GENAI-POWERED PRONUNCIATION PRACTICE: ADAPTING ASR-BASED LESSONS FOR VOICE-ACTIVATED CONVERSATIONAL ASSISTANTS

William Gottardi, Universidade Federal de Santa Catarina and Concordia University
Rosane Silveira, Universidade Federal de Santa Catarina
Walcir Cardoso, Concordia University

Recent years have witnessed constant and substantial improvements in GenAI-powered virtual assistants, providing new opportunities for L2 pronunciation practice. Building on these developments, this study explores the pedagogical affordances of Voice-Activated Conversational Assistants (VACAs) for L2 pronunciation teaching and learning by designing prompts for three VACA-based pronunciation lessons developed for Google Gemini Live. Grounded in established frameworks in CALL and L2 pronunciation pedagogy, the lessons were designed using prompt engineering techniques, namely: defining a persona, establishing constraints, creating a clear context, specifying a response format and feedback focus, and providing predefined sentences for the pronunciation practice. Based on these insights, the proposed VACA-based pronunciation lessons are meant to complement teacher-led instruction, while also enabling learners to practice independently.

INTRODUCTION

Recent advancements in Generative Artificial Intelligence (GenAI)-powered virtual assistants have opened new possibilities for second language (L2) pronunciation instruction, particularly because of their embedded Automatic Speech Recognition (ASR) feature, which has shown significant potential for enhancing L2 pronunciation pedagogy (e.g., Gottardi & Silveira, 2024, 2025; Papin & Cardoso, 2025).

One effective way to integrate GenAI-powered ASR into pronunciation instruction is through Voice-Activated Conversational Assistants (VACAs), such as ChatGPT's advanced voice mode and Google Gemini Live. VACAs offer users the possibility to orally interact with the virtual assistant, supporting a multimodal conversation and providing immediate feedback. In addition, these GenAI-based tools may provide learners with new opportunities for engaging in meaningful human-machine interactions that account for language-related episodes (Chapelle, 2025), and for both instructed and autonomous learning (Godwin-Jones, 2024). Although the versatility of these resources may offer significant benefits to L2 pronunciation pedagogy, further research into the affordances of such tools is needed to offer insights to both learners and educators on how to integrate VACAs into pronunciation practice effectively.

In light of these considerations, the objective of this study is to explore the affordances of VACAs for pronunciation practice by adapting ASR-based pronunciation lessons to VACA-based pronunciation lessons. In addition, we evaluate the appropriateness of VACAs as a Computer-Assisted Language Learning (CALL) resource – grounded in Chapelle's (2001)

Gottardi, W., Silveira, R., & Cardoso, W. (2026). GenAI-powered pronunciation practice: Adapting ASR-based lessons for voice-activated conversational assistants. In P. Trofimovich, A. Bodea, T.-N. N. Le, R. Suzuki, & C. Teló (Eds.), *Proceedings of the 16th Pronunciation in Second Language Learning and Teaching Conference* (pp. 1–13). Concordia University, Montreal, Canada, October 2025.
https://doi.org/10.31274/psllt.22791

criteria – and examine how their affordances can be leveraged as a pronunciation teaching material following insights from Celce-Murcia et al.'s (2010) communicative framework for teaching pronunciation. A detailed design rationale is provided in the following section.

