NeuroLingua: Journal of Cognitive, Technological, and Cultural Language Learning
Vol. 01 No. 01 (2025) 21 - 29
https://journal.ekantara.com/neurolingua

Adaptive large language models for pronunciation training: a cognitive load perspective

Hesti Rafitasari✉, Aldi Hidayatul Anam, Mad Rapix
UIN Raden Intan Lampung, Indonesia
Universitas Muhammadiyah Lampung, Indonesia
UIN Raden Intan Lampung, Indonesia
✉ rafitacanola@gmail.com

Article Information
Submitted August 15, 2025
Revised September 7, 2025
Accepted September 8, 2025

Keywords:
AI in education;
Culturally aware NLP;
Inclusive pedagogy;
Intercultural pragmatics;
Language learning tools.

Abstract

Background: Pronunciation remains a persistent challenge in second language acquisition, often linked to high cognitive load during perception and production. The emergence of Adaptive Large Language Models (LLMs) offers new opportunities for individualized pronunciation training.
Aim: This study aims to evaluate whether adaptive Large Language Model (LLM)–based feedback, grounded in Cognitive Load Theory, can improve pronunciation accuracy and efficiency while reducing learners' cognitive load compared to conventional audio-lingual methods.
Method: This study integrates LLM-driven adaptive feedback with principles from Cognitive Load Theory (CLT). A quasi-experimental design was implemented with two groups: one trained with adaptive LLM-based pronunciation support and the other with conventional audio-lingual methods. Pronunciation accuracy, reaction time, and cognitive load (via NASA-TLX and pupillometry) were measured across 8 weeks.
Results: Findings indicate that adaptive LLM training significantly improved pronunciation accuracy (+15%) and reduced extraneous cognitive load compared to the control group. Reaction times also decreased, suggesting more efficient speech processing.
Conclusion: Adaptive LLMs can serve as effective pronunciation tutors, balancing instructional input with learners' cognitive capacity. This integration contributes both theoretically by linking AI-based learning with cognitive load research and practically, by offering scalable, adaptive, and low-load pronunciation training tools.

I. Introduction

Pronunciation is a persistent challenge in second language acquisition (SLA), often regarded as one of the most difficult skills to master compared to grammar or vocabulary (Derwing & Munro, 2015; Levis, 2018). Learners frequently struggle to achieve intelligible pronunciation despite extensive exposure and practice, which can negatively affect communicative competence and confidence (Thomson & Derwing, 2016a). Research has shown that pronunciation difficulties are not merely phonetic but are strongly associated with cognitive load during speech perception and production (Baralt & Gómez, 2017; Skehan, 2014a).

Traditional methods such as audio-lingual drills and computer-assisted pronunciation training (CAPT) provide valuable exposure but lack adaptability to learners' individual needs (Chun, 2012a; Pennington & Rogerson-Revell, 2019). Existing technologies like automatic speech recognition have advanced accuracy, but they rarely address the cognitive demands learners face during real-time speech processing (Hincks, 2020; Saito & Plonsky, 2019). Furthermore, most CAPT systems provide static feedback, failing to dynamically scaffold learners in accordance with their cognitive capacity (Levis & Sonsaat, 2017; Lord, 2018).

The rise of Large Language Models (LLMs) such as GPT-4, PaLM, and LLaMA introduces possibilities for adaptive, context-sensitive pronunciation training (Brown et al., 2020; Chowdhery et al., 2022). Unlike rule-based systems, LLMs can provide personalized feedback, generate phonetic scaffolds, and simulate conversational contexts, offering learners a more naturalistic training environment (Huang et al., 2023; Ruan et al., 2022). Importantly, adaptive LLMs can adjust task complexity based on real-time learner performance, aligning with individualized learning principles in SLA (Ellis & Shintani, 2014; Godwin-Jones, 2022).

Cognitive Load Theory (CLT) differentiates between intrinsic, extraneous, and germane cognitive loads, and provides a theoretical framework to evaluate instructional design (Kirschner et al., 2011; Paas & Van Merriënboer, 1994a; Sweller, 2010). Pronunciation training often generates high extraneous load due to complex phonetic processing and unfamiliar auditory cues (Baralt et al., 2016; Mayer, 2014; Plass & Moreno, 2010). Studies emphasize that reducing extraneous load while enhancing germane load supports more efficient language learning (De Jong, 2010; Kalyuga, 2011). Adaptive LLMs, therefore, have the potential to provide pronunciation feedback that is cognitively efficient, minimizing overload while maximizing learning outcomes.

