy and confidence in communication by combining clear structure, active practice, and real-world language use.

3. Lesson 3: Nouns and Articles in English

a. Lesson Plan Overview

This lesson focuses on enhancing students' understanding of nouns and articles in English, the correct application of the Present Simple tense in both positive and negative forms, and expanding vocabulary around irregular plurals, jobs, and family relationships. The aim is to equip students with the skills necessary to form accurate and grammatically correct sentences using proper nouns and articles, while also enriching their vocabulary with important everyday terms. By the end of the lesson, students should feel confident using nouns and articles correctly in conversation and writing.

b. Methodology

The lesson adopts an inductive teaching method, beginning with an authentic text to expose students to real-world usage of nouns and articles in context. From this, students derive grammatical rules themselves, which promotes deeper understanding and retention. The lesson will incorporate several interactive activities to reinforce learning. First, students will engage in a Cloze Exercise, where they will fill in missing articles (definite and indefinite) within a text. This activity challenges students to recognize the appropriate use of articles and encourages them to apply the concepts in context.

Next, students will use Interactive Flashcards to match nouns with the correct articles. These flashcards help students visually and cognitively reinforce the rules for article use. The third activity, Grammar Treasure Hunt, involves solving clues that require students to identify the correct nouns and articles. The clues will be tailored to different contexts, pushing students to apply their knowledge in varied situations. To conclude, students will participate in Creative Writing, where they will be asked to write short descriptions or stories using the correct nouns and articles. This activity will allow students to practice applying their knowledge creatively and personalize the lesson content.

c. Content Breakdown

The lesson begins with an introduction to nouns and articles. Students learn that nouns name people, places, things, or ideas, and that articles help specify them. Both concrete nouns (e.g., book, car) and abstract nouns (e.g., love, freedom) are introduced. Students explore how to use the definite article (the) for specific nouns (The book on the table is mine) and indefinite articles (a, an) for general ones (a cat, an apple). Practical examples show how articles change meaning and help make sentences clearer.

Next, students review the Present Simple tense, used for regular actions and general truths. They practice forming affirmative sentences (I work, She plays) and negative ones using do/does + not (I do not like coffee, She does not enjoy reading). Exercises converting positive to negative sentences reinforce understanding and accuracy.

The following section introduces irregular plurals, focusing on nouns that don't follow the regular -s/-es rule, such as child → children, foot → feet, mouse → mice, and tooth → teeth. A short quiz helps students recognize and use these forms correctly in context.

In Jobs Vocabulary, students learn common job titles and their meanings, including teacher, doctor, engineer, and chef. They match titles with definitions and build simple sentences such as The doctor treats patients or She is an engineer, applying new vocabulary in context.

The final section covers Family Vocabulary, introducing terms like mother, father, sibling, and cousin. Students create family trees to visualize relationships and practice sentences such as My mother is a teacher or My cousin is younger than me. This helps them talk confidently about family and personal connections.

d. Interactive Phase

During the Interactive Phase, students will engage with various activities designed to help them apply the concepts they've learned. The first activity is the Cloze Exercise, where students will fill in missing articles (a, an, the) in a text. This task will help them practice identifying when and where articles are needed. In the Interactive Flashcards activity, students will match nouns with their corresponding articles. This exercise provides both visual and cognitive reinforcement of article usage. Next, in the Grammar Treasure Hunt, students will solve clues to practice identifying and using nouns and articles in different contexts. For example, a clue might challenge them to determine the correct article to use based on the noun provided, like '___ apple' or '___ books.'

Lastly, in the Creative Writing task, students will write short stories or descriptions using the correct nouns and articles they've learned. For example, they might describe their family, their daily routine, or their favorite hobbies, making sure to use the proper nouns and articles in each sentence. This activity allows students to apply their knowledge in a personalized way and express themselves in English.

e. Consolidation Phase

In the Consolidation Phase, students will revisit the Grammar Treasure Hunt activity, solving more clues that require them to apply their knowledge of nouns and articles. This reinforces the lesson's concepts in a fun and engaging way. The Creative Writing task will also be revisited, giving students the opportunity to refine their writing and reinforce their understanding of the lesson's grammar points.

f. Didactic choices

This lesson focuses on teaching nouns and articles while reinforcing the Present Simple tense and expanding vocabulary on irregular plurals, jobs, and family. Using an inductive approach, students analyze an authentic text to discover grammar rules, promoting active engagement and deeper understanding.