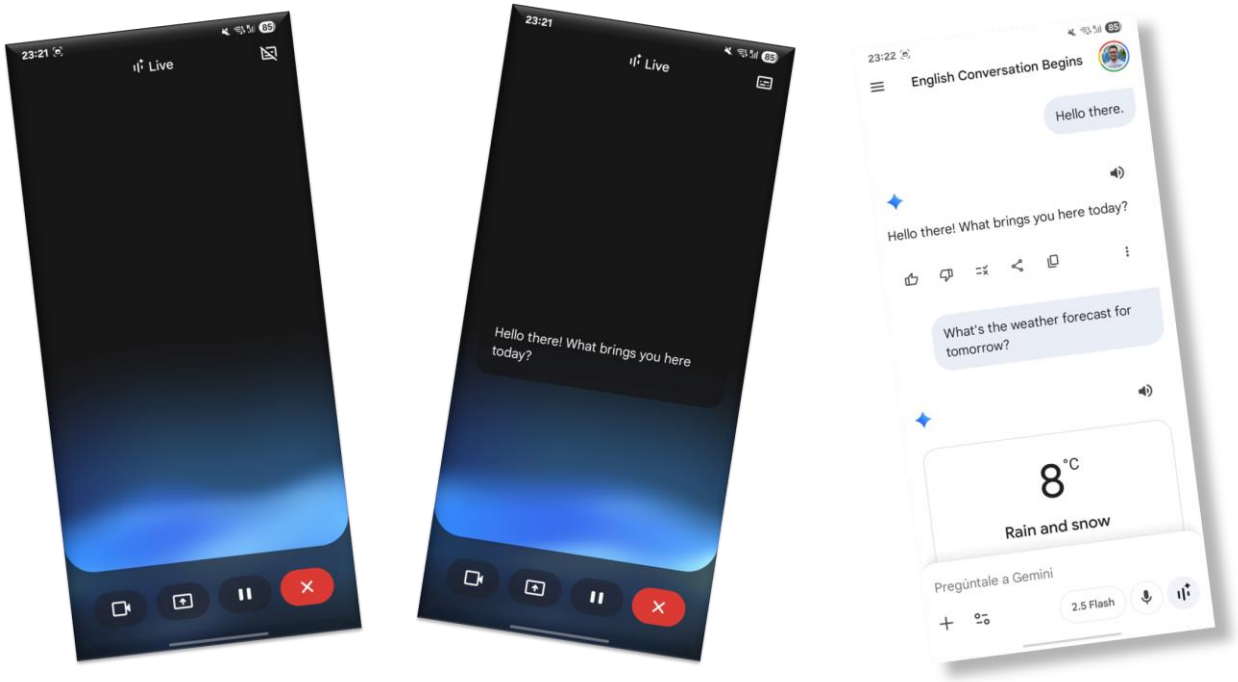
DESIGN RATIONALE

This practice-oriented study examines a free and emerging digital resource and its affordances for L2 pronunciation teaching and independent learning, grounded in Cardoso's (2022) chronological framework for investigating Computer-Assisted Language Learning (CALL) tools. This four-level framework encompasses the (a) development of a new pedagogical tool; (b) exploration of potential and affordances of a new technology; (c) assessment of suitability of a tool by investigating its usability as well as learner's attitudes and acceptance; and (d) assessment of pedagogical effectiveness of a given technology (Cardoso, 2022). This research focuses on the Exploration stage, the second level of the framework, by examining a specific VACA (i.e., Google Gemini Live) as a CALL tool for pronunciation practice and adapting three ASR-based pronunciation lessons for use in this context, as outlined below.

Voice-Activated Conversational Assistant

For designing the VACA-based pronunciation lessons, we adopted the free version of Google Gemini Live (GGL), running on Gemini 2.5 Flash, a Multimodal Large Language Model (MLLM) that can act like an Intelligent Personal Assistant (see Moussalli & Cardoso, 2020 for a pedagogical overview of these assistants) such as Amazon Alexa or Microsoft Cortana. However, GGL can also rely on additional multimodal capabilities (e.g., pictures, files, mobile device's screen and camera sharing) and is not limited to spoken interaction. These multimodal texts can enrich the quality of the feedback and the interaction by adapting and personalizing the pronunciation lesson. During the design of the proposed lessons (conducted in October 2025), Google Gemini's conversational feature (Live)[1] was available exclusively via its mobile applications (iOS and Android). Figure 1 illustrates the GGL interface across three stages: the active Live mode, the Live mode with real-time subtitles enabled during an ongoing conversation, and the transition back to the text-based mode showing the chat history (after exiting the Live mode).

Figure 1
Google Gemini's Conversational Feature (Google Gemini Live)



VACA & CALL Principles

In order to assess GGL suitability for the VACA-based pronunciation lessons, we adopted the six criteria to evaluate CALL resources and activities proposed by Chapelle (2001), namely, language learning potential (Does the tool provide opportunities for focus on form?), learner fit (Is the tool appropriate for learners' proficiency levels and needs?), meaning focus (Does the activity direct attention to meaning?), authenticity (Does the task resemble real-world language use?), positive impact (Does the tool have beneficial effects on learning and instruction?), and practicality (Is the tool feasible to implement given technical and contextual constraints?).