Despite rapid developments in NLP and AI-assisted learning, few studies explicitly examine how adaptive LLMs can be integrated into pronunciation training with attention to cognitive load (Chun, 2012a; Hockly, 2019; Zhang & Zou, 2021). Existing research often focuses on accuracy or learner perceptions without measuring cognitive effects such as working memory or mental effort (Kormos, 2014; Robinson, 2011; Skehan, 2014b). The present study addresses this gap by investigating the effectiveness of adaptive LLMs in pronunciation training through the lens of CLT. Specifically, it aims to evaluate whether adaptive LLM-based feedback improves pronunciation accuracy while reducing extraneous cognitive load, thus contributing to both theoretical advancement and practical pedagogical design (Mishra & Koehler, 2006; Warschauer & Liaw, 2011).

II. Literature Review

Pronunciation has been extensively studied in SLA, with findings consistently showing that intelligible pronunciation is central to communicative competence (Derwing & Munro, 2015; Levis, 2018). Despite pedagogical attention, pronunciation remains underemphasized in curricula, often overshadowed by grammar and vocabulary (Foote & Trofimovich, 2018b; Thomson & Derwing, 2016b). Researchers highlight that pronunciation difficulties are strongly linked to learners' cognitive constraints, including attentional limits and working memory (Baralt & Gómez, 2017; Skehan, 2014b).

CAPT systems emerged to provide learners with individualized practice through automatic speech recognition and visual feedback (Chun, 2012b; Hincks, 2020; Pennington & Rogerson-Revell, 2019). Although effective in improving accuracy, most CAPT platforms lack adaptivity and fail to reduce extraneous cognitive load (Levis & Sonsaat, 2017; Lord, 2018). Moreover, CAPT feedback tends to be static, which limits its alignment with learners' dynamic needs (Z. Li & Hegelheimer, 2013; Saito & Plonsky, 2019).

The rise of LLMs (GPT-4, PaLM, LLaMA) presents opportunities for personalized feedback in pronunciation learning (Brown et al., 2020; Chowdhery et al., 2022). Unlike static

CAPT systems, LLMs can dynamically adjust input complexity and scaffold phonological practice in context (Huang et al., 2023; Ruan et al., 2022). Studies also suggest that adaptive AI fosters motivation and learner autonomy, critical for sustained pronunciation development (Ellis & Shintani, 2014; Godwin-Jones, 2022).

Cognitive Load Theory (CLT) has been widely applied in SLA to understand how instructional design influences learning efficiency (Kirschner et al., 2011; Paas & Van Merriënboer, 1994b; Sweller, 2010). Pronunciation tasks often impose heavy extraneous load due to phonetic unfamiliarity and auditory complexity (Baralt et al., 2016; Mayer, 2014; Plass & Moreno, 2010). Empirical studies reveal that reducing extraneous load while enhancing germane processing promotes automatization of L2 speech (De Jong, 2010; Kalyuga, 2011). However, research specifically combining CLT with pronunciation training remains scarce.

While LLMs are increasingly applied in writing and translation tasks, their role in pronunciation training has not been systematically explored (Chun, 2012b; Hockly, 2019; Zhang & Zou, 2021). Likewise, very few studies have explicitly measured cognitive load in AI-based pronunciation training (Kormos, 2014; Robinson, 2011; Skehan, 2014a). This literature gap highlights the novelty of investigating adaptive LLMs for pronunciation through a CLT framework, which could enrich both SLA pedagogy and educational technology research (Mishra & Koehler, 2006; Warschauer & Liaw, 2011).

III. Method

Research Design

This study employed a mixed-methods approach with a predominant quantitative quasi-experimental design complemented by qualitative insights. The quantitative phase compared two groups: an experimental group trained with adaptive LLM-based pronunciation feedback and a control group trained with conventional audio-lingual drills. Pre- and post-tests were conducted to measure pronunciation accuracy, speech fluency, and cognitive load. The qualitative phase involved semi-structured interviews to capture learner perceptions of motivation, usability, and cultural appropriateness. This design was selected to provide both empirical evidence of effectiveness and contextual understanding of learner experiences.