A cloze exercise lets students apply article rules in context, while interactive flashcards offer repeated, feedback-driven practice. A grammar treasure hunt adds a playful, game-based element that motivates learners to apply new knowledge.

Role-play activities help students describe people and objects using correct nouns and articles, strengthening practical communication. Creative writing tasks, such as short stories or descriptions, encourage free use of grammar and vocabulary, building confidence and fluency.

Practice on irregular plurals is reinforced through a short quiz, while jobs and family vocabulary activities, like matching, quizzes, and family tree creation, make learning interactive and memorable.

The interactive phase integrates grammar, vocabulary, and speaking through varied tasks such as cloze exercises, flashcards, and role-play. The consolidation phase focuses on creativity and real-world application through games and storytelling, supporting long-term retention.

Finally, a lesson summary reviews key points and promotes continued practice, emphasizing language learning as an ongoing process. Overall, the lesson balances structure, engagement, and real-life relevance to build confidence and proficiency in English.

4. Lesson 4: The Verb *To Have*

This lesson is designed to introduce the verb ‘to have’ in the present tense, focusing on its use in expressing possession, relationships, and experiences. Students will learn how to use the verb to talk about what they own, their relationships with others, and their experiences. The lesson will also expand vocabulary with terms related to free time activities, love stories, and musical genres and instruments.

a. Main Objectives

The primary objectives of the lesson are for students to become proficient in using the verb *to have* in the present tense, correctly constructing affirmative, negative, and interrogative sentences. Students will also expand their vocabulary by learning terms related to free time activities such as ‘watch TV,’ ‘read books,’ and ‘play sports.’ Additionally, they will be introduced to common phrases related to love stories, including ‘fall in love,’ ‘get married,’ and ‘have a date.’ Furthermore, students will learn vocabulary associated with musical genres and instruments.

b. Methodology

The lesson follows a clear structure, combining direct instruction with guided practice and opportunities for autonomous application. Much like the lesson on the verb ‘to be,’ the approach centers around using contextualized examples to illustrate the correct use of ‘to have.’ Activities are designed to engage students in playful and interactive ways, making the learning process fun and effective. Games like *Guess Who?* and exercises where students draw and describe fantasy characters provide opportunities to practice the verb ‘to have’ in meaningful contexts. These activities are intended to promote natural acquisition by encouraging students to use language in a relaxed and enjoyable environment.

c. Content Breakdown

In the introductory phase of the lesson, students are introduced to the forms and usage of the verb *to have* in various sentence structures. The lesson begins with an explanation of the different forms of ‘to have’: *have* is used with the subjects *I*, *you*, *we*, and *they*, while *has* is used with *he*, *she*, and *it*. For example, *I have a meeting tomorrow*, or *She has a beautiful voice*. The lesson also covers how to construct sentences in the present tense. Students will be introduced to affirmative, negative, and interrogative sentence structures. Affirmative sentences are used to state something that is true, such as *I have a book*. Negative sentences use ‘do not’ (don’t)

or ‘does not’ (doesn’t) to indicate something that is not true, like ‘They do not have any pets.’ In interrogative sentences, ‘do’ or ‘does’ is used to form questions, such as ‘Do you have any questions?’

d. Interactive Phase

The interactive phase provides students with opportunities to practice the verb ‘to have’ in context. One of the activities in this phase is a cloze exercise, where students complete sentences by filling in the blanks with the correct form of ‘to have.’ For example, students will fill in sentences like, ‘They ___ a new puppy,’ or ‘She ___ many friends at school.’ Another example could be, ‘Do you ___ any free time this weekend?’ This exercise reinforces the correct use of ‘have’ and ‘has’ in different contexts. Additionally, students will be given interactive flashcards to match subjects with the appropriate form of the verb ‘to have.’ The flashcards will include both vocabulary terms and sentences, allowing students to practice forming sentences using the new vocabulary.