These criteria are operationalized through GGL's affordances, as it allows instructors and designers to prompt which linguistic features the lesson will address (learner fit) and to specify how detailed the corrective feedback will be (language learning potential); it presents some level of adaptability (learner fit; e.g., enable/disable subtitles, change the English variety), being particularly suitable for more advanced learners who are able to negotiate meaning and elaborate their responses (learner fit; language learning potential). Depending on the way the lesson/prompt is designed (e.g., a role-play activity), it is possible to direct learners' attention toward meaning (meaning focus). GGL's multimodal features allow for personalized practice sessions by incorporating the learner's surrounding environment (e.g., sharing a mobile device's screen or camera) or attaching files to the chat session (e.g., a restaurant menu, a shopping list). These features can be used to create a stronger correspondence between the designed CALL task and L2 tasks that learners may face outside the classroom (authenticity; e.g., ordering food at a restaurant, going to the grocery store). In addition, GGL can be used as an autonomous or self-

directed learning tool, relying on its interactive nature and immediate feedback (positive impact). Lastly, GGL requires standard technological resources for Mobile-Assisted Language Learning (e.g., a mobile device with internet access, the installed application, an active user account) and speech-based pronunciation practice (e.g., a stable connection, quiet environment for optimal speech recognition), which can be argued as common requirements within this context (practicality).

In summary, the potential affordances of GGL for pronunciation practice are numerous. Nevertheless, the quality of its output is shaped by the clarity of the prompt used during the VACA-based pronunciation lesson. In turn, the quality of the prompt itself will depend on the pedagogical planning of the lesson. A lesson plan that lacks a clear learning objective and that neither targets specific linguistic forms nor respects learners' individual characteristics compromises the prompt design, resulting in a deficient interaction with the application. With this in mind, the background and rationale for lesson and prompt design are presented next.

Lesson Design and Rationale

To explore GGL's affordances for pronunciation teaching, we adapted three ASR-based pronunciation lessons (originally designed in a previous study; Gottardi, 2023) to capitalize on the tool's interactive nature. The original ASR-based lessons, which served as the reference for the new VACA-based lessons, were previously assessed by twelve in-service L2 English teachers who indicated their perceived classroom applicability, usability, and instructional relevance. These original lessons aligned with established frameworks in CALL (the six criteria for CALL task appropriateness; Chapelle, 2001) and L2 pronunciation pedagogy (the Communicative Framework for Teaching Pronunciation; Celce-Murcia et al., 2010). These lessons addressed pronunciation issues that commonly affect the intelligibility of Brazilian learners of L2 English (Gonçalves & Silveira, 2015; Silveira et al., 2017), the population for whom these lessons were designed.

We selected three lessons from the original study to be redesigned by integrating the VACA's speech and interactive capabilities, which enable both real-time oral exchanges (e.g., filling in gaps using learner's personal information) and immediate oral feedback (e.g., via GGL's accurate interpretation of learner's input, through signaling communication breakdowns by VACA-initiated phrases such as "Can you repeat what you just said?").

Furthermore, the redesigned VACA-based lessons also follow insights from the Communicative Framework for Teaching Pronunciation (Celce-Murcia et al., 2010). This framework comprises five phases (description and analysis, listening discrimination, controlled practice, guided practice, communicative practice) suggesting a progressive improvement of phonological competence. While the first phase focuses on the development of phonological awareness and the second on perception skills, the final three phases emphasize oral production, moving from controlled exercises (e.g., reading sentences aloud) toward more communicative practice centered on oral interaction (e.g., role-play activities). The lessons proposed in this study primarily focus on phases 3 (Controlled Practice) and 4 (Guided Practice). Nevertheless, whenever GGL runs out of the pre-defined sentences within the prompts, it will start suggesting new sentences. Then, learners must decide and negotiate the sentences they want to practice

next. Thus, the lessons open space for extended oral interaction, guided by learners' readiness to pursue further practice (phase 5, Communicative Practice).

Prompt Design and Rationale

The interaction prompts with GGL were carefully designed to reflect CALL/Second Language Acquisition pedagogical principles (e.g., providing access to meaningful input, a focus on both linguistic form and meaning; see Celce-Murcia et al., 2010; Chapelle, 2001), while targeting pronunciation issues that may affect intelligibility. The prompts also incorporate the implementation procedures described in the original ASR-based pronunciation lessons to define the context, learning objectives, and the instructions provided to GGL (see Gottardi, 2023).