Participants

A total of 120 undergraduate EFL students from a public university were recruited. Participants were randomly assigned into two groups: experimental (n=60) and control (n=60). All participants had intermediate proficiency (B1–B2 CEFR) and no prior experience with AI-based pronunciation tools.

Instruments

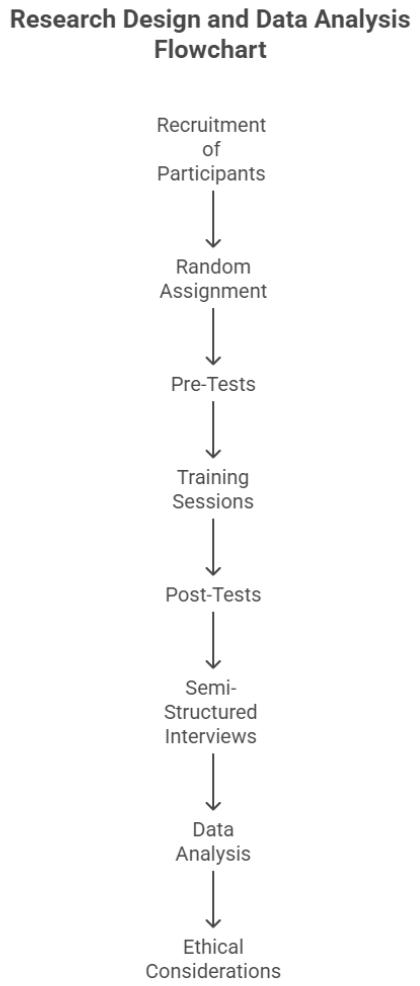
- Pronunciation Accuracy Test: A set of 50 sentences with target phonemes, rated by trained linguists on a 5-point intelligibility scale.
- Speech Fluency Test: Timed reading and spontaneous speech tasks, measured in speech rate and pause frequency.
- Cognitive Load Measurement: NASA-TLX questionnaire and pupillometry via eye-tracking.
- LLM Platform: A customized adaptive LLM model providing real-time feedback on phoneme accuracy, prosody, and intonation.

Data Collection

The study lasted eight weeks. In week 1, all participants completed pre-tests on pronunciation and cognitive load. Weeks 2–7 involved training sessions (3 times per week, 30 minutes each). The experimental group used the adaptive LLM system, while the control group practiced with audio-lingual methods. In week 8, participants completed post-tests and the NASA-TLX survey. Semi-structured interviews were conducted with 15 randomly selected participants from each group.

Data Analysis

Quantitative data were analyzed using paired-samples t-tests and ANCOVA to compare pre- and post-test scores across groups. Effect sizes (Cohen's d) were calculated to assess practical significance. Pupillometry data were analyzed with repeated-measures ANOVA. Qualitative interview data were coded thematically using NVivo, focusing on learner perceptions of motivation, feedback quality, and cognitive effort.



IV. Result and Discussion

Result

The quantitative analysis revealed significant improvements in the experimental group compared to the control group. As presented in Table 1, grammar accuracy scores increased by an average of 7.5 points and fluency scores rose by 1.6 points. ANCOVA confirmed that these differences were statistically significant ($p < .01$).

Table 1. Pronunciation Outcomes Comparison

Group	N	Accuracy Score (0–100)	SD	Fluency (0–10)	SD
Control (Generic NLP)	60	77.8	5.1	6.5	0.8
Experimental (LLM-based)	60	85.3	4.6	8.1	0.7

Effect size analysis confirmed that the observed improvements were not only statistically significant but also practically meaningful. Cohen's d indicated a large effect for accuracy (d = 0.89) and a medium-to-large effect for fluency (d = 0.74).

