DOI: 10.54254/2753-7048/2025.CB25093

The Impact of GPT-based Dialogue Systems vs. Traditional Word-list Memorization on Second Language Vocabulary Acquisition

Yueru Zhao

School of Foreign Studies, Nanjing University, Nanjing, China
221108122@smail.nju.edu.cn

Abstract. In an age where technology is rapidly changing how individuals interact with core issues, such as education, respective institutions need to embrace these trends. Today, learning a second language is characterized by shifts in artificial intelligence, contesting with the existing traditional systems. This study compares GPT-based dialogue systems with traditional word-list practice for second language (L2) English vocabulary acquisition. Based on a 2 × 3 mixed ANOVA and independent-samples t-tests, the study analyzed pre-test, post-test, and delayed post-test data gained from 60 participants (n = 30 each for GPT group and traditional groups). Findings indicate that the GPT group significantly outperform the word-list group with a greater immediate learning gains and retention and robust effect sizes. Analysis was grounded in theories of comprehensible input, declarative/procedural knowledge, and scaffolding, positioning the study as an effective advocate for AI-driven tools in language education. These tools can be integrated in modern L2 approaches, but effective adoption needs to consider their current limitations. The potential of GPT-based techniques reveal that stakeholders need to embrace these systems even as they maintain traditional approaches.

Keywords: GPT-based dialogue systems, L2 vocabulary acquisition, language learning technology, educational technology

1. Introduction

Second language (L2) acquisition is increasing at an unprecedented rate as people seek to adapt to an increasingly globalized world. Data from the European Union revealed a growing trend with 87% of secondary school students learning English as a foreign language in 2022 alone with 49% of this demographic learning two or more languages [1]. These figures reveal the potential of significant learner demographics that could pressure current traditional approaches to L2 vocabulary acquisition. However, attributable to technological advancements, the potential for specialized L2 solutions is more feasible. An example is artificial intelligence (AI) that could be adapted for vocabulary acquisition and bridge gaps posed by traditional approaches. While traditional methods (e.g., word-list memorization) have been widely used for decades, their effectiveness is steadily declining within the long-term retention context. Modern learners are looking for more interactive

DOI: 10.54254/2753-7048/2025.CB25093

and contextualized approaches to language learning. To meet these shifting demands, AI-powered tools, such as GPT-based dialogue systems may present solutions for personalized and interactive learner experiences.

This study seeks to explore the potential of GPT-based learning and whether it is more effective than traditional approaches by answering the question: How effective is L2 English vocabulary learning through GPT-based solutions in comparison to traditional approaches? This study employs a comparative approach that is meant to assess the strengths and weaknesses of the two strategies. To achieve this, the study is guided by an objective that seeks to compare the immediate L2 vocabulary acquisition (recognition and recall) between learners using the GPT-based dialogue system and those using traditional word-list memorization. This study is among the few that provide a comparative analysis of traditional and AI-based methods for L2 vocabulary acquisition. Therefore, findings are likely to significantly influence language education. If GPT-based dialogue system is found to be more effective, it could encourage the adoption of AI-powered tools in language learning curricula. However, if traditional methods are equally or more effective, then it could highlight continued relevance of these approaches in the digital age.

2. Literature review

2.1. Traditional approaches to L2 vocabulary acquisition

L2 acquisition has often relied on conventional approaches that involved accessing physical locations and materials. While modern technologies exist to revolutionize this, these traditional approaches are still popular. Traditional methods rely heavily on incidental learning, where vocabulary is acquired without explicit intention. Instead, they rely on implicit cognitive processes to enhance vocabulary retention. An example is reading comprehension tasks that allow learners to create provisional lexical entries and enhance vocabulary retention [2]. Repetition, another approach to traditional L2 vocabulary acquisition is also considered effective because it is simple and familiar to most learners, thus reinforcing word knowledge. Prior study also suggest that varied exposure to vocabulary in different contexts can create better learning outcomes and active participation.

The potential of traditional learning approaches heightens effective L2 vocabulary acquisition through enhanced deep cognitive processing skills. According to Ender, these strategies are effective for long-term retention and effective use of new vocabulary [2]. Traditional metacognitive approaches provide good conditions for predicting vocabulary knowledge acquisition. These conventional methods face obstacles which reduce their effectiveness when used for long-term language retention [3]. The development of technology-based solutions is proposed to resolve these matters.