Moreover, to ensure a more pedagogically sound interaction with GGL, specific prompt engineering techniques were adopted, following insights from Gadesha (2025) and Google Cloud (2025b, 2025c). According to Jurafsky and Martin (2025, p. 151), prompt engineering refers to "the process of finding effective prompts for a task". In this case, the task was intimately related to the learning objective of each pronunciation lesson. The main techniques consist of defining a persona, establishing constraints, creating a clear context, specifying a response format/feedback focus, and providing GGL with a model of sentences to be used during pronunciation practice, also called few-shot prompting. Finally, the prompt design incorporated current recommendations and guidelines for the responsible use of AI (Google Cloud, 2025a; Lu et al., 2024; UNESCO, 2024)

Finally, during the iterative process of designing and testing the interaction prompts (October 2025), we captured screenshots to document examples of user-VACA interactions. The following section presents the VACA-based lessons and their corresponding prompts, using these screenshots to illustrate some of the affordances of GGL as a pedagogical tool.

VACA-BASED PRONUNCIATION LESSONS

This section presents the three designed VACA-based pronunciation lessons. The interaction prompts for each lesson explicitly stated their learning focus and target pronunciation feature. They were designed mainly for adults, especially because of the required technology knowledge needed to use these virtual assistants, and privacy issues related to the use of Large Language Models. However, they can be adapted to different audiences and used with different proficiency levels, including beginners. In addition, the duration of each lesson may vary, since GGL can continue the practice session if the learner decides to (in a self-directed learning environment), or the teacher finds it appropriate, according to the level of each student or group of students (in a teacher-guided practice context). Nevertheless, learners can pause the practice session and resume whenever it is more convenient or opt to restart the session from the beginning.

The following subsections present the three lessons designed for this study. The prompts related to the lessons were formatted into a table to clarify their design. For each lesson, screenshots illustrate GGL's behaviors and the nature of the interactions during the practice sessions.

Lesson 1 – Personal Information

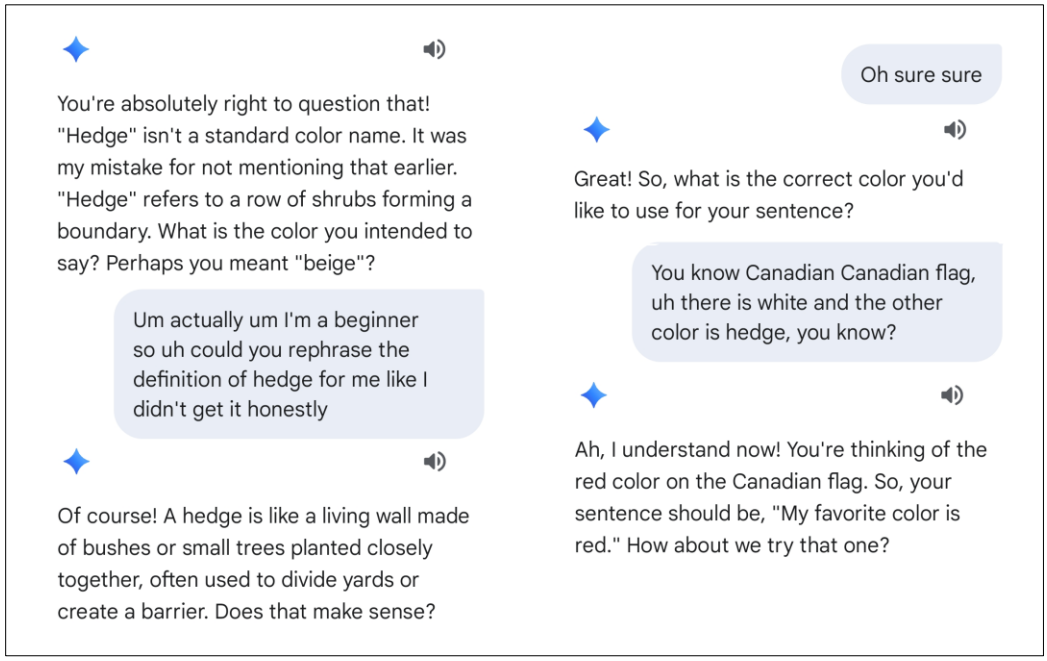
This lesson suggests practicing basic communicative skills through the exchange of personal information with GGL. The prompt can be seen in Table 1. It can work as a pronunciation self-assessment lesson, whereby the learner is invited to reflect on some of their pronunciation patterns that might disrupt oral communication (the target feature). Any communication breakdown will be signaled by GGL during the oral interaction (e.g., clarification questions such as "Can you repeat?") or via chat history after the practice session is resumed. This lesson can be easily tailored to learners' needs, preferences, and proficiency level by modifying the pre-defined set of sentences. To protect learners' sensitive information, a fictional character can be used instead of real personal information. This alternative also creates opportunities for further linguistic development by allowing learners to work within a fictional context and interact with fictional characters.