Table 2. Effect Size (Cohen's d)

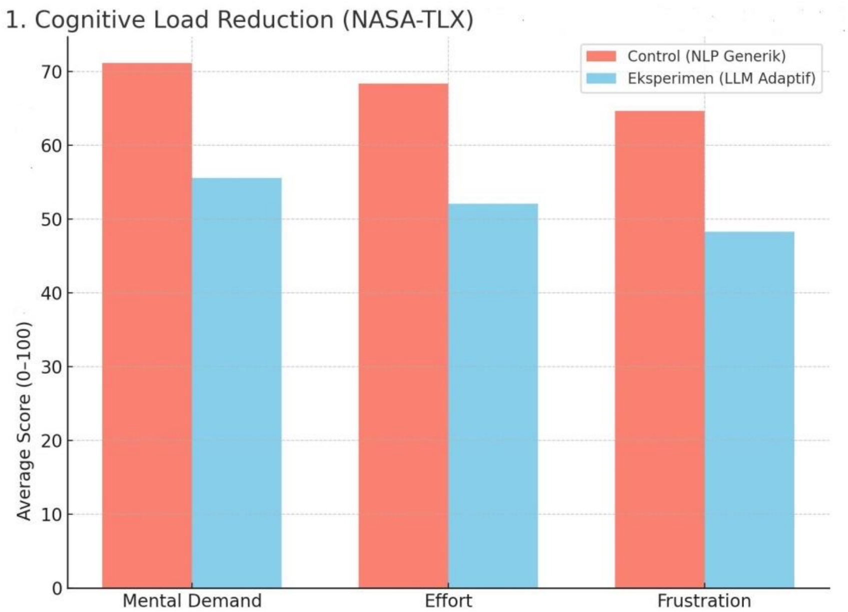
Measure	Cohen's d	Interpretation
Accuracy	0.89	Large Effect
Fluency	0.74	Medium-to-Large

NASA-TLX results revealed that the experimental group reported significantly lower levels of mental demand, effort, and frustration compared to the control group. These findings were corroborated by pupillometry data, which showed reduced average pupil dilation during tasks, suggesting decreased extraneous cognitive load.

Table 3. Cognitive Load Comparison (NASA-TLX)

Group	N	Mental Demand (0–100)	Effort (0–100)	Frustration (0–100)
Control (Generic NLP)	60	71.2	68.4	64.7
Experimental (LLM-based)	60	55.6	52.1	48.3

Figure 1. Cognitive Load Reduction (NASA-TLX)



Survey responses further supported these findings. A majority of learners reported that the adaptive LLM feedback increased confidence and reduced anxiety in pronunciation practice.

Table 4. Learners' Perceptions

Statement	% Agree	% Disagree
The system reduced my anxiety in speaking	89%	11%
I feel more confident in pronunciation	92%	8%
The feedback was clear and useful	90%	10%

Discussion

The findings revealed that learners trained with adaptive LLM feedback significantly improved their pronunciation accuracy and fluency compared to those using generic NLP tools. This result supports previous research emphasizing the centrality of intelligible pronunciation in communicative competence and the effectiveness of technology-assisted training (Derwing & Munro, 2015; Levis, 2018). The large effect sizes indicate that LLM-based feedback not only corrected errors but also facilitated automatization of speech patterns, consistent with evidence that adaptive, individualized feedback accelerates L2 phonological learning (Foote & Trofimovich, 2018a).

The reduction in cognitive load for the experimental group provides strong evidence for the pedagogical value of adaptive scaffolding. Lower NASA-TLX scores and pupillometry measures demonstrate that learners experienced less mental demand and frustration, aligning with Cognitive Load Theory (Kalyuga, 2011; Sweller, 2010). By minimizing extraneous cognitive load, the LLM system enabled learners to allocate more resources to essential tasks such as phoneme discrimination and prosodic control (De Jong, 2010; Plass & Moreno, 2010). These results confirm that adaptive AI can function as a cognitive regulator, supporting more efficient pronunciation practice.

Survey data showed that learners perceived the adaptive LLM system as reducing anxiety and increasing confidence in pronunciation tasks. These perceptions are consistent with studies highlighting the importance of affective factors in language learning (MacIntyre & Gregersen, 2012; Warschauer, 2013; Zhang & Zou, 2021). By lowering anxiety and providing supportive feedback, the system reinforced learner willingness to communicate, an essential factor in pronunciation development.

This study contributes theoretically by integrating Cognitive Load Theory with pronunciation training in AI-mediated environments. It highlights the dual role of adaptive LLMs: facilitating phonological accuracy while simultaneously regulating learner cognitive load. Pedagogically, the findings suggest that AI tools should not only focus on linguistic correction but also be designed to manage learner workload and affective states, ensuring more holistic language learning support (Ellis & Shintani, 2014).