e. Consolidation Phase

In the consolidation phase, the focus shifts to reinforcing what has been learned through fun and engaging activities. The game ‘Guess Who?’ will be used, where students describe a person using sentences with the verb ‘to have.’ For example, they might say, ‘This person has two pets,’ or ‘This person has a lot of friends.’ The goal of the game is to encourage students to use the verb ‘to have’ correctly while maintaining a fun, interactive atmosphere. Another activity in the consolidation phase is the ‘Game of the Goose,’ a board game where students roll dice and answer questions or complete sentences using the verb ‘to have.’ These games provide a hands-on way for students to apply their knowledge in an interactive and enjoyable setting.

f. Resources

A variety of interactive resources support this lesson to reinforce understanding and engagement. A digital and print cloze exercise allows students to fill in blanks and receive instant feedback, while interactive flashcards, online or physical, help them match subjects with the correct form of to have and related vocabulary. Games such as Guess Who? and The Game of the Goose provide fun, communicative opportunities to apply grammar in real contexts.

The lesson is structured to build a strong understanding of to have in expressing possession, relationships, and experiences, while expanding vocabulary on free time activities, love stories, and musical genres and instruments. The introduction focuses on mastering the present tense forms, have for I, you, we, they and has for he, she, it, through contextualized examples in affirmative, negative, and interrogative sentences (e.g., I have a book, They don’t have pets, Do you have any questions?).

During the interactive phase, students complete cloze exercises such as They ___ a new puppy or Do you ___ any free time this weekend?, connecting grammar with real-life communication. Flashcards reinforce recall through visual associations, promoting multisensory learning and long-term retention.

The consolidation phase emphasizes active participation. In Guess Who?, students describe people using to have sentences, improving fluency and listening skills. The Game of the Goose adds a playful, competitive element as students complete sentences with the correct verb form. Both activities reflect a communicative language teaching approach, prioritizing meaningful use of English.

All selected resources, cloze tasks, flashcards, and games, encourage repetition, feedback, and engagement, making grammar practice both effective and enjoyable.

Overall, this lesson offers a comprehensive and interactive approach to mastering to have in the present tense. By combining clear instruction with creative, contextualized practice, it enables students to use the structure confidently and naturally in everyday communication.

5. Lesson 5: The Past Simple Tense

a. Lesson Plan Overview

This lesson focuses on introducing the Past Simple tense in English, highlighting its importance for discussing past actions. It also aims to teach word formation using prefixes and suffixes, the use of past time expressions (e.g., yesterday, last week, ago), and the past forms of common verbs like ‘go,’ ‘have,’ and ‘get.’ Vocabulary related to daily routines will be explored as well. The lesson follows a step-by-step approach that moves from explanation and guided practice to independent application, ensuring students build a solid understanding of both the structure and usage of the Past Simple tense.

b. Main Objectives

The main objectives of the lesson are to help students understand and correctly use the Past Simple tense, both in affirmative and negative forms, as well as in questions. Students will also focus on word formation, particularly how suffixes and prefixes can change the meaning of words. They will learn past time expressions such as ‘yesterday,’ ‘last week,’ and ‘ago’ to indicate when actions occurred. Additionally, the lesson will provide practice using the Past Simple forms of the verbs *go*, *have*, and *get*, while expanding vocabulary related to daily routines (e.g., wake up, have breakfast, go to school/work).

c. Methodology

This lesson takes a gradual approach to learning the Past Simple tense. The lesson structure includes direct instruction, guided practice, and opportunities for independent application. Initially, the rules for forming the Past Simple tense will be presented clearly, and students will be given a series of interactive activities to practice their understanding. In the guided practice phase, cloze exercises will provide structured support to reinforce learning. In the autonomous practice phase, students will apply their knowledge independently, working on sentence transformation tasks and engaging in creative activities that require them to use the Past Simple tense in context. The goal is for students to feel comfortable with the tense and its applications by the end of the lesson.

d. Content Breakdown

The introductory phase focuses on forming and using the Past Simple tense. Students first learn to create affirmative sentences by using the past form of verbs, such as I played tennis yesterday or She went to the market. Regular verbs take -ed (e.g., play → played), while irregular verbs have unique forms that must be memorized (e.g., go → went).

Next, students study negative sentences, using did not (didn’t) plus the base form of the verb (I didn’t go to the park, They didn’t eat lunch). For interrogative sentences, they place did at the beginning, followed by the base form (Did you go to school?).