Table 1
Lesson 1 Prompt

Context	Hello there! Let's speak in English. I am a beginner learner of English as a second language, and I want to practice my pronunciation skills. Bear in mind that, as I am a learner, some pauses and interruptions in my speech are expected.
Learning focus	The learning focus of the lesson is to raise awareness about pronunciation patterns that might disrupt communication.
Instruction	I will provide you with some sentences below. In those sentences, there are some blanks I want to fill in with my personal information. The topic of the vocabulary that is expected to be used is expressed between brackets representing the blanks; therefore, check if the word I say makes sense according to the topic between the brackets. These sentences are part of a self-assessment of my pronunciation. I want to say each sentence, one at a time, following the numeric order.
Feedback focus	Please provide me with feedback on any pronunciation patterns that may cause communication breakdowns. The goal is to help me improve my pronunciation rather than accommodate unintelligible speech.
Interaction rules	You will suggest the next sentence only if I say "next", or else we will continue practicing the same sentence. Finally, I want to practice using Google Gemini Live.
Sentences (few-shot prompting):	Sentences: 1. My favorite food is [food]. 2. My favorite drink is [drink]. 3. My favorite color is [color]. 4. [new sentence].

As an example of a practice session (Figure 2), one can see how GGL interacts with the learner when a communication breakdown occurs. In this interaction, the expected response was the color *red*, but the application misinterpreted the learner's pronunciation as *hedge*, reflecting a pronunciation pattern commonly observed among Brazilian learners of L2 English (i.e., /r/ to [h] substitution and /d/ palatalization). Following this, GGL offered an opportunity to negotiate meaning, enabling the learner to explain what color they referred to.

Figure 2
Lesson 1 Practice Session



Lesson 2 – Vowel Contrast

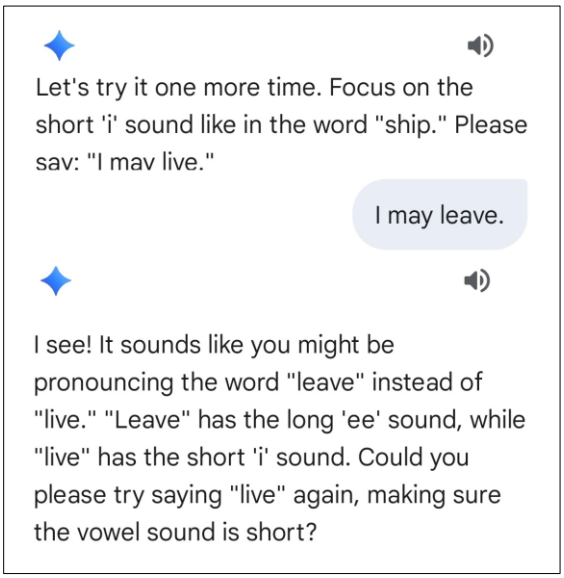
This is a classic pronunciation lesson designed to practice the target /ɪ/-/i/ vowel contrasts in an ambiguous linguistic context, offering a practical way to provide extended pronunciation practice of these challenging contrasts. However, this lesson might be overwhelming for learners if they are unable to perceive or produce the minimal pairs intelligibly. Therefore, teacher guidance might be necessary to help learners overcome the target pronunciation difficulties, set a realistic learning goal, and establish an appropriate practice time. Table 2 illustrates the prompt, which can be adapted for other vowel contrasts by changing the instructions and feedback as well as the pre-set sentences within the prompt.