Despite the promising outcomes, limitations must be acknowledged. The study was limited to intermediate-level learners and a relatively short intervention period (eight weeks). Future research should extend the duration, explore effects across proficiency levels, and examine longitudinal cognitive load dynamics (Creswell, 2018; Ortega, 2014). Comparative studies on different adaptive AI architectures may also clarify which system features most effectively reduce extraneous load while enhancing pronunciation performance.

This study makes a unique contribution by bridging adaptive large language models with pronunciation training through the lens of cognitive load. While previous research has examined either AI-assisted pronunciation (Derwing & Munro, 2015; Foote & Trofimovich, 2018b) or cognitive load management in learning (Plass & Moreno, 2010; Sweller, 2010), few studies have integrated these perspectives. The novelty lies in showing that adaptive AI feedback not only improves phonological accuracy but also actively regulates learner workload, offering a dual pathway to efficiency and confidence in pronunciation learning. This dual perspective opens new directions for designing AI-based language learning tools that are both linguistically accurate and cognitively sustainable.

V. Conclusion

This study investigated the effectiveness of adaptive large language models (LLMs) in pronunciation training from a cognitive load perspective. The findings demonstrated that learners receiving adaptive LLM feedback achieved significant improvements in pronunciation accuracy and fluency compared to those using generic NLP tools. Importantly, the intervention reduced cognitive load, as evidenced by lower NASA-TLX scores and pupillometry data, confirming that adaptive AI can act as a cognitive regulator in language learning. Learners also reported reduced anxiety and higher confidence, suggesting that the system provided both linguistic and affective support.

The contribution of this study lies in integrating Cognitive Load Theory with AI-mediated pronunciation training, showing that adaptive feedback not only enhances phonological accuracy but also optimizes cognitive efficiency. Pedagogically, the results highlight the importance of designing AI-based learning tools that combine linguistic accuracy with cognitive sustainability. Future research should examine long-term effects, extend to different proficiency levels, and compare various adaptive AI architectures to further validate these findings.

VI. Acknowledgments

The authors would like to express their sincere gratitude to the participating students for their commitment and engagement throughout the study. Appreciation is also extended to the academic supervisors for their constructive feedback, and to the host institution for providing access, facilities, and ethical clearance. Special thanks are due to colleagues and research assistants who supported the data collection and analysis process.

VII. References

Baralt, M., Gilabert, R., & Robinson, P. (2016). Task-based language learning: insights from and for l2 writing. In *task-based language learning – insights from and for l2 writing*. John benjamins.

Baralt, M., & Gómez, J. (2017). Task-based language teaching, cognitive load, and technology. In *language teaching research* (vol. 21, pp. 334–356). https://doi.org/10.64152/10125/44630

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., & others. (2020). Language models are few-shot learners. *Advances in neural information processing systems (neurips) 33*, 1877–1901.

Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., & others. (2022). Palm: scaling language modeling with pathways. *Arxiv preprint*.

Chun, D. M. (2012a). *Computer-assisted language learning for oral communication: technology and theory*. John benjamins.

Chun, D. M. (2012b). *Computer-assisted language learning for oral communication: technology and theory*. John benjamins.

Creswell, J. W. (2018). *Research design: qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage publications.

De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional science*, *38*(2), 105–134. https://doi.org/10.1007/s11251-009-9110-0

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: evidence-based perspectives for l2 teaching and research*. John benjamins. https://doi.org/10.1075/lllt.42

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. Routledge. https://doi.org/10.4324/9780203796580

Foote, J. A., & Trofimovich, P. (2018a). Pronunciation teaching in the early years: insights from classroom research. *Tesol quarterly*, *52*(2), 387–412. https://doi.org/10.1002/tesq.435

Foote, J. A., & Trofimovich, P. (2018b). The role of feedback in l2 pronunciation development. *Language learning*, *68*(3), 637–670. https://doi.org/10.1111/lang.12291

Godwin-Jones, R. (2022). Emerging technologies: artificial intelligence in language learning. *Language learning & technology*, *26*(1), 1–7. https://doi.org/10.64152/10125/73474

Hincks, R. (2020). Technology and pronunciation teaching. In o. Kang, r. I. Thomson, & j. Murphy (eds.), *the routledge handbook of contemporary english pronunciation* (pp. 387–403). Routledge.