Technology eliminates the challenges posed by traditional approaches. For instance, mobile-assisted language learning addresses the accessibility and affordability gaps that make traditional L2 expensive. Chang and Hung found that these technological interventions had a large positive effect on L2 acquisition (mean effect size = .993) [4]. Yu and Trainin also found that these strategies had a moderate effect size (d = 0.64) in technology-assisted vocabulary learning [5]. Specifically, they found that mobile-assisted learning demonstrated greater learning outcomes than computer-assisted instruction. Such insights suggest that technology, when adapted for L2 vocabulary acquisition, can lead to enhanced outcomes for learners.

DOI: 10.54254/2753-7048/2025.CB25093

2.2. Gpt-based dialogue systems on L2 vocabulary acquisition

GPT-based systems stand out as a major technological advancement which proves suitable for vocabulary learning. GPT-based systems use artificial intelligence together with large language models (LLMs) for developing different tools for various tasks. Language learning benefits from GPT-systems which demonstrate effective ability to teach receptive vocabulary and productive vocabulary acquisition [6]. This is specifically true since GPT-based systems foster consumer interaction without intermediaries, thus contributing to long-term retention and incidental vocabulary learning. The language learning systems enhance L2 vocabulary acquisition when combined with conventional educational technologies [7]. From their findings, the authors achieved the development of vocabulary learning assistants by utilizing this combined approach [7]. The assistants would enable learners to submit questions that generate time-sensitive feedback which helps guide their learning process.

The strategic implementation of automation makes GPT-based systems highly suitable for L2 vocabulary acquisition. According to Timpe-Laughlin and Dombi, learners benefit from personalized feedback through this feature, guiding development in particular language skills [8]. Requests made by learners tailor feedback more effectively as it provides them with exact learning-use case feedback. As Hatmanto and Sari puts it, GPT systems provide specific learning activities which align with user proficiency ranges and individual learning speed [9]. Hu and Škultéty demonstrated a particular application of GPT-based dialogue systems that replicate natural human conversations [10]. The computerized learning environment functions optimally because it provides learners with a reassuring environment to practice both speech and listening skills. The strategy provides advantageous conditions for teaching L2 vocabulary acquisition because it promotes target language proficiency through increased learner confidence and better fluency.

The role-playing features of GPT systems have also been explored in literature. In this context, these systems employ strategies that could mimic real-life language usage in specific learning contexts [11]. This approach poses specific value for learners through their instant conversational abilities and strategic competence skills. These models gain additional strength because they could match the education principles that exist in traditional L2 approaches. For instance, Hatmanto and Sari illustrate this quality, revealing that GPT-based systems match the guidelines of Communicative Language Teaching and Task-Based Learning educational approaches [9]. L2 vocabulary acquisition benefits from their application which helps students develop active learning techniques alongside autonomy and authentic usage. Moreover, Timpe-Laughlin et al. supported the potential integration of technology-based systems, revealing that these systems when integrated into classroom settings, can help teachers supplement traditional teaching methods [12]. For instance, GPT-based tasks can be incorporated into flipped classroom models to enhance the overall learning experience.

However, the adoption of GPT-based systems for L2 vocabulary acquisition still lags due to its associated challenges. For instance, these systems are characterized by high word-error-rates, implying that they could produce wrong information to language learners, especially in the context of non-native or non-fluent speech [13]. As the technology is still developing, these systems are currently marred by occasional inaccuracy and inconsistency that is particularly problematic for language learners who rely on precise and correct information to build their vocabulary [14]. Language learning is also a holistic approach that requires effective pedagogy, reliable content, and a supportive teacher-student relationship. According to Nouzri et al., GPT-based systems eliminate this element, thus making it insufficient for comprehensive language learning [15]. This observation demonstrates that despite their novelty, GPT-based systems may not have a significant improvement effect on language learning.