By offering GGL a set of sentences, the system is better able to understand which words are expected, thereby fine-tuning the quality of the feedback and, consequently, the learning experience. Figure 3 displays the type of feedback GGL can provide to learners, illustrating the orthographic feedback available after leaving the conversation mode. During practice, learners have access to the feedback orally, provided by GGL's Text-to-Speech (TTS) feature, but they can also enable the live captions feature to read what GGL is saying (e.g., to support the establishment of grapheme-to-phoneme rules in English).

Table 2
Lesson 2 Prompt

Context	Hello there! Let's speak in English. I am a beginner learner of English as a second language, and I want to practice my pronunciation skills. Bear in mind that, as I am a learner, some pauses and interruptions in my speech are expected.
Learning focus	The learning focus of the activity is to raise awareness about the acoustic features of these vowels and how to distinguish them in production.
Instruction	I will provide you with some sentences below for a minimal pair practice session involving the production of the vowel contrast /ɪ/ and /i/. In each sentence, the last word contains the target feature. I want to say each sentence, one at a time, following the numeric order.
Feedback focus	Please provide me with feedback related to those words containing the vowel contrast /ɪ/ and /i/ that are expected in each sentence. The goal is to help me improve my pronunciation rather than accommodate unintelligible speech.
Interaction rules	You will suggest the next sentence only if I say "next", or else we will continue practicing the same sentence. Finally, I want to practice using Google Gemini Live.
Sentences (few-shot prompting):	Sentences: 1. Look at that sheep! 2. Look at that ship! 3. I may live. 4. I may leave. 5. It is a sin. 6. It is a scene. 7. [new minimal pair sentence]

Figure 3
Lesson 2 Practice Session



Lesson 3 – Tongue Twister

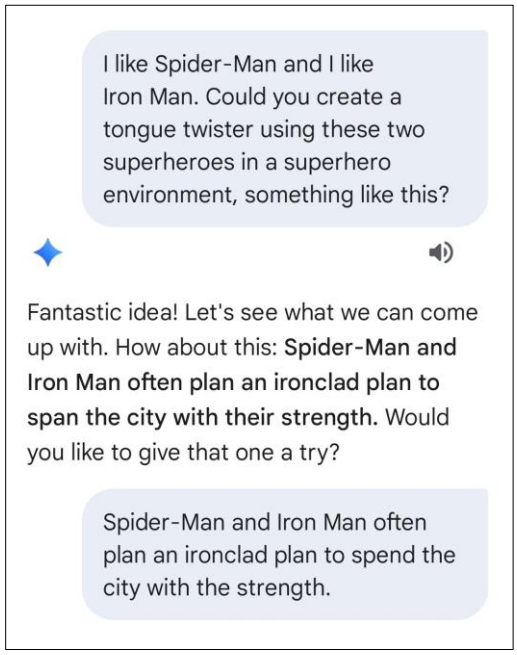
The last lesson offers a dynamic and adaptable guided practice using tongue twisters. The prompt (Table 3) offers a few initial suggestions of tongue-twisters, but they can be tailored to learners' needs. This prompt was designed to offer a potentially never-ending tongue-twister practice for difficult pronunciation targets, giving learners the possibility to practice different pronunciation features over time.

Table 3
Lesson 3 Prompt

Context	Hello there! Let's speak in English. I am a beginner learner of English as a second language, and I want to practice my pronunciation skills. Bear in mind that, as I am a learner, some pauses and interruptions in my speech are expected.
Learning focus	The learning focus of the activity is to use tongue twisters to improve my pronunciation accuracy and speaking fluency by targeting difficult patterns that hinder intelligibility and communication.
Instruction	I will provide you with some tongue twisters below. I want to say each sentence, one at a time, following the numeric order.
Feedback focus	Please provide me with feedback on any unexpected pronunciation pattern that deviates from the expected tongue twister. The goal is to help me improve my pronunciation rather than accommodate unintelligible speech.
Interaction rules	You will suggest the next sentence only if I say "next", or else we will continue practicing the same sentence. Finally, I want to practice using Google Gemini Live.
Sentences (few-shot prompting):	Tongue twisters: 1. I was thinking of thanking you. 2. She sells seashells by the seashore. 3. I rarely see a real rat in the mirror. 4. [new tongue twister].