The section also includes word formation, showing how prefixes and suffixes change word meanings and help expand vocabulary.

Finally, students learn common past time expressions such as yesterday, last week, and ago, practicing them in context to describe when actions happened.

e. Interactive Phase

In the interactive phase, students will practice forming and using the Past Simple tense through a variety of exercises. One key activity is an interactive cloze exercise where students fill in the blanks with the correct past form of verbs. Example sentences might include ‘Yesterday, I _____ (go) to the market,’ or ‘Last week, they _____ (watch) a great movie.’

Additionally, students will participate in a flashcard game, where they match present tense verbs with their corresponding past tense forms. For example, they will match ‘go’ with ‘went’ and ‘have’ with ‘had.’ This exercise will help solidify their understanding of verb conjugations and improve their ability to recall the past forms quickly.

f. Consolidation Phase

In the consolidation phase, the focus shifts to reinforcing and applying the Past Simple tense through more complex activities. Students will work on sentence transformation exercises, where they convert affirmative sentences into negative sentences. For example, they might change ‘I went to the market yesterday’ into ‘I didn’t go to the market yesterday.’ This will help solidify their understanding of the negative structure in the Past Simple tense.

Students will also engage in autonomous practice, completing sentences in the Past Simple tense. They might fill in sentences like ‘Last weekend, I _____ (go) to the beach’ or ‘They _____ (get) a surprise gift last month.’ This phase encourages independent application of knowledge and helps students practice using the tense in more varied contexts.

g. Resources

This lesson on the Past Simple tense uses a range of interactive resources to support learning. A cloze exercise helps students practice verb forms, flashcards reinforce the link between present and past verbs, and a sentence transformation task guides them in forming negatives. An autonomous worksheet provides opportunities for independent practice. All materials are available in digital and print formats to suit different learning settings. The lesson begins with a clear introduction to the function and formation of the Past Simple, focusing on regular (play → played) and irregular (go → went) verbs, as well as past time expressions (yesterday, last week, ago). This phase establishes a foundation for accurate use in real communication.

In the interactive phase, students complete cloze sentences such as Yesterday, I ____ to the market or Two days ago, she ____ a letter to her friend, followed by a flashcard game to strengthen recall. These activities promote meaningful, communicative practice.

The consolidation phase focuses on applying knowledge through sentence transformation (I went to the market → I did not go to the market) and autonomous practice, encouraging self-correction and confidence.

Finally, creative writing and short presentations allow students to describe daily routines and past experiences, linking grammar to personal expression.

By combining explanation, guided practice, and creative use, the lesson builds both grammatical accuracy and communicative fluency, ensuring students can use the Past Simple confidently in real-life contexts.

6. Lesson 6: Vocabulary improvement and lessons recap

a. Lesson Plan Overview

This lesson focuses on vocabulary related to houses and furniture, prepositions of place, and places in a city. Additionally, students will practice forming and using the negative form of verbs in the Present Simple tense. The methodology of the lesson is based on direct instruction of grammatical rules, followed by guided practice and consolidation activities. To make learning more engaging, interactive games and activities such as the *True or False* game and error correction exercises are included.

b. Main Objectives

The primary objectives of this lesson are to help students learn and use vocabulary related to houses and furniture, prepositions of place, and places in a city. The lesson also aims to enable students to correctly form negative sentences using *do not* (don’t) and *does not* (doesn’t). Through interactive activities, students will practice the negative form of verbs in the Present Simple tense, reinforcing their understanding through games and contextual exercises.

c. Methodology

The lesson follows a structured approach, starting with direct instruction on grammatical rules. In the introductory phase, the rules for forming the negative form of verbs in the Present Simple tense will be clearly explained. Guided practice will follow, giving students opportunities to apply the rules in structured exercises. To further engage students, consolidation activities such as interactive games, error correction exercises, and a debate will provide opportunities for applying the learned content in real communicative contexts. This approach ensures that students not only understand grammatical rules but also become comfortable using them in practical, everyday situations.

d. Content Breakdown

The introductory phase introduces the negative form of verbs in the Present Simple tense. Students will learn how to form negative sentences using *do not* (don't) and *does not* (doesn't). For example, 'I do not like apples' or 'She does not go to the park.' This section will focus on helping students understand the difference between affirmative and negative sentences and how to construct them correctly.