DOI: 10.54254/2753-7048/2025.CB25093

2.3. Theoretical framework

This study is informed by a structured theoretical framework that considers perspectives from three foundations: (i) second language acquisition (SLA); (ii) cognitive psychology; and (iii) educational technology. Under SLA, the study relies on the Input Hypothesis [16]. According to this theory, it's critical to strike a balance between known and unknown knowledge to promote the easy assimilation of fresh perspectives [16]. This hypothesis is directly applicable to the current study, specifically since it considers the immediate learning outcome. For instance, learners acquire language most efficiently when exposed to input that is slightly beyond their current level but still understandable. Traditional word lists provide input but often lack the context that limits comprehension. However, a GPT-based dialogue system can deliver contextualized input through simulated conversations.

The study also relies on Anderson's declarative vs. procedural knowledge [17]. The process of acquiring a language is naturally multifaceted, requiring learners to master both general factual knowledge and task-specific procedural skills. Declarative knowledge involves knowing "that" (e.g., word meanings), while procedural knowledge is knowing "how" (e.g., using words in context). Through sustained practice and effort, declarative and procedural knowledge could be enhanced in memory with learner performance becoming more reliable and rapid [17]. In the context of this study, declarative knowledge is illustrated in traditional approaches (word-list memorization) while the dialogue system fosters both as it requires contextual application. Since L2 vocabulary acquisition is educational in nature, this study also considers Vygotsky's concept of Zone of Proximal Development [18]. This construct essentially describes the gap between what learners can do and what they can accomplish with the guidance of more experienced parties [18]. In most L2 contexts, learning is optimized when the learner has some level of current knowledge that is later supported through scaffolding. A GPT-based system can adapt to the level of a learner and provide them with personalized scaffolding as opposed to the static nature of word lists.

3. Method

3.1. Design

The study employed a randomized, two-group, repeated-measures design with three measurement occasions, including pre-test, immediate post-test, delayed post-test (1 week later). The independent variable was learning condition (GPT-based chatbot vs. traditional word-list/flashcard practice) while the dependent variable explored vocabulary knowledge indexed by total percentage scores on a 40-item test (20 multiple-choice recognition items and 20 cued-recall items). This design facilitated examination of three constructs, including (i) overall learning gains, (ii) differential improvement between conditions, and (iii) retention after a two-week delay.

3.2. Participants

Sixty intermediate L2 English learners from aged at least 18 years old were recruited from university language-support programs and represented diverse L1 backgrounds (East Asian = 31.7%, South Asian = 25%, Southeast Asian = 20%, European non-English = 8.3%, Arabic = 6.7%, Latin American = 3.3%, and Other = 5%). Random assignment was conducted using stratified sampling by proficiency level, producing 30 learners per condition.

DOI: 10.54254/2753-7048/2025.CB25093

3.3. Materials and measures

The experimental tool was a custom interface to OpenAI's GPT model with prompts designed to guide it to present individualized definitions, example sentences, and follow-up elaborations for each of the 20 target words. The same 20 target words with dictionary definitions and L2 example sentences were used for the control units. The Vocabulary Test (Forms A, B, C) contained identical target items. Scores were computed as the sum of correct recognition and recall items (maximum = 40, reported as percentage). The forms also gathered demographic questions, focusing on the age, L1 category, years of formal English study, weekly L2 exposure, technology-use frequency, and vocabulary-learning motivation. The post-study instrument contained five global usability items plus condition-specific items (four for GPT, three for flashcards) and single-choice questions on boredom, repetition, and technical difficulties (see Appendix 3).

4. Results & discussion

A 2 × 3 mixed ANOVA was conducted with Group (GPT vs. Word List) as the between-subjects factor and Time (Pre-test, Post-test, Delayed post-test) as the within-subjects factor. Table 1 illustrates an overview of these groups.

Table 1: Descriptive statistics

	Group	Mean	SD
Pre-Test Total(%)	Experimental (GPT)	12.20	4.715
	Control (Word List)	11.57	4.321
	Total	11.88	4.495
Post-Test Total(%)	Experimental (GPT)	30.33	5.492
	Control (Word List)	23.57	5.905
	Total	26.95	6.604
Delayed Post-Test Total(%)	Experimental (GPT)	25.67	5.006
	Control (Word List)	19.23	5.151
	Total	22.45	5.990

As Table 1 shows, the pre-test scores were comparable between the GPT group and the Word List group. However, the GPT group showed substantially higher scores than the Word List group at both the immediate post-test and the delayed post-test. The overall scores increased from pre-test to post-test and then decreased slightly at delayed post-test, while remaining significantly higher than pre-test levels. Further examination of the Group × Time interaction revealed that while both groups improved from pre-test to post-test (see Table 1), the GPT group showed significantly greater gains ($\Delta = 18.13$) compared to the Word List group ($\Delta = 12.00$). Similarly, retention at delayed post-test was superior for the GPT group, with scores 13.47 points higher than pre-test, compared to a 7.67-point improvement for the Word List group.