For this activity, learners can request that GGL generate customized tongue-twisters containing words or phrases that might be either meaningful or challenging to them (see Figure 4 for an example). In this context, learners can receive aural input via GGL's TTS feature and orally produce the tongue-twister via GGL's ASR feature, thus creating an interactive input/output loop. In the same practice session, learners can practice difficult words or phrases repeatedly until they feel ready to proceed with the whole sentence. As a result, learners can create tongue-twisters by suggesting to GGL the use of specific words or phrases, emphasizing personal preferences and interests (e.g., Spider-Man and Iron Man – see Figure 4), or professional and personal needs (e.g., difficult words or phrases encountered on learners' daily routines).

Figure 4
Lesson 3 Practice Session



CONCLUSIONS AND PEDAGOGICAL IMPLICATIONS

The proposed VACA-based pronunciation lessons are intended to support both classroom-based, teacher-guided instruction as well as autonomous, self-directed pronunciation practice. They offer educators suggestions on how to integrate a widely accessible VACA (Google Gemini Live – GGL) into L2 pronunciation practices along with ready-to-use and editable prompts. As discussed in this article, VACAs can be a convenient CALL resource that fulfills the criteria for CALL task appropriateness (Chapelle, 2001, 2017). Moreover, specific VACA affordances, such as real-time interaction and the ability to save and resume progress, make these pronunciation lessons easier to integrate into instruction than TTS- and ASR-based activities.

During L2 interactions, learners can ask VACAs for linguistic support, clarifications, and explicit information about linguistic forms (e.g., pronunciation features), facilitating their progress toward the lesson's objective. In addition, oral interaction with the application enables learners to negotiate meaning during the practice session while receiving immediate oral feedback (via TTS; e.g., when the tool requests clarification) or delayed orthographic feedback (via the chat history). These practice sessions can be personalized to focus on linguistic form and meaning, allowing learners to practice and engage in L2 interactions multiple times, both within and beyond the confines of the L2 classroom. To capitalize on the conversational affordances of VACAs, this study proposes lessons designed to support controlled and guided pronunciation practice (Celce-Murcia et al., 2010), while also offering opportunities for learner interaction with the tool.

While VACAs are a versatile resource, they are subject to specific limitations. As an off-the-shelf virtual assistant, GGL is designed for general user assistance rather than pronunciation assessment or instruction. For this reason, GGL may prioritize task completion over the detection

of unintelligible speech, occasionally accommodating learner output to maintain conversational flow. Nonetheless, the use of engineered prompts that explicitly define the pedagogical objectives and constrain undesired behaviors may override VACA's native tendencies, ensuring the interaction remains pedagogically aligned. Finally, teachers may help circumvent some of these limitations by verifying the adequacy of the VACA's feedback, assisting learners in determining the appropriate duration for each practice session, and helping them overcome communication breakdowns.

To conclude, more research is necessary to investigate the affordances and limitations of VACAs for pronunciation teaching, including how instructors can cope with constant upgrades and new software releases. An interesting line of research is to explore both teachers and learners' perceptions of such technology and its role in L2 pronunciation pedagogy.

ACKNOWLEDGMENTS

The authors acknowledge the financial support of CAPES (Gottardi), CNPq (Silveira), and SSHRC (Cardoso). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 (Gottardi).

ABOUT THE AUTHORS

William Gottardi (prof.williamgottardi@outlook.com) is a doctoral candidate at the English Graduate Program at the Universidade Federal de Santa Catarina and an Academic Visitor at Concordia University. His research interests include digital technologies for second language (L2) pedagogy, L2 pronunciation teaching, and CALL/ICALL.

Rosane Silveira (rosanesilduarte@gmail.com) is an associate professor at the Department of Foreign Languages and Literature at the Universidade Federal de Santa Catarina. She also holds a research fellowship granted by the National Council for Scientific and Technological Development (CNPq). She conducts and supervises graduate and undergraduate research in the area of L2 Phonetics and Phonology.

Walcir Cardoso (walcir.cardoso@concordia.ca) is a Professor of Applied Linguistics at Concordia University, Canada. He conducts research on the L2 acquisition of phonology, morphosyntax and vocabulary, and the effects of computer technology (e.g., clickers, text-to-speech synthesizers, automatic speech recognition, intelligent personal assistants) on L2 learning.