The lesson will also introduce vocabulary related to houses and furniture (e.g., kitchen, bedroom, table, chair), prepositions of place (e.g., in, on, at, under, next to, in front of, behind, between), and places in a city (e.g., bank, post office, park, library). This vocabulary will be integrated into sentences using the negative form of verbs, allowing students to apply both the grammatical structure and new vocabulary together.

e. Interactive Phase

In the interactive phase, students will practice using the negative form of verbs in the Present Simple tense. One activity will involve transforming affirmative sentences into negative ones. For example, students will start with a sentence like 'I live in a house' and change it to 'I do not live in a house.'

Additionally, students will participate in a multiple-choice activity where they select the correct negative form of a verb. This exercise will help reinforce the rules of negation, such as choosing the correct form of 'do' or 'does' with different subjects (e.g., 'She does not like apples' vs. 'I do not like apples').

The vocabulary learned earlier will be incorporated into these activities, with students using prepositions and place-related terms in sentences like, 'The table is not in the kitchen' or 'The library is not next to the bank.'

f. Consolidation Phase

The consolidation phase focuses on reinforcing and applying the knowledge gained during the lesson. One activity will be the *True or False* game. In this game, students will hear sentences in the Present Simple tense, some in the affirmative and some in the negative. They must listen carefully and decide whether the sentences are true or false based on their understanding of the negative form of verbs. For example, if the sentence is 'She doesn't go to work on Sundays,' students must decide if the sentence matches what they know about her schedule.

Another key activity will involve error correction. Students will be provided with sentences that contain incorrect negative forms (e.g., 'I doesn't like pizza'). They will need to identify and correct the errors, reinforcing the rules of negation and helping to develop self-correction skills. This activity will help solidify students' understanding and their ability to apply the grammatical rules independently.

Finally, students will engage in a debate on a topic of interest. In this debate, they will be encouraged to use the negative form of verbs to express disagreement or opposing opinions. For example, 'I do not think this is a good idea' or 'I don't agree with that point.' This activity promotes spontaneous oral production and the use of the language in real communicative contexts, allowing students to practice their negative sentence structures in a natural and meaningful way.

g. Resources

The lesson will make use of various resources to support student learning. This includes:

Cloze Exercises: These exercises will involve transforming affirmative sentences into negative ones, with fill-in-the-blank options for students to complete.

Multiple-Choice Activity: A digital or printed quiz will allow students to choose the correct negative form of verbs.

Flashcards and Vocabulary Lists: Visual aids for learning the vocabulary related to houses, furniture, prepositions of place, and places in a city.

Error Correction Exercises: Worksheets or digital resources with incorrect negative sentences for students to identify and correct.

Debate Topics: Pre-prepared topics that encourage students to use the negative form of verbs in a real-life context.

By the end of the lesson, students will have mastered the use of the negative form of verbs in the Present Simple tense. They will be able to construct negative sentences confidently using *do not* (*don't*) and *does not* (*doesn't*). Additionally, students will have expanded their vocabulary related to houses and furniture, prepositions of place, and places in a city. Through interactive activities like the True or False game, error correction tasks, and debates, students will have reinforced their grammatical knowledge and vocabulary in engaging, real-life contexts. This comprehensive approach ensures that students gain a strong foundation in both grammar and vocabulary, equipping them to use the Present Simple tense effectively in daily communication.

h. Motivation of choices

This lesson is designed to strengthen students' understanding of grammar while keeping learning engaging and relevant. The focus on the Present Simple tense – negative form addresses an essential aspect of everyday communication, helping learners express disagreement, absence, and non-existence clearly and accurately.

Vocabulary related to houses, furniture, prepositions of place, and places in a city makes the lesson practical and immediately useful. Describing one's surroundings supports meaningful communication and helps students construct more precise sentences.

Interactive tasks such as True or False and sentence transformation activities encourage active participation and immediate application of grammar and vocabulary. A strong emphasis on error correction and self-correction fosters reflection and independence, helping learners develop accuracy and confidence.