Mauchly's test indicated that the assumption of sphericity was violated ($W = .277$, $\chi2(2) = 73.247$, $p < .001$), so the Greenhouse-Geisser correction was applied ($\varepsilon = .580$). The analysis revealed a significant main effect of Time, $F(1.161, 67.310) = 532.639$, $p < .001$, partial $\eta2 = .902$, indicating that vocabulary scores changed significantly across the three time points for all participants (see Table 2).

DOI: 10.54254/2753-7048/2025.CB25093

Table 2: Test of within-subject effects between the two groups

Measure: MEASURE_1							
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Time	Sphericity Assumed	7178.178	2	3589.089	532.639	<.001	.902
	Greenhouse-Geisser	7178.178	1.161	6185.281	532.639	<.001	.902
	Huynh-Feldt	7178.178	1.190	6032.745	532.639	<.001	.902
	Lower-bound	7178.178	1.000	7178.178	532.639	<.001	.902
Time * Group	Sphericity Assumed	356.844	2	178.422	26.479	<.001	.313
	Greenhouse-Geisser	356.844	1.161	307.485	26.479	<.001	.313
	Huynh-Feldt	356.844	1.190	299.902	26.479	<.001	.313
	Lower-bound	356.844	1.000	356.844	26.479	<.001	.313
Error(Time)	Sphericity Assumed	781.644	116	6.738			
	Greenhouse-Geisser	781.644	67.310	11.613			
	Huynh-Feldt	781.644	69.012	11.326			
	Lower-bound	781.644	58.000	13.477			

The main effect of Group was also significant, $F(1, 58) = 14.656$, $p < .001$, partial $\eta^2 = .202$, demonstrating that the GPT group showed overall better performance than the Word List group as Table 3 shows.

Table 3: Tests of between-subjects effects

Measure: MEASURE_1						
Transformed Variable: Average						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	75112.939	1	75112.939	1150.521	<.001	.952
Group	956.806	1	956.806	14.656	<.001	.202
Error	3786.589	58	65.286			

Most importantly, there was a significant Time × Group interaction, $F(1.161, 67.310) = 26.479$, $p < .001$, partial $\eta^2 = .313$. This interaction indicates that the pattern of change in vocabulary scores over time differed between the two instructional methods. Pairwise comparisons with Bonferroni

DOI: 10.54254/2753-7048/2025.CB25093

correction showed that all three time points differed significantly from each other ($p < .001$ for all comparisons) as Table 4 shows.

Table 4: Pairwise comparisons based on Bonferroni correction

Measure: MEASURE_1						
(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
1	2	-15.067*	.588	<.001	-16.515	-13.618
	3	-10.567*	.541	<.001	-11.900	-9.233
2	1	15.067*	.588	<.001	13.618	16.515
	3	4.500*	.190	<.001	4.032	4.968
3	1	10.567*	.541	<.001	9.233	11.900
	2	-4.500*	.190	<.001	-4.968	-4.032

From these results, it is evident that GPT-based solutions are significantly more effective than traditional word-list practice for L2 English vocabulary learning. This is supported by evidence from the immediate learning gains and retention over time. The advantage of the GPT-based system over traditional approaches suggest that as learners receive intelligible input that is just a little bit above their current competency level, language acquisition is most successful [16]. The current findings suggest that learners who encountered this input was presented to them through contextualized and adaptive interactions, which is different from the static and decontextualized nature of word lists. This is complemented by literature that found GPT systems effective in simulating natural conversations and deliver personalized feedback [8,10].

An independent-samples t-test was also conducted to compare vocabulary learning gains between the two instructional methods. Analysis examined immediate learning gains and retention over time as illustrated in Table 5.