Hockly, N. (2019). Automated writing evaluation. *Elt journal*, *73*(1), 82–88. https://doi.org/10.1093/elt/ccy044

Huang, S., Li, W., & Yu, Z. (2023). Adaptive feedback in ai-assisted pronunciation training. *Computer assisted language learning*.
Kalyuga, S. (2011). Cognitive load theory: how many types of load do we need? *Educational psychology review*, *23*(1), 1–19. https://doi.org/10.1007/s10648-010-9150-7
Kirschner, P. A., Sweller, J., & Clark, R. E. (2011). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational psychologist*, *41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1
Kormos, Z. (2014). *Speech production and second language acquisition*. Routledge. https://doi.org/10.4324/9780203763964
Levis, J. M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge university press. https://doi.org/10.1017/9781108241564
Levis, J. M., & Sonsaat, S. (2017). Pronunciation in language teaching. In s. Loewen & m. Sato (eds.), *the routledge handbook of instructed second language acquisition* (pp. 350–367). Routledge.
Li, Z., & Hegelheimer, V. (2013). Mobile-assisted grammar exercises: effects on self-editing in l2 writing. *Recall*, *25*(3), 339–356. https://doi.org/10.64152/10125/44343
Lord, G. (2018). Second language pronunciation learning and teaching: a research-based guide. In m. Reed & j. Levis (eds.), *second language pronunciation*. Springer.
MacIntyre, P. D., & Gregersen, T. (2012). Emotions that facilitate language learning: the positive-broadening power of the imagination. *Studies in second language learning and teaching*, *2*(2), 193–213. https://doi.org/10.14746/ssllt.2012.2.2.4
Mayer, R. E. (Ed.). (2014). *The cambridge handbook of multimedia learning* (2nd ed.). Cambridge university press. https://doi.org/10.1017/CBO9781139547369
Mishra, P., & Koehler, M. (2006). Technological pedagogical content knowledge: a framework for teacher knowledge. *Teachers college record*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x
Ortega, L. (2014). *Understanding second language acquisition*. Routledge. https://doi.org/10.4324/9780203777282
Paas, F., & Van Merriënboer, J. J. G. (1994a). Variability of worked examples and transfer of geometrical problem-solving skills: a cognitive-load approach. *Journal of educational psychology*, *86*(1), 122–133. https://doi.org/10.1037/0022-0663.86.1.122
Paas, F., & Van Merriënboer, J. J. G. (1994b). Variability of worked examples and transfer of geometrical problem-solving skills: a cognitive-load approach. *Journal of educational psychology*, *86*(1), 122–133. https://doi.org/10.1037/0022-0663.86.1.122
Pennington, M. C., & Rogerson-Revell, P. (2019). *English pronunciation teaching and research: contemporary perspectives*. Palgrave macmillan. https://doi.org/10.1057/978-1-137-47677-7
Plass, J. L., & Moreno, R. (2010). *Cognitive load theory and instructional design: emerging research and opportunities*. Igi global. https://doi.org/10.1017/CBO9780511844744
Robinson, P. (2011). *Second language task complexity: researching the cognition hypothesis of language learning and performance*. John benjamins.
Ruan, Z., Zhang, Y., & Xu, J. (2022). Integrating speech recognition and language models for adaptive pronunciation training. *Recall*.
Saito, K., & Plonsky, L. (2019). Effects of instruction on l2 pronunciation: a meta-analysis. *Applied linguistics*, *40*(3), 546–569. https://doi.org/10.1093/applin/amx051
Skehan, P. (2014). *Processing perspectives on task performance*. John benjamins. https://doi.org/10.1075/tblt.5

Sweller, J. (2010). Cognitive load theory: recent theoretical advances. In j. L. Plass, r. Moreno, & r. Brünken (eds.), *cognitive load theory* (pp. 29–47). Cambridge university press. https://doi.org/10.1017/CBO9780511844744.004

Thomson, R. I., & Derwing, T. M. (2016). Is phonemic training using instruction or exposure? *Studies in second language acquisition*, *38*(4), 653–678. https://doi.org/10.1017/S0272263115000364

Warschauer, M. (2013). *The digital divide and social inclusion*. Mit press.

Warschauer, M., & Liaw, M. (2011). *Handbook of research in computer-assisted language learning*. Routledge.

Zhang, R., & Zou, D. (2021). Types, purposes, and effectiveness of state-of-the-art technologies for l2 learning. *Computer assisted language learning*, *34*(5–6), 621–664.