UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE COMUNICAÇÃO E EXPRESSÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INGLÊS

WILLIAM GOTTARDI

AUTOMATIC SPEECH RECOGNITION AS A PRONUNCIATION TEACHING RESOURCE: WHAT ARE THE IN-SERVICE TEACHERS' PERCEPTIONS OF THIS SPEECH TECHNOLOGY?

FLORIANÓPOLIS
2023

William Gottardi

Automatic Speech Recognition as a Pronunciation Teaching Resource: What Are the In-Service Teachers' Perceptions of This Speech Technology?

Dissertação submetida ao Programa de Pós-graduação em Inglês: Estudos Linguísticos e Literários da Universidade Federal de Santa Catarina para a obtenção do título de Mestre em Inglês: Estudos Linguísticos e Literários.

Orientadora: Profa. Dra. Rosane Silveira

Florianópolis

2023

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Gottardi, William
Automatic Speech Recognition as a Pronunciation Teaching Resource: What Are the In-Service Teachers' Perceptions of This Speech Technology? / William Gottardi ; orientadora, Rosane Silveira, 2023.
132 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Comunicação e Expressão, Programa de Pós Graduação em Inglês: Estudos Linguísticos e Literários, Florianópolis, 2023.

Inclui referências.

1. Inglês: Estudos Linguísticos e Literários. 2. L2 Pronunciation Teaching. 3. Automatic Speech Recognition. 4. Teacher Cognition. 5. CALL. I. Silveira, Rosane. II. Universidade Federal de Santa Catarina. Programa de Pós Graduação em Inglês: Estudos Linguísticos e Literários. III. Título.

William Gottardi

Automatic Speech Recognition as a Pronunciation Teaching Resource: What Are the In-Service Teachers' Perceptions of This Speech Technology?

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Profª. Drª. Rosane Silveira

Orientadora e Presidente

Universidade Federal de Santa Catarina (UFSC)

Prof. Dr. Celso Henrique Soufen Tumolo

Universidade Federal de Santa Catarina (UFSC)

Prof. Dr. Walcir Cardoso

Concordia University (Montreal, Canada)

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título Mestre em Inglês: Estudos Linguísticos e Literários.

Documento assinado digitalmente
Celso Henrique Soufen Tumolo
Data: 11/04/2023 10:55:24-0300
CPF: ***.275.298-**
Verifique as assinaturas em https://v.ufsc.br

_____

Prof. Dr. Celso Henrique Soufen Tumolo

Coordenação do Programa de Pós-Graduação em Inglês

Documento assinado digitalmente
Rosane Silveira
Data: 11/04/2023 11:30:46-0300
CPF: ***.936.839-**
Verifique as assinaturas em https://v.ufsc.br

_____

Profª. Drª. Rosane Silveira

Orientadora

Florianópolis, 2023.

ACKNOWLEDGMENTS

First, I want to thank my family. Starting with my parents because they always have been so supportive of my professional and personal development. I want to thank my brother for being such a great educator. He is an inspiration to me. Also, I want to thank my girlfriend who has helped me countless times during this past year. Without her support and her company, this journey would have been much harder. I will also never forget the support and friendship of Micaella de Lima. She is a friend who has become a part of my family.

Many people have helped me during my studies. From my former higher education institution, Universidade Regional de Blumenau (FURB), I owe a special thank-you note to Professor Dra. Cyntia Bailer. She was the one who encouraged me to start my master's at UFSC and has helped me become a better student and educator. I also want to express my gratitude to my former professors, Marina Borgmann Cunha, and Diva Rangel Martinelli. Their passion for teaching has always inspired me. In Blumenau, I also had the opportunity to learn so much from Paulina Gajardo, *mi maestra*, not only about Spanish but also about being a teacher who is always willing to learn and to try to empower students.

I am grateful to the MA program I joined (PPGI) because I had the opportunity to have classes with incredible professors. Special thanks to Professor Celso Henrique Soufen Tumolo and Professor Hanna Kivistö de Souza who had worked with me on other research projects. During my MA, I met an MA candidate, Janaina Fernanda de Almeida, who has helped me immensurable during this academic journey and has become a great friend. Similarly, I want to thank Luana Garbin Baldissera. She has given me advice since the MA selection process.

Finally, I want to express my gratitude to my advisor, Professor Rosane Silveira, who has guided me throughout this journey, giving me support and solid recommendations for my thesis' improvement. In a similar vein, I want to thank, once more, Professor Celso Henrique Soufen Tumolo and Professor Walcir Cardoso for having kindly accepted the invitation to be part of this thesis' defense committee. They also provided me with important feedback during the design of my research. Lastly, at UFSC, I had the opportunity to receive a CAPES scholarship which allowed me to dedicate myself to my research. In short, without all these people, this research would never have happened.

RESUMO

Pronúncia é um elemento chave para alcançar uma comunicação oral bem-sucedida. No entanto, em muitos programas de língua, a pronúncia é um componente comumente negligenciado (DERWING, 2010) principalmente devido a restrições de tempo e à falta de conhecimento sobre implicações pedagógicas para o ensino de pronúncia. As tecnologias digitais podem ajudar no ensino da pronúncia, especialmente o Reconhecimento Automático da Fala (ASR – Automatic Speech Recognition). Este estudo procurou explorar as possibilidades do ASR para o ensino da pronúncia a partir da perspectiva dos professores de inglês atuantes e investigar a avaliação dos professores sobre atividades de pronúncia baseadas no uso de ASR. Para isso, uma oficina on-line foi realizada para dois grupos diferentes a fim de reunir as percepções dos participantes sobre essa tecnologia. Além disso, sete atividades de pronúncia baseadas em ASR foram disponibilizadas para a avaliação dos participantes. As atividades foram criadas respeitando os seis critérios de adequação às tarefas CALL (CHAPELLE, 2001) e as fases 3 a 5 da estrutura comunicativa para o ensino da pronúncia do inglês (CELCE-MURCIA; BRINTON; GOODWIN, 2010). Antes da oficina, os participantes responderam a um questionário online para coletar informações contextuais e outro questionário online foi aplicado após a oficina para coletar suas percepções em relação ao uso de ASR para o ensino de pronúncia. Um total de 5 participantes estiveram presentes na primeira sessão e outros 7 na segunda (n=12). Todos os participantes eram professores de inglês, do setor público ou privado. Os dados dos questionários foram analisados quantitativamente através de estatísticas descritivas. Uma pergunta aberta do questionário final, as transcrições das gravações das oficinas e os registros de bate-papo salvos foram analisados qualitativamente. Os dados foram categorizados durante uma análise sistemática de conteúdo, codificados e depois analisados qualitativamente. As percepções dos participantes indicam que o ASR pode 1) ser usado como uma sugestão de ferramenta para aprendizagem autônoma; 2) ajudar a incentivar os alunos a produzir mais linguagem oral na língua-alvo fora da sala de aula; 3) fornecer aos alunos um feedback ortográfico relevante; 4) ser usado como um suplemento fora da sala de aula; e 5) ser um recurso auxiliar adequado para o ensino da pronúncia em aulas regulares, especialmente, em um ambiente híbrido. Além disso, os participantes relataram ter percepções positivas em relação às atividades e como elas foram apresentadas durante a oficina. Eles também indicaram que as atividades forneceram aos professores suporte técnico, procedimentos e técnicas para utilizar o ASR de forma prática. Conclui-se que as possibilidades do uso de ASR para o ensino da pronúncia são numerosas. No entanto, a mediação dos professores é de suma importância para um melhor resultado. Assim, os resultados apresentados podem ajudar os professores a contornar algumas das limitações do ASR e se beneficiar de suas potencialidades. Em resumo, esta pesquisa contribui com o campo da linguística aplicada ao oferecer ideias sobre como usar o ASR para o ensino da pronúncia e apoio pedagógico para que os professores possam usar esta tecnologia com confiança em suas aulas de inglês.