REFERENCES

Cardoso, W. (2022). Technology for speaking development. In T. Derwing, M. Munro, & R. Thomson (Eds.), *The Routledge handbook of second language acquisition and speaking* (pp. 299–313). Routledge.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation hardback with audio CDs (2): A course book and reference guide* (2nd ed.). Cambridge University Press.

Chapelle, C. A. (2001). *Computer applications in second language acquisition*. Cambridge University Press.

Chapelle, C. A. (2017). Evaluation of technology and language learning. In C. A. Chapelle & S. Sauro (Eds.), *The handbook of technology and second language teaching and learning* (pp. 378–392). Wiley.

Chapelle, C. A. (2025). Generative AI as game changer: Implications for language education. *System*, *132*, Article 103672. https://doi.org/10.1016/j.system.2025.103672

Gadesha, V. (2025). *Prompt Engineering Techniques*. IBM. https://www.ibm.com/think/topics/prompt-engineering-techniques

Godwin-Jones, R. (2024). Distributed agency in second language learning and teaching through generative AI. *Language Learning & Technology*, *28*(2), 5–31. https://doi.org/10.64152/10125/73570

Gonçalves, A. R., & Silveira, R. (2015). Intelligibility research in Brazil: Empirical findings and methodological issues. *Revista Horizontes de Linguistica Aplicada*, *14*(1), 51–81. https://doi.org/10.26512/rhla.v14i1.1384

Google Cloud. (2025a). *Gemini for Google Cloud and responsible AI*. Google Cloud Documentation. https://docs.cloud.google.com/gemini/docs/discover/responsible-ai

Google Cloud. (2025b). *Introduction to prompting*. Google Cloud Documentation. https://docs.cloud.google.com/vertex-ai/generative-ai/docs/learn/prompts/introduction-prompt-design

Google Cloud. (2025c). *Prompt Engineering for AI Guide*. Google Cloud. https://cloud.google.com/discover/what-is-prompt-engineering

Gottardi, W. (2023). *Automatic speech recognition as a pronunciation teaching resource: What are the in-service teachers' perceptions of this speech technology?* [Unpublished master's thesis]. Universidade Federal de Santa Catarina (UFSC). https://tede.ufsc.br/teses/PPGI0228-D.pdf

Gottardi, W., & Silveira, R. (2024). Automatic speech recognition as a pronunciation teaching resource: In-service teachers' perceptions. In D. J. Olson, J. L. Sturm, O. Dmitrieva, & J. M. Levis (Eds)., *Proceedings of 14th Pronunciation in Second Language Learning and Teaching Conference* (pp. 1–15). Purdue University, September 2023. https://doi.org/10.31274/psllt.17567

Gottardi, W., & Silveira, R. (2025). A theoretical assessment of ASR-based pronunciation activities for classroom learning. *Texto Livre*, *18*, 56636. https://doi.org/10.1590/1983-3652.2025.56636

Jurafsky, D., & Martin, J. H. (2025). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition: Unpublished manuscript*. https://web.stanford.edu/~jurafsky/slp3/ed3book_Jan25.pdf.

Lu, Q., Zhu, L., Whittle, J., & Xu, X. (2024). *Responsible AI: Best practices for creating trustworthy AI systems*. Addison-Wesley.

Moussalli, S., & Cardoso, W. (2020). Intelligent personal assistants: Can they understand and be understood by accented L2 learners? *Computer Assisted Language Learning*, *33*(8), 865–890.

Papin, K., & Cardoso, W. (2025). Assessing the pedagogical potential of Google Translate's speech capabilities: Focus on French pronunciation. *CALICO Journal*, *42*(1), 1–24. https://doi.org/10.3138/calico-2025-0117

Silveira, R., Delatorre, F., Reis, L. P., & Gonçalves, A. R. (2017). Percepção, produção e inteligibilidade do inglês falado por usuários brasileiros. *Perspectivas Atuais de Aprendizagem e Ensino de Línguas. Florianópolis: LLE/CCE/UFSC*, 237–283.
UNESCO. (2024). *AI competency framework for teachers*. https://doi.org/10.54675/ZJTE2084

[1]For an overview of GGL's main features, technical requirements, and how to access it on your mobile device, check Gemini's help center at https://support.google.com/gemini/answer/15274899?hl=en&co=GENIE.Platform%3DAndroid