A debate in the consolidation phase offers authentic practice of the negative form in real communication, promoting fluency, collaboration, and spontaneous language use.

The lesson follows a scaffolded progression, from clear explanation to guided and autonomous practice, allowing students to build confidence gradually and retain concepts effectively.

By embedding grammar and vocabulary in real-life contexts, such as describing homes, cities, and routines, the lesson enhances motivation, comprehension, and long-term retention.

Overall, this structure combines clarity, interaction, and relevance, helping students develop both competence and confidence in using English effectively.

7. Lesson 7: Subject Pronouns & Object pronouns

a. Main Objectives

The main objective of this lesson is to ensure students understand and correctly use subject and object pronouns in English. In addition, students will expand their vocabulary, focusing on categories such as food (fruits, vegetables, meat, dairy), countable and uncountable nouns, drinks (water, juice, coffee, tea), and holidays (Christmas, Easter, Halloween). These areas are chosen to ensure practical and meaningful learning, helping

students to understand the usage of pronouns in everyday contexts while enriching their vocabulary with terms that are widely applicable in daily conversations.

b. Methodology

This lesson follows an inductive approach, where students encounter subject pronouns in meaningful contexts before identifying their function and use. Rather than starting with isolated grammar rules, the lesson introduces pronouns through dialogues, visuals, and interactive exercises, helping students understand how they naturally appear in conversation and writing. This method promotes active discovery and long-term retention.

The lesson begins with an animated video or interactive presentation that introduces subject pronouns (I, you, he, she, it, we, they) in everyday situations. Students observe how pronouns replace nouns to make speech more natural, for instance, using *he* instead of repeating *John*. This visual and auditory input helps students grasp the concept through clear, real-life examples.

On the online platform, students complete a series of interactive exercises to reinforce their understanding. The first is a matching activity, where they pair pronouns with images of people or objects (e.g., a boy with *he*, a girl with *she*). This strengthens recognition and connection between pronouns and real-world referents.

Next, a sentence completion exercise has students fill in blanks with the correct pronoun, for example, *___ is a teacher* → *She is a teacher*. This reinforces comprehension and sentence-building skills.

Finally, a sentence creation task encourages independent production, where students write their own examples using subject pronouns. They can share their sentences in an online class forum for peer and teacher feedback, promoting collaboration and self-correction. This final activity helps consolidate learning and builds confidence in using pronouns accurately in communication.

c. Content Breakdown

To consolidate their understanding, students will participate in a series of listening exercises. In these exercises, students will listen to short dialogues or descriptions where subject pronouns are used. Afterward, they will be asked to identify the subject pronouns within the dialogue, testing their ability to recognize pronouns in context. This helps transition the focus from written exercises to real-life listening comprehension, reinforcing the importance of subject pronouns in both spoken and written English.

The next step will be a writing activity, where students will be asked to write a short text about themselves or someone they know, using subject pronouns correctly. This task encourages students to personalize the material and apply it in a more creative, real-world context. It also gives them the opportunity to practice using subject pronouns in longer sentences, promoting fluency and coherence in their writing.

Finally, interactive quizzes will be available on the platform, offering multiple-choice or fill-in-the-blank questions to assess the students' understanding of subject pronouns. These quizzes are designed to be engaging and to offer immediate feedback, allowing students to track their progress and revisit areas they find challenging. The quizzes help verify the students' understanding and serve as a review tool to ensure that the lesson's content has been internalized.

d. Resources:

The lesson will utilize a variety of resources to enhance learning. These include the animated video or interactive presentation for the introductory phase, which will be designed to visually and auditorily demonstrate the use of subject pronouns in context. The online platform will offer matching exercises, sentence completion tasks, and sentence creation activities, all designed to provide interactive, scaffolded practice. In the consolidation phase, students will have access to listening exercises, writing prompts, and interactive quizzes to reinforce the lesson's content in both receptive (listening) and productive (writing) tasks. These resources will cater to different learning styles, ensuring that students can engage with the material in multiple ways and solidify their understanding of subject pronouns.

e. Motivation of Choices

The approach taken in this lesson was designed to ensure that students can grasp the function of subject pronouns in a contextually rich and engaging manner. By using real-life examples through dialogues and images, the lesson makes abstract concepts feel more practical and relatable. The gradual progression from presentation to guided practice and independent application allows students to build their understanding step-by-step. Interactive exercises ensure active student engagement and provide opportunities for immediate feedback, reinforcing key concepts as students practice. The choice of online resources such as quizzes and sentence creation tasks ensures that students are consistently challenged to apply their knowledge while receiving timely support, helping to motivate continuous learning and improvement.