Palavras-chave: Ensino de pronúncia em L2. Reconhecimento automático da fala. Cognição do professor.

ABSTRACT

Pronunciation is a key element in achieving successful spoken communication. Yet, in many language programs, pronunciation is a commonly neglected component (DERWING, 2010) mainly due to time constraints and a lack of knowledge concerning the pedagogical implications to teach pronunciation effectively. Digital technologies may aid pronunciation teaching, especially Automatic Speech Recognition (ASR). This study sought to explore the affordances of ASR for pronunciation teaching from the perspective of in-service English teachers and investigate teachers' appraisal of ASR-based pronunciation activities designed to be implemented in L2 English classes. To achieve these objectives, this research followed a mixed-method approach. An online workshop session was delivered to two different groups of participants in order to gather participants' perceptions of ASR technology for pronunciation teaching. In addition, seven ASR-based pronunciation activities were made available for their appraisal. The activities were designed respecting the six criteria for CALL task appropriateness (CHAPELLE, 2001) and phases 3 to 5 of the communicative framework for teaching English pronunciation (CELCE-MURCIA; BRINTON; GOODWIN, 2010). Participants answered an online background questionnaire before the workshop session to collect background information, and an online survey after the workshop session to collect their perceptions of ASR for pronunciation teaching. A total of 5 participants joined the first session while 7 joined the second session (n=12). All participants were in-service English teachers from both the public and private sectors. Data from the background questionnaire and the online survey were analyzed quantitatively by running descriptive statistics. An open-ended question from the online survey, the transcriptions of the workshop session recordings, and the saved chat logs were analyzed qualitatively. Data were categorized during a systematic content analysis, coded, and then analyzed qualitatively. The overall perceptions of the participant-teachers indicate that ASR can 1) be used as a tool for self-studying; 2) help encourage learners to produce more output outside the classroom; 3) provide students with relevant orthographic feedback; 4) be used as an out of class supplement; and 5) be an adequate auxiliary resource for pronunciation teaching in regular classes, especially, in a hybrid environment. In addition, participant-teachers reported holding positive attitudes towards the ASR-based pronunciation activities and how they were presented during the workshop session. They also indicated that the activities provided teachers with procedures and techniques to use ASR for pronunciation teaching in a practical way. It is concluded that the affordances of ASR for pronunciation teaching are numerous. Yet, teachers' guidance is of paramount importance for an optimal result. Thus, the presented findings and pedagogical implications may help teachers to circumvent some of the ASR limitations and benefit from ASR affordances. All in all, this research contributes to the field of applied linguistics by offering insights into how to use ASR technology for pronunciation teaching and what further support teachers need in order to use this technology confidently in their L2 English classes.

Keywords: L2 Pronunciation Teaching. Automatic Speech Recognition. Teacher Cognition.

TABLE OF CONTENTS

1 INTRODUCTION

1.1 STATEMENT OF THE PROBLEM

In many language programs, pronunciation is a commonly neglected component (DERWING, 2010). Moreover, studies suggest further research in pronunciation teaching (COSTA, 2016); investigation into how pronunciation pedagogy courses can support foreign language teachers (BUSS, 2017), and a need of collecting data regarding pronunciation teaching to provide solid recommendations to teachers (MUNRO; DERWING, 2015).

Bearing in mind this scenario, Celce-Murcia et al (1996) state that teachers must have an up-to-date repertoire of procedures and techniques to teach pronunciation appropriately. Intriguingly, few education programs offer courses regarding second language (L2[1]) pronunciation teaching (BAKER, 2014), which may cause teachers to lack confidence in teaching it and to express their willingness to have more education related to it (BUSS, 2017; COSTA, 2016).

Furthermore, Pennington and Rogerson-Revell (2019) point out that there is also a lack of knowledge concerning pedagogical implications because research studies usually offer insufficient information related to how teachers can adopt the study's suggestions into their practices, considering their teaching context. The authors further acknowledge that these research studies might not even be available for teachers since many research-oriented journals are not easily accessible. Taking this into consideration, there might be a mismatch between teacher education, research, and pronunciation teaching. This research seeks to contribute towards understanding this mismatch and propose pedagogical uses of a speech technology for L2 pronunciation teaching to L2 English teachers.

To add more complexity to pronunciation teaching, time constraint imposes a barrier to it. Munro and Derwing (2015) emphasize the need of making smart choices regarding the English curriculum. The authors indicate that, besides knowledge to teach pronunciation, teachers must optimize their practices to be able to handle students' communicative needs, including pronunciation. Considering the aforementioned arguments, the use of speech technologies might be an alternative to circumvent this time constraint and provide learners with more pronunciation practice opportunities. In addition, once the digital technology is so interwoven into our daily lives, teaching without any use of it would create an abnormal learning environment (CHUN; SMITH; KERN, 2016).