Table 5: Group statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Gain_Immediate	Experimental (GPT)	30	18.1333	4.74693	.86667
	Control (Word List)	30	12.0000	4.34702	.79365
Gain_Retention	Experimental (GPT)	30	13.4667	4.48548	.81893
	Control (Word List)	30	7.6667	3.87150	.70684

DOI: 10.54254/2753-7048/2025.CB25093

Table 6: Independent samples t-test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						One-Sided p	Two-Sided p			Lower	Upper
Gain_Immediate	Equal variances assumed	.05	.82	5.22	58	<.001	<.001	6.13	1.18	3.78	8.49
	Equal variances not assumed			5.22	57.56	<.001	<.001	6.13	1.18	3.78	8.49
Gain_Retention	Equal variances assumed	.57	.46	5.36	58	<.001	<.001	5.80	1.08	3.63	7.97
	Equal variances not assumed			5.36	56.79	<.001	<.001	5.80	1.08	3.63	7.97

Table 7: Independent samples effect sizes

		Standardizer	Point Estimate	95% Confidence Interval	
				Lower	Upper
Gain_Immediate	Cohen's d	4.55137	1.348	.781	1.905
	Hedges' correction	4.61130	1.330	.771	1.880
	Glass's delta	4.34702	1.411	.781	2.025
Gain_Retention	Cohen's d	4.18975	1.384	.814	1.945
	Hedges' correction	4.24492	1.366	.804	1.919
	Glass's delta	3.87150	1.498	.855	2.125

The Levene's test indicated that equal variances could be assumed for immediate learning gains, ($F = .053$, $p = .818$). The experimental group using GPT-based solutions ($M = 18.1333$, $SD = 4.7469$) demonstrated significantly higher immediate vocabulary gains than the control group using traditional word-list practice ($M = 12.0000$, $SD = 4.3470$), $t(58) = 5.2190$, $p < .001$, 95% CI [3.781, 8.486] as Table 6 shows. There was a larger effect size as indicated by the Cohen's $d = 1.348$ (see Table 7). This effect size is significantly larger than what has been established in previous literature, further implying that GPT systems are a particularly effective tool. For retention scores, Levene's test similarly indicated that equal variances could be assumed ($F = .566$, $p = .455$) as illustrated in Table 6. The experimental group ($M = 13.4667$, $SD = 4.4855$) demonstrated significantly greater retention of vocabulary knowledge compared to the control group ($M = 7.6667$, $SD = 3.8715$), $t(58) = 5.361$, $p < .001$, 95% CI [3.6356, 7.9654].

DOI: 10.54254/2753-7048/2025.CB25093

Independent samples t-test also implies that GPT-based solutions provide significantly greater benefits for L2 English vocabulary learning compared to traditional word-list practice. This is particularly demonstrated in the immediate learning outcomes and longer-term retention where the experimental group outperforms the control group by 51.11% on immediate gains and 75.65% on retention measures. These figures illustrate affective declarative that requires learners to apply vocabulary in dialogue [17]. In context, this a strategy that promotes productive vocabulary acquisition through active usage, explaining the observed 75.65% higher retention scores in the GPT group [6]. The results on immediate gains have been complemented by research that demonstrated the effectiveness of GPT-based systems in adjusting feedback and activities to learner proficiency and pace, thus accelerating vocabulary acquisition [9]. This is why the GPT group achieved a 51.11% higher immediate gain over the Word List group since it dynamically bridged the gap between current and potential abilities. Traditional approaches often lack the interactivity and contextual depth of GPT systems [3]. The current findings corroborate this as Word List group’s gains, though significant, were markedly lower than those of the GPT group, suggesting that rote memorization is less effective than interactive, technology-mediated learning.

5. Conclusion

The aim of this study was to explore the potential of GPT-based L2 vocabulary acquisition approaches over tradition strategies. Findings reveal that GPT-based systems are indeed superior over traditional word-list practice for L2 English vocabulary acquisition. Participants using the GPT system demonstrated significantly higher immediate learning gains and better retention, as confirmed by the comprehensive analyses. These results complement initial insights from literature and theoretical frameworks that explain the strengths of GPT systems. These results have vital implications for L2 acquisition, considering that technological adoption is being embraced in this field. This means that when integrated effectively in language education, AI-based technologies could enhance vocabulary learning efficiency. These systems eliminate the barriers posed by traditional systems, suggesting that are suited for effective learner-paced language acquisition. However, implementation and adoption need to consider that this trend is limited by inaccuracies that could impede effectiveness. The study also follows a narrow approach in the vocabulary domain, which is not the whole breadth of language learning. The sample size is also smaller, limiting generalizability to a wider demographic of L2 students. Furthermore, individual differences in prior familiarity with GPT technology and a potential novelty effect were not controlled, possibly influencing the observed learning outcomes. The short intervention period and the specificity of the tools and vocabulary also mean results primarily reflect short-term effects within this particular context. Future research should explore broader applications (e.g., pronunciation and cultural competence) to fully assess the potential GPT-based systems have in language learning. Despite these slight limitations, this study positions GPT-based systems as a transformative tool that could improve outcomes in L2 acquisition.