8. Lesson 8: Prepositions of Place

a. Main Objectives

The primary objective of this lesson is to ensure students understand and correctly use prepositions of place in English, including terms such as *in*, *on*, *at*, *under*, *next to*, *in front of*, *behind*, and *between*. In addition to grammatical accuracy, the lesson aims to broaden students' vocabulary in related thematic areas such as personality adjectives (e.g., *kind*, *funny*, *smart*, *brave*), weather conditions (e.g., *sunny*, *rainy*, *cloudy*, *windy*), adventure-related terms (e.g., *travel*, *explore*, *discover*), and common adverbs (e.g., *quickly*, *slowly*, *carefully*, *loudly*). The vocabulary themes have been selected to provide meaningful linguistic contexts that support the accurate and natural use of prepositions, while also enriching students' everyday communication skills.

b. Methodology

This lesson uses a visual and kinesthetic approach to help students develop an intuitive understanding of prepositions of place through concrete, hands-on learning.

The introductory phase features an animated video or slideshow illustrating real-world examples such as a book on the table, a cat under the chair, or a child behind the door. These visuals make abstract spatial relationships clear and memorable.

Next, students engage in interactive online exercises that promote active learning. In the first task, a matching activity, students pair images with the correct prepositions (e.g., a lamp on a desk → *on*), reinforcing recognition through visual association.

They then complete multiple-choice exercises, selecting the appropriate preposition to describe positions in given images. This helps strengthen the connection between visual cues and accurate language use.

Finally, in the sentence completion phase, students apply their knowledge in context. Using visuals and partial prompts, they fill in blanks with the correct prepositions (e.g., The dog is ____ the sofa → *under*). This guided practice supports accuracy and prepares them for independent application.

c. Content Breakdown

To consolidate the learning outcomes, the lesson progresses to more productive and integrative tasks. In the image description activity, students are shown various scenes and asked to describe the spatial arrangement of people or objects using prepositions of place. These descriptions can be submitted as written texts or recorded orally, allowing students to develop both their writing and speaking skills while reinforcing correct prepositional usage.

In a creative mapping activity, students are encouraged to produce a map of a familiar space, such as their home, school, or neighborhood, and describe the positions of various objects or people using prepositions of place. This activity integrates creativity with spatial reasoning and language production, offering a

personalized and meaningful application of the lesson content. Maps may be created using digital tools or by hand, with uploads facilitated via the online learning platform.

In the final consolidation phase, students complete listening comprehension exercises. These involve listening to short recordings in which objects or people are described in relation to their surroundings using prepositions. Students must identify which prepositions were used, thereby practicing real-time auditory recognition and reinforcing the connection between spoken input and grammatical structures.

d. Resources

A diverse range of resources supports this lesson. These include the animated video or multimedia presentation used during the introductory phase, which visually demonstrates each preposition's meaning and typical usage. Interactive online activities, including matching tasks, multiple-choice questions, and sentence completion exercises, provide scaffolded, engaging opportunities for grammar practice. Additional resources such as image description prompts, map creation tools, and audio clips for listening comprehension support both receptive and productive language skills. These multimodal resources ensure the lesson caters to a variety of learning styles and preferences, maximizing accessibility and student engagement.

e. Motivation of Choices

The instructional strategies and resources employed in this lesson were selected to support an immersive and context-rich learning experience. Visual and kinesthetic methods were prioritized to make abstract spatial concepts more concrete and memorable, particularly for younger learners or those at early proficiency levels. By encouraging students to interact with visual stimuli, engage in creative mapping, and describe real-life contexts, the lesson promotes a deeper understanding of prepositions of place through experiential learning. The progression from recognition to guided application and finally to creative production ensures a gradual, supported development of competence. The inclusion of immediate feedback mechanisms through quizzes and interactive tools helps to reinforce learning while maintaining student motivation and promoting self-assessment.