Digital technologies demonstrate great potential for both pronunciation teaching and learning (REVELL-ROGERSON, 2021). Automatic Speech Recognition (ASR) technology, for example, can facilitate pronunciation improvement and provide learners with instant feedback (GOLONKA et al., 2014). This technology is now available for free as a built-in feature in varied websites and programs. One ASR application is to transcribe an oral input into sequences of words (YU; DENG, 2015). Therefore, a convenient use of ASR technology is through *dictation tools*[2] that

[1] The term second language (L2) will be used throughout this research with no distinction between English as a second, foreign, additional language, or lingua franca. For a detailed definition of these terms see (JORDÃO, 2014).

[2] In this study a distinction is made between *dictation tools* and *speech recognition*. The former is related to the ASR application for voice dictation. The letter is a broader term that encompasses, for instance, voice dialing, voice search, digital assistance (e.g., Siri on iPhone, Google Now on Android, or Cortana on Windows 10 OS), voice control in a

are flexible to use (MCCROCKLIN, 2019) and can foster learner autonomy (KIM, 2006; LIAKIN; CARDOSO; LIAKINA, 2017; MCCROCKLIN, 2016; MROZ, 2018). These programs are especially suitable for learners with little access to the target language outside the classroom once it offers an opportunity for learners to produce more output with instant orthographic feedback (LIAKIN; CARDOSO; LIAKINA, 2017).

Taking into consideration all the ideas previously presented, this research seeks to investigate the use of ASR from the perspective of in-service English teachers. Such technology is particularly important due to its potential to aid pronunciation learning and teaching as an in- and out-classroom resource (CARDOSO, 2022). This flexibility may represent an opportunity to circumvent the time constraint and the lack of resources that may hinder pronunciation teaching in L2 English classes. In the next section, the objectives of this research are presented in detail.

1.2 OBJECTIVES

The objective of this research is twofold:

1. Explore the affordances of ASR for pronunciation teaching from the perspective of in-service English teachers.

2. Investigate teachers' appraisal of ASR-based pronunciation activities designed to be implemented in L2 English classes.

1.3 RESEARCH QUESTIONS

In order to achieve the objectives previously stated, this research intends to answer the following research question:

1. What are the in-service English teachers' perceptions of ASR for pronunciation teaching after attending a workshop on how to use this speech technology regarding the following constructs?
    a) Teacher Development Needs
    b) ASR Accessibility
    c) ASR Affordances
2. How do in-service teachers appraise the ASR-based pronunciation activities designed to be implemented in L2 English classes?

1.4 SIGNIFICANCE OF THE RESEARCH

This study intends to present theoretical insights into possible classroom practices and supply teachers with useful resources to be incorporated into their daily practices. As Costa (2016,

---

smart home and home automation, in-vehicle navigation, and entertainment (LI et al., 2016). Many ASR applications do not provide any orthographic feedback as opposed to ASR-based dictation tools.

p. 63) discusses, teachers should have freedom to choose what to teach to help learners to achieve communicative competence. Thus, it is expected that this study will provide teachers with relevant information regarding ASR technology applied to pronunciation teaching so they can decide what path they want to follow to achieve their pedagogical goals. Furthermore, it is hoped that this research offers practical suggestions on how to implement ASR technology in L2 English classes in order to help learners develop intelligible and comprehensible speech. Although this research focuses on English as the target language, the discussion can encompass different L2s supported by the ASR technology.

1.5 SUMMARY AND ORGANIZATION OF THE CHAPTERS

This first chapter presented the contextualization of the research, the objectives, and the research questions. The following section presents the review of the literature (chapter 2). Then, the method is explained (chapter 3). Next, the results are presented and discussed (chapter 4). The last one (chapter 5) presents the conclusion drawn by this study.

2 REVIEW OF THE LITERATURE

This chapter presents the review of the literature that this research was based on. First, an overall context related to pronunciation teaching is provided (section 2.1), followed by its relationship with technology (section 2.2). Then, we turn our focus to the speech technology that is the central part of this research: Automatic Speech Recognition (section 2.3). Finally, a summary of the chapter is provided for the reader (section 2.4).

2.1 PRONUNCIATION TEACHING

The term *pronunciation* captures "all aspects of how we employ speech sounds for communicating" (BURNS; SEIDLHOFER, 2020, p. 247). Therefore, it encompasses more than single sounds, comprehending different *features*: segments (consonants, vowels), suprasegmentals/ prosody (syllables, rhythms, and intonation), and connected speech phenomena (CARDOSO, 2017). Pronunciation is linked to language attitudes and identity issues for learners and teachers, which may impact teachers' confidence and, consequently, their willingness to teach pronunciation (PENNINGTON; ROGERSON-REVELL, 2019). Also, pronunciation is highly connected to social issues once it "cannot be ignored in view of globalization, the growing needs of cross-cultural communication, and issues of speaker identity and social integration" (CHUN, 2020, p. 213). In short, pronunciation is a key element aiming at successful spoken communication (PENNINGTON; ROGERSON-REVELL, 2019).

Moreover, speech is the main medium of language (SLABAKOVA, 2016). In the area of Second Language Acquisition (SLA), there are different dimensions of L2 speech to consider. Munro and Derwing (1995) define the terms *intelligibility*, *comprehensibility*, and *accentedness*. For the authors, intelligibility is "the extent to which a speaker's message is actually understood by a listener". Comprehensibility refers to the degree of difficulty in understanding speech based on the listener estimation. Lastly, accentedness is how different the speech is in relation to an expected production pattern. To put it succinctly, Derwing (2010, p. 29) simplifies these dimensions as "accent is difference, comprehensibility is effort, and intelligibility is actual understanding".

For some authors, L2 pronunciation teaching and research should focus on enhancing intelligibility and comprehensibility contrary to accent improvement or native-likeness (O'BRIEN et al., 2018). As Levis (2020) contends, pronunciation teaching that focus on intelligibility is consistent with realistic goals since accent should be seen as a variation to embrace instead of a problem to overcome. Nowadays, this focus on intelligibility and comprehensibility rather than accent improvement as the goals of L2 speakers has been accepted by most L2 teachers (DERWING, 2010). Focusing on intelligibility is also important for overall communication. Sicola and Darcy (2015, p. 471) state that "pronunciation difficulties in a second language (L2) can seriously impede intelligibility. […] Lack of intelligible pronunciation is also accompanied by comprehension difficulties when L2 learners listen to spoken English".

Regarding pronunciation teaching, it is important not only to help raise students' awareness regarding different sounds and sound features - which benefits their production and understanding of spoken English but also to improve their overall speaking skills (HARMER, 2015).