References

[1] Eurostat. (2024, September). Foreign language learning statistics. https: //ec.europa.eu/eurostat/statistics-explained/index.php?title=Foreign_language_learning_statistics

[2] Ender, A. (2016). Implicit and Explicit Cognitive Processes in Incidental Vocabulary Acquisition. Applied Linguistics, 37(4), 536–560.

[3] Kashani, S., & Shafiee, S. (2016). A comparison of vocabulary learning strategies among elementary Iranian EFL learners. Journal of Language Teaching and Research, 7(3), 511–518.

[4] Chang, M.-M., & Hung, H.-T. (2019). Effects of technology-enhanced language learning on second language acquisition: A meta-analysis. Educational Technology and Society, 22(4), 1–17.
[5] Yu, A., & Trainin, G. (2022). A meta-analysis examining technology-assisted L2 vocabulary learning. ReCALL, 34(2), 235–252.
[6] Zhang, Z., & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. Heliyon, 10(3), e25370.
[7] Zografos, G., & Moussiades, L. (2023). A GPT-Based Vocabulary Tutor. Augmented Intelligence and Intelligent Tutoring Systems. ITS 2023. Lecture Notes in Computer Science (Vol. 13891, pp. 270–280). Springer, Cham.
[8] Timpe-Laughlin, V., & Dombi, J. (2020). Exploring L2 learners' request behavior in a multi-turn conversation with a fully automated agent. Intercultural Pragmatics, 17(2), 221–257.
[9] Hatmanto, E. D., & Sari, M. I. (2023). Aligning Theory and Practice: Leveraging Chat GPT for Effective English Language Teaching and Learning. International Conference on Environment and Smart Society (ICEnSO 2023) (Vol. 440). E3S Web of Conferences.
[10] Hu, W.-C., & Škultéty, R. (2024). Unlocking the Learning Potential: ChatGPT as a Virtual Platform for Cross-Interaction in English Language Learning. Engineering Proceedings, 74(1), 59.
[11] Wang, C., & Li, X. (2025). Software Review: Empowering Language Education With D-ID Creative Reality Studio's Multimodal Capabilities. International Journal of Computer-Assisted Language Learning and Teaching, 15(1). 1-11.
[12] Timpe-Laughlin, V., Sydorenko, T., & Daurio, P. (2022). Using spoken dialogue technology for L2 speaking practice: What do teachers think? Computer Assisted Language Learning, 35(5–6), 1194–1217.
[13] Mai, L., & Carson-Berndsen, J. (2024). Enhancing conversation smoothness in language learning chatbots: An evaluation of GPT4 for ASR error correction. ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 11001-11005). IEEE.
[14] Alhazmi, A. A., & Muftah, M. (2025). Evaluating ChatGPT's Reliability in Second Language Acquisition (SLA): Insights on Language Skills and Technology's Role. Electronic Journal of E-Learning, 23(1), 96–112.
[15] Nouzri, S., EL Fatimi, M., Guerin, T., Othmane, M., & Najjar, A. (2025). Beyond Chatbots: Enhancing Luxembourgish Language Learning Through Multi-Agent Systems and Large Language Model. PRIMA 2024: Principles and Practice of Multi-Agent Systems. PRIMA 2024. Lecture Notes in Computer Science (Vol. 15395, pp. 385–401). Springer, Cham.
[16] Krashen, S. (1985). The input hypothesis: Issues and implications. Longman.
[17] Anderson, J. (1993). Rules of the Mind. Lawrence Erlbaum.
[18] Vygotsky, L. S. (1978). Mind in society. Harvard University Press.