Notwithstanding, research shows that “teachers’ beliefs about pronunciation teaching may reflect a neglect of pronunciation in the classroom” (COSTA, 2016, p. 61). The author further claims that “pronunciation teaching is a major gap in teacher education” (COSTA, 2016, p. 61). In addition, the author indicates that there might be a lack of deep knowledge of pronunciation features (segments and suprasegments), which results in a didactical limitation. In a similar vein, Sicola and Darcy (2015, p. 472) state that "the first challenge is the lack of teacher training in pronunciation". In Costa’s (2016) study, the participant-teachers did not use a great variety of resources and/or techniques to teach pronunciation. Thus, besides the lack of training and knowledge about pronunciation features, there might be a lack of knowledge about techniques and/ or resources to teach pronunciation.

For example, in her research with five English teachers, Baker (2014) reports that most of the techniques used by the teachers to teach pronunciation were *controlled* ones (e.g., repetition practice, oral reading, tongue twister practice). Such exercises “allow students to develop skill in perception and/or production with a feature of pronunciation” (LANE; BROWN, 2010, p. 11). Baker (2014) also revealed the limited knowledge of the teachers regarding a variety of techniques, especially *guided/ semi-controlled* (e.g., information-gap exercises, cued dialogues, and strip stories). This limited repertoire is likely to overlook important phases during the acquisition of new pronunciation features. Nonetheless, Lane and Brown (2010) contend that:

> The fact that pronunciation gains in controlled activities may not carry over in communication does not mean that controlled activity have no value; on the contrary, they provide practice opportunities that can gradually lead to more automatic use of the new pronunciation as well as to skills for self-correcting. However, controlled activities should not be the end of the lesson. Our students are not studying English to become proficient to readers of word lists (p. 11 – 12).

It is important to stress that, with the research insights that indicate factors impacting learner’s intelligibility, a diverse range of innovative techniques can be embraced by L2 teachers, mirroring, and shadowing, for example, for practicing suprasegmentals with the objective of successful communication (BRINTON, 2017). Although innovative techniques are available to L2 teachers, there is also room for “old-fashioned” practices since “drills and other repetition exercises may be important to develop new motor skills, particularly regarding the articulation of new sounds” (ROGERSON-REVELL, 2011, p. 235).

In regard to pronunciation teaching techniques, Celce-Murcia et al (2010) suggest a diverse range of techniques and practice materials. A summary of those techniques is presented in Table 1. Considering all these resources and tips, the authors further state that pronunciation teaching should focus not only on isolated words or sentences but rather work on the discourse level as well. In addition, pronunciation teaching should be open to receiving insights from other disciplines like speech pathology, drama, and neurolinguistics.

Table 1 – Pronunciation teaching techniques for L2 teachers

Pronunciation Teaching Technique	Definition
1- Listen and imitate	Students listen to a teacher-provided model and repeat or imitate it.
2- Phonetic training	Use of articulatory descriptions, articulatory diagrams, and a phonetic alphabet.
3- Minimal-pair drills	Students should distinguish between similar and problematic sounds in the target language through listening discrimination and spoken practice. Minimal-pair drills typically begin with word-level drills and then move on to sentence-level drills.
4- Contextualized minimal pairs	In this technique, the teacher establishes the setting (e.g., a blacksmith shoeing a horse) and presents key vocabulary; students are then trained to respond to a sentence stem with the appropriate meaningful response (a or b).
5- Visual aids	Enhancement of the teacher's description of how sounds are produced by audiovisual aids such as sound-color charts, charts, rods, pictures, mirrors, etc.
6- Tongue twisters	A technique from speech correction strategies for native speakers (e.g., "She sells seashells by the seashore.").
7- Developmental approximation drills	Speakers are taught to retrace the steps that many English-speaking children follow as they acquire certain sounds in their first language.
8- Practice of vowel shifts and stress shifts related by affixation	To raise awareness, the teacher points out the rule-based nature of vowel and stress shifts in words related etymologically.
9- Reading aloud / recitation	Passages or scripts for learners to practice and then read aloud, focusing on stress, timing, and intonation.
10- Recordings of learners' production	Audio and video recordings of rehearsed and spontaneous speeches, free conversations, and role plays.

Source: Celce-Murcia et al (2010, p. 9 – 10)

Table 1 presents the wide variety of techniques that L2 teachers can employ while teaching pronunciation. However, the possibilities are not limited to this list. For example, *shadowing* can improve significantly different aspects of L2 learners' pronunciation (FOOTE; MCDONOUGH, 2017). This technique consists of the student repeating either along with or slightly after someone's

speech (input) (CELCE-MURCIA; BRINTON; GOODWIN, 2010). Moreover, Brinton (2017) indicates that such a technique can be an alternative resource for autonomous practice.

Regarding learning autonomy, Burns and Seidlhofer (2020, p. 255) advocate that teaching "with the aim of fostering learner autonomy and enabling students to develop strategies for coping on their own and for continuing to learn is perhaps the most valuable thing that can be developed in learners". In fact, considering the time constraints of the regular curriculum, extended practice outside the classroom may be necessary. For optimal results in learners' productions, Carlet and Kivistö-de Souza (2018) indicate that learners need to continue their learning process outside the classroom. They further defend the use of autonomous activities that focus on increasing learners' awareness about phonology. The authors explain that L2 phonological awareness "can be developed through any activity that brings a specific aspect into the language learners' consciousness" (CARLET; KIVISTÖ-DE SOUZA, 2018, p. 104). Likewise, Brinton (2017) points out that teachers encourage learners to be more autonomous by creating out-of-class learning opportunities.

As aforementioned, a lack of time dedicated to pronunciation teaching may be a constraint for successful pronunciation improvement. Lack of time due to an overloaded syllabus might be an issue for many English teachers (COLLINS; MUÑOZ, 2016; PENNINGTON; ROGERSON-REVELL, 2019). Time limitation may hinder the learning process because of the insufficient linguistic input received and fewer opportunities to produce output. Notwithstanding, time dedicated to pronunciation learning is particularly important once this component of L2 learning requires hours of input and output practice (EVERLY, 2019).

According to Roccamo (2014, p. 183), "many foreign language instructors recognize the importance of pronunciation skills, but do not teach them because they are worried about time constraints or are unsure of how to get started". The author also suggests a module-style pronunciation instruction that "teaches pronunciation in just ten minutes per class period" (p. 188). This idea of integrating pronunciation teaching in every lesson is also supported by Pennington and Rogerson-Revell (2019, p. 186) who advocate that "integrating pronunciation teaching into other areas of language learning is a way of reminding learners of its broader significance". The authors argue that some time should be invested in spreading awareness regarding pronunciation in the early stages of the language learning process. Addressing pronunciation issues in every lesson could be a solution to help learners improve them (SICOLA; DARCY, 2015). All in all, Rogerson-Revell (2011) suggests that teachers should prioritize pronunciation teaching because:

> For many students of English the real challenge to mastery of the language is not the grammar, not the vocabulary but the pronunciation. It is an area traditionally thought of as 'difficult' and frequently ignored by teachers and learners alike. Nevertheless, it leads to breakdowns in communication and once fossilized, poor pronunciation is immensely difficult to remedy. With the advent of more communicative methodologies, dedicated pronunciation coursebooks and interactive resources, a principled and engaging approach to teaching pronunciation can be adopted. The important thing is that pronunciation should not be relegated to a five minute slot at the end of a lesson but integrated into all aspects of teaching (p. 261).

In addition, communication is the primary purpose of language use. Hence, the main goal of L2 teaching should be using language to communicate (CELCE-MURCIA; BRINTON; GOODWIN, 2010). It is possible to establish, thus, that "the goal of pronunciation instruction should be to help learners realize their full communicative potential in second language acquisition"

(MUNRO, 2008, p. 213). Following this principle, Celce-Murcia et al (2010) proposed a framework for pronunciation teaching based on the tenets of communicative language teaching (CLT). Table 2 summarizes this framework which considers the learners' gradual development of pronunciation features from controlled to automatic knowledge of the L2 phonology (CELCE-MURCIA; BRINTON; GOODWIN, 2010).

Table 2 – Communicative framework for teaching English pronunciation

1- Description and analysis	Oral and written illustrations of how the feature is produced and when it occurs within spoken discourse.
2- Listening discrimination	Focused listening practice with feedback on learners' ability to correctly discriminate the feature.
3- Controlled practice	Oral reading of minimal-pair sentences, short dialogues, etc., with special attention paid to the highlighted feature in order to raise learner's consciousness.
4- Guided practice	Structured communication exercises, such as information-gap activities or cued dialogues, that enable the learner to monitor for the specified feature.
5- Communicative practice	Less structured, fluency-building activities (e.g., role play, problem solving) that require the learner to attend to both form and content of utterances.

Source: Celce-Murcia et al (2010, p. 45)

The above-mentioned framework suggests a 5-phase division of the pronunciation lesson. The first phase focuses on phonological awareness, the second focuses on perception skills while the three last phases predict the learner's production. The framework recognizes that pronunciation learning is not a linear process. Therefore, teachers might need to revisit certain phases during their teaching process. The first phase (description and analysis) refers to the introduction of new pronunciation features while calling the attention of the learner to the articulatory features associated with them. Diagrams and charts are useful tools during this stage of the pronunciation lesson. The second phase (listening discrimination) entails focused listening. In this phase, learners are asked to identify or distinguish a particular new feature from other similar features. Such activities allow learners to train their listening skills related to the target feature and raise their consciousness regarding its importance (CELCE-MURCIA; BRINTON; GOODWIN, 2010).

Moving to the third phase (controlled practice), it is when the learner's attention is directed to the production of specific pronunciation features. Controlled activities focus on the target-like production of a particular number of sound features. Guided practice (fourth phase), which may also be called structured or semicontrolled practice, provides much of the language to the learner as well as the learning context; however, learners have the possibility of expressing meaning by informing specific words or phrases. The fifth and last phase (communicative practice) refers to activities that require learners to use novel phonological features in genuine communicative language use. During the communicative practice, learners' attention should be directed toward the content of their message and the linguistic form (CELCE-MURCIA; BRINTON; GOODWIN, 2010). All in all, this communicative framework:

> […] recognizes the key role played by each phase in the acquisition of new pronunciation features. It also recognizes that learners' progression from one phase to another may be a gradual one and that the three phases of practice facilitate learners' movement from

> controlled to automatic processing/production of L2 phonology. It is important to note that the application of this framework should be viewed as extending over the course of several lessons rather than just one since it takes time and effort for learners to acquire a new feature and automatize it in their spoken production. […] Ultimately, the most important contribution of this framework is the recognition that practice must extend beyond the controlled phase of repetition or oral reading and instead, once learners have gained more control over the feature, extend to having them use the newly acquired feature in more creative or communicative exchanges (CELCE-MURCIA; BRINTON; GOODWIN, 2010, p. 45).

Besides the presented framework, the authors advocate in favor of setting realistic goals for pronunciation teaching, "focusing on the targeted communicative needs of their learners, not the language as a whole" (CELCE-MURCIA; BRINTON; GOODWIN, 2010, p. 283). Considering the context where this study was conducted (L1 Brazilian Portuguese speakers) and taking into consideration that L2 learners need intelligible speech rather than native-sounding speech (MUNRO, 2008), much research has been conducted to identify pronunciation issues that may affect intelligibility and cause communication breakdowns of Brazilian learners of L2 English (DELATORRE, FERNANDA; GONCALVES, ROBERTO; SILVEIRA, 2017; GONÇALVES; SILVEIRA, 2015; SILVEIRA et al., 2017). Thus, Brazilian teachers might be able to perform research-based decisions on which pronunciation features they intend to address. In addition, Zimmer et al (2009) offer a variety of pronunciation activities designed for Brazilian learners and explain the transfer processes from Brazilian Portuguese into American English.

Considering the arguments presented in this section, it seems clear that pronunciation teaching and learning is of pivotal importance when it comes to having an intelligible speech aiming at successful communication. As Pennington and Rogerson-Revell (2019, p. 219) contend, "having a critical awareness of pedagogic resources and related research can help teachers make informed choices about what and how to teach". Pronunciation learning is highly affected by learners' exposure to the L2 as well as the extent to which they actually use it (LANE; BROWN, 2010). Therefore, digital resources may help learners to have more L2 exposure and teachers to achieve their pedagogical goals especially considering "the notoriously short time for pronunciation instruction in language curriculums" (KIVISTÖ-DE SOUZA; GOTTARDI, 2022, p. 778). The next section will discuss the relationship between technology and pronunciation teaching.

2.2 TECHNOLOGY AND PRONUNCIATION TEACHING

Language classrooms without technology[3] creates an artificial and limited learning environment considering that it is so pervasive in human activities. Nowadays, digital technologies grant access for people to "speak or write either synchronously or asynchronously, with participants either at a distance or in close proximity" (CHUN; KERN; SMITH, 2016, p. 66). When it comes to L2 teaching, its use during the lessons does not guarantee that bad pedagogy becomes good (GOLONKA et al., 2014). In regard to pronunciation teaching more specifically, Munro and Derwing (2015, p. 393) caution that "the potential of technology for pronunciation instruction has yet to be effectively tapped". The use of digital technology in L2 classes should not be considered

[3] This study will refer to the term *technology* as *digital technology*. Digital technologies will be considered as "any materials created digitally or converted to digital format, or any platform, application or program available digitally" (TUMOLO; FINARDI, 2021, p. 11). Also, *digital resources* and *digital tools* will be used with no distinction.

the solution to all problems, but rather a way of supporting specific pedagogical goals (CHUN; SMITH; KERN, 2016). Moreover, digital technologies may allow learners to be more autonomous toward their learning process, besides promoting learning beyond the classroom (LAI, 2018). In a similar vein, Thomson and Derwing (2015, p. 336) contend that "computer based approaches may promote greater learner autonomy and, most importantly, afford the possibility of individualized instruction". Finally, technology can also offer "the learner limitless choice of what, when, and how to learn" (REVELL-ROGERSON, 2021, p. 201).

The complex nature of technology and pedagogy demanded the creation of a new subfield of applied linguistics called Computer-Assisted Language Learning (CALL), which studies the relationship between technology and Second Language Acquisition (SLA) (MARTINS; MOREIRA, 2012). As Davies (2006, p. 261) defines, CALL is "an approach to language teaching and learning in which computer technology is used as an aid to the presentation, reinforcement, and assessment of material to be learned". Historically, CALL has been influenced not only by SLA and language pedagogy but also by the evolution of computer technology (DAVIES; OTTO; RÜSCHOFF, 2013).

CALL can refer to different types of practice from websites to multimedia programs. These types of practice may include, for example, self-studying activities on a website, multimedia components to practice listening skills, computer-generated feedback on learners' production, and internet collaboration among L2 learners (SUVOROV; CHAPELLE, 2020). Thus, technology has transformed its auxiliary role in the curriculum into an indispensable source of authentic content and language learning opportunities (OTTO, 2017). Moreover, technology has taken a meaningful role in pronunciation teaching due to the great variety of input that is made available to the learners and immediate feedback (BALDISSERA; TUMOLO, 2021). For all these reasons, many teachers, parents, and learners see technology as a valuable learning ally, providing varied learning experiences (DUCATE; ARNOLD, 2006).

Considering the advances in technology and SLA research, CALL researchers have begun to explore other venues. For instance, Computer-Assisted Pronunciation Teaching (CAPT) investigates the use of digital resources' affordances to help learners develop their production and perception of rhythm and melody of a target L2 (CHUN, 2020). On another venue, Mobile Assisted Language Learning (MALL) emerged due to the evolution of technology to a mobile form. In MALL, the personal and portable devices allow learners to practice a target L2 at a suitable place and time for them (BALDISSERA; TUMOLO, 2021). Finally, CALL researchers also examine the effectiveness of other new technologies (e.g., virtual worlds for L2 learning, and digital game-based learning) (SUVOROV; CHAPELLE, 2020).

In recent decades, the field of CALL has advanced quickly (PENNINGTON; ROGERSON-REVELL, 2019) and it now embodies a broad array of practices, predicting varied pedagogical effectiveness and outcomes (CHAPELLE, 2013). Also, there are many digital resources to support the development of an L2 available to teachers, learners, and for autonomous learning (TUMOLO, 2014). More importantly, the use of technology with a clear purpose can provide teachers and learners with meaningful classroom experiences (PERNA; DELGADO; SILVA, 2022).

However, it is important to stress that "no discussion of the applications of new technologies to language learning would be complete without making reference to scenarios in which access to

these means are lacking" (BRETT; GONZÁLEZ-LLORET, 2009, p. 366). In addition, new digital resources are made available to teachers and learners all the time. Although more CALL resources might be a positive asset to them, they can also represent an extra challenge for them to keep posted in relation to the latest releases and updates (SOLEIMANI, 2021). Evaluating CALL software and the activities designed to use CALL resources is, hence, necessary. In fact, it is what "teachers need to concentrate more on, in order to ensure that technology is not used for technology's sake" (STANLEY, 2013, p. 9).

Regarding the evaluation of CALL learning activities, comprehending how appropriate the designed activity is can be challenging as many teachers might have limited experience with the use of technology for teaching (HUBBARD, 2006). In order to evaluate the appropriateness of CALL activities, Chapelle (2001) suggests a set of criteria. These criteria are displayed in Table 3.

Table 3 – The six criteria for CALL task appropriateness

1- Language learning potential	The degree of opportunity present for beneficial focus on form.
2- Learner fit	The amount of opportunity for engagement with language under appropriate conditions given learner characteristics.
3- Meaning focus	The extent to which learners' attention is directed toward the meaning of the language.
4- Authenticity	The degree of correspondence between the CALL activity and target language activities of interest to learners out of the classroom.
5- Positive impact	The positive effects of the CALL activity on those who participate in it.
6- Practicality	The adequacy of resources to support the use of the CALL Activity.

Source: Chapelle (2001, p. 55)

As stated by Chapelle (2001), these criteria can be applied to evaluating educational software and activities that teachers plan to utilize in their classes. From all the criteria afore-presented, *language learning potential* "should be considered the most critical for CALL activities" (CHAPELLE, 2001, p. 58); however, the importance of each criterion may differ according to the objective of the task. The first criterion (language learning potential) evaluates whether the activity is a proper language learning activity instead of a simple opportunity for language use, that is, the activity can promote a beneficial focus on linguistic forms. The second criterion (learner fit) refers to how appropriate the activity is for the learner, considering the difficulty level, and the learner's individual differences and characteristics.

The third criterion (meaning focus) evaluates if the activity directs learner's primary attention to the meaning of the language required to accomplish the task. This criterion also takes into account activities that involve reading and writing skills since the language is used with a clear purpose for interpreting and constructing meaning. The fourth criterion (authenticity) analyses whether the learning task is similar to a task that learners might encounter outside the L2 classroom. The fifth criterion (positive impact) evaluates if the learning task teaches more than language, that is, supports learners to develop or improve their metacognitive strategies, allowing them to claim agency over their learning process. The sixth and last criterion (practicality) evaluates whether the

learning task is easy for the teachers and learners to implement considering the particular constraints of an L2 class or the language program. In addition, it takes into account the availability of software, hardware, and knowledgeable personnel that is needed for the proper execution of the learning task (CHAPELLE, 2001).

All things considered, the benefits of current digital resources to pronunciation teaching are enormous (O'BRIEN et al., 2018). As Celce-Murcia et al (2010, p. 361) point out, "when combined with traditional classroom instruction, pronunciation technology may provide the key to individual improvement". Because of it, the authors further state that teachers should demand access to up-to-date technology from their institutions and programs. In short, CALL and the advances in computer technology have considerably contributed to the improvement of L2 teaching and learning (CARDOSO, 2022). Nonetheless, a particular speech technology, Automatic Speech Recognition (ASR), may be a suitable CALL resource for L2 teaching and learning - as it may fulfill most of the criteria proposed by Chapelle (2001) - especially for pronunciation practice, as we shall see in detail during the next section.

2.3 AUTOMATIC SPEECH RECOGNITION

Simply put, ASR is "an independent, machine-based process of decoding and transcribing oral speech" (LEVIS; SUVOROV, 2013, p. 316); that is, ASR can transcribe speech (oral input) into word sequences (written output) (YU; DENG, 2015). It can be embedded into smart devices containing Intelligent Personal Assistants (IPAs) (e.g., Siri, Alexa, Microsoft Cortana, Google Assistant), used to auto-generate captions, applied to human-computer interaction, or dictation tools (INCEOGLU; LIM; CHEN, 2020; JURAFSKY; MARTIN, 2021; MOUSSALLI; CARDOSO, 2020). The process of speech recognition, sometimes called speech-to-text, is "to map acoustic waveforms to sequences of graphemes" (JURAFSKY; MARTIN, 2021, p. 575). In addition to this definition, it is possible to comprehend speech recognition as "synonymous with intelligibility" (DIB, 2019, p. 63).

ASR technology is not recent once the first ASR systems started being developed in the early 1950s. The first ASR system was developed by Bell Laboratories in 1952 and was called *Audrey System*. This system could recognize any of the 10 digits from the speech of one single speaker (JURAFSKY; MARTIN, 2021). Then, ASR systems from the 60s were able to identify and respond to 16 words in English. In the 70s, this number increased to one thousand words. The accuracy rate, however, had considerably decreased. However, over the last decades, much progress has been made. By the early 2000s, the speech recognition accuracy rate reached over 80%, and, more recently, over 95% (CARDOSO, 2022).

Although this complex process still faces many challenges, ASR technology has improved to the extent that it is suitable for many tasks. This is due to the development of more robust algorithms and new model techniques, the expanding demand for ASR integration with smart devices, and the improvement in noisy speech recognition (JURAFSKY; MARTIN, 2021; LEVIS; SUVOROV, 2013). Furthermore, the accuracy rates of the ASR system are continually being increased by the use of machine learning combined with relevant acoustic information (ASHWELL; ELAM, 2017).

In general, ASR systems share some particular characteristics related to four dimensions of the speech recognition task. The first dimension is vocabulary size. A large vocabulary makes the task much harder. The second is regarding the speaker. The speech derived from Human-Machine Communication (HMC) is recognized more easily than Human-Human Communication (HHC) since the latter contains more lexical variation and a higher speech rate. The third is channel and noise. The task of speech recognition is easier in a quiet place with a proper input device (e.g., head-mounted microphones). The fourth and last dimension is related to the speaker's characteristics and accent. Depending on how the system was trained, the task may be more challenging. Speech recognition of children is usually harder to perform because most systems are trained on adult speech. Highly accented speech may also be troublesome. In addition, speakers talking to the ASR system using the same dialect or variety that it was trained on will make the task easier (JURAFSKY; MARTIN, 2021).

Bearing in mind these dimensions, some limitations of ASR systems might be inevitable despite the current technological advancements. Under specific circumstances, they still perform poorly and face some issues. The most problematic ones include spontaneous speech that might be disfluent, with variable speed, or with emotion; very noisy environment; side talks and multitalker speech; highly accented speech, and far-field microphones (YU; DENG, 2015). Moreover, a proper ASR system has to be capable of recognizing units (e.g., phonemes, words, phrases) accurately, and consider language complexity and syntactic and semantic ambiguity (LEVIS; SUVOROV, 2013). Hence, to evaluate the accuracy of ASR systems, there is a standard evaluation metric called Word Error Rate (WER) based on "how much the word string returned by the recognizer (the hypothesized word string) differs from a reference transcription" (JURAFSKY; MARTIN, 2021, p. 567).

The lower the WER is, the better. The WER will vary according to the task and context. For example, WER on read speech (LibriSpeech audiobook corpus) is around 2%. However, such a task requires massive computational resources. The WER is much higher when transcribing telephone conversations, 5.8% for Switchboard, and 11% for CALLHOME corpora. Transcribing speech from a 4-people dinner party can have a WER higher than 81% due to the context where the conversational task took place (JURAFSKY; MARTIN, 2021). In addition, another noteworthy limitation when using ASR systems to transcribe speech is the "unreliability of the internet connection, particularly during peak periods" (LIAKIN; CARDOSO; LIAKINA, 2015, p. 18). Bearing in mind all the characteristics and limitations of ASR systems, the next section discusses the applicability of ASR technology for pronunciation improvement.

2.3.1 ASR for pronunciation teaching

Much research has been made trying to comprehend the affordances of ASR for pronunciation teaching. The most common affordance is regarding learner autonomy. ASR can offer extensive output practice in different learning contexts, which may promote learner autonomy in and out of the classroom (KIM, 2006; LIAKIN; CARDOSO; LIAKINA, 2015, 2017; INCEOGLU; LIM; CHEN, 2020; MCCROCKLIN, 2018, 2014, 2016; MROZ, 2018; VAN LIESHOUT; CARDOSO, 2022). In addition, using ASR on mobile phones enables learners to practice their speech limitlessly (WALESIAK, 2021) with feedback provided by the app (BALDISSERA;

TUMOLO, 2021). This feedback generated by the ASR program can also facilitate the evaluation of pronunciation once ASR "provide a way of evaluating learners' output in an objective, immediate and automatic manner" (FOUZ-GONZÁLEZ, 2015, p. 323). Finally, ASR programs can be a functional tool to be used outside the classroom (INCEOGLU; LIM; CHEN, 2020; LIAKIN; CARDOSO; LIAKINA, 2015, 2017).

From a pedagogical point of view, Gottardi et al (2022, p. 6) suggest that "teachers and learners can focus on the strengths of this technology to fulfill specific learning goals". The authors also point out that this speech technology can offer learners extra opportunities to produce oral output hence practicing pronunciation and speaking skills. As a result, the potential of ASR mainly for pronunciation development is prodigious. Such a statement is supported by Liakin et al (2017, p. 14) as ASR provides "endless opportunities for output by allowing students to practice the target language anytime, anywhere, and at their own pace". This extended output production is also important for the development of morphosyntactic knowledge of the learner (CHAPELLE, 2003). Finally, ASR-based pronunciation practice can create a safe opportunity for learners to practice oral skills and develop their self-monitoring and analysis skills related to their pronunciation (WALLACE, 2016).

From the learners' point of view, research shows that learners demonstrated an overall positive attitude toward ASR for pronunciation practice (INCEOGLU; LIM; CHEN, 2020; KIM, 2006; MROZ, 2018; WALESIAK, 2021). There is also evidence that ASR-based pronunciation practice can reduce learners' foreign language anxiety and improve their confidence and willingness to communicate (CHEN, 2011; INCEOGLU; LIM; CHEN, 2020; KIM, 2006; MROZ, 2018). Furthermore, Liakin et.al (2017) report that learners enjoy using ASR due to the mobile-enhanced learning environment that this speech technology provides them with in addition to fostering their autonomy. Lastly, there is evidence of oral production improvement after ASR-based pronunciation practice (CUCCHIARINI; NERI; STRIK, 2009; NERI; CUCCHIARINI; STRIK, 2006, 2008).

Regarding ASR orthographic feedback, Liakina and Liakin (2022, p. 294) point out that "ASR is a very promising technology that should allow students to get immediate feedback on their pronunciation, thus making them more independent in learning this aspect". The authors further state that participants (fifty-seven adult L2 French learners) of the action research study indicated that the corrective feedback provided by the application (ASR included) was useful. Furthermore, research has proven that ASR-generated feedback can aid the improvement, mainly on the segmental level, of problematic speech sounds (NERI; CUCCHIARINI; STRIK, 2006, 2008). Nevertheless, Liakin et al (2017) reported that learners felt frustrated after not being able to comprehend why the ASR applications did not understand them and how to use ASR-generated feedback to improve their pronunciation. Hence, teachers still play a pivotal role while using ASR for pronunciation teaching (GOTTARDI; ALMEIDA; TUMOLO, 2022; KIVISTÖ-DE SOUZA; GOTTARDI, 2022; LIAKINA; LIAKIN, 2022; WALLACE, 2016).

Some studies also investigated the accuracy of ASR-based dictation tools applied for pronunciation practice. The programs under investigation obtained a high score considering learners with different native languages. For instance, the accuracy of Google Docs' voice typing feature for non-native speech (Spanish, Mandarin Chinese, Arabic, French, and Ewe) ranged from 88.61% (sentences) to 93.47% (free speech) (MCCROCKLIN; HUMAIDAN; EDALATISHAMS, 2019).

Likewise, Google's accuracy in a sentence-reading task was over 90% for all groups (English – 96.2%, Chinese – 90.9%, and Spanish – 92.73) (MCCROCKLIN; EDALATISHAMS, 2020). Finally, Brazilian Portuguese L1 speakers and Spanish L1 speakers were highly intelligible to two different ASR programs: VoiceNotebook, which uses Google's speech recognition technology, (80%), and Microsoft Word Dictation (89%) (KIVISTÖ-DE SOUZA; GOTTARDI, 2022). It is important to stress that, in the past, ASR could recognize nonnative speech with around 20% less accuracy if compared to native speech. Nonetheless, with ASR technology's recent improvements, Google Voice Typing could reduce this accuracy gap to 3-5% (MCCROCKLIN; EDALATISHAMS, 2020).

Furthermore, ASR seems to meet, to some extent, all the six criteria for CALL task appropriateness proposed by Chapelle (2001), which were presented in the previous section. ASR can be used to practice specific pronunciation patterns with instant orthographic feedback (language learning potential); it is flexible and can be adaptable to use with a wide range of text passages (learner fit); it can be used to improve speech intelligibility and work on pronunciation issues that might disrupt communication (meaning focus); it can transcribe different type of texts, including authentic passages from the internet or learner-made compositions (authenticity); it can improve learners' autonomy and have a positive impact towards their learning process (positive impact), and ASR can be found on different platforms, from operating systems (Windows Speech Recognition on Windows 10) and mobile phones (Google Keyboard on Android smartphones) to websites (Google Translate's ASR feature) (practicality). Likewise, van Lieshout and Cardoso (2022) state that Google Translator's speech features, which contain an embedded ASR tool, also meet most of the criteria defined by Chapelle (2001).

Going back to the communicative framework for teaching pronunciation proposed by Celce-Murcia et al (2010), presented in section 2.1, ASR-based pronunciation practice can be used for phases three to five (controlled practice, guided practice, and communicative practice). Those phases predict the oral production of the target pronunciation features to be improved. Therefore, learners could use ASR-based dictation tools to practice their oral production by dictating to the ASR program. Moreover, as Celce-Murcia et al (2010, p. 48) caution, "providing systematic feedback to the learner is essential in all the above phases of the lesson". ASR's transcription can be used as instant orthographic feedback during the activity or as delayed feedback after the whole practice. Hence, learners can receive pronunciation feedback provided by the ASR, not being restricted to receiving it from the teacher or other students. In summary, effective pronunciation teaching may take advantage of technology with a clear pedagogical purpose (CELCE-MURCIA; BRINTON; SNOW, 2014). ASR can be a powerful ally to this end.

Notwithstanding, ASR should be used as a complement to in-class pronunciation instruction and, by any means, substitutes for pronunciation instruction provided by a teacher (CUCCHIARINI; STRIK, 2018). In a similar vein, Kivistö-De Souza and Gottardi (2022, p. 777) advocate that teacher's guidance is necessary to "circumvent the limitations of the programs, tailor the passages according to the learner's needs and proficiency level and provide appropriate feedback on the intelligibility breakdowns indicated by the programs". In short, ASR can be a useful teacher and learner's ally. However, the use of digital resources is not a replacement for the teacher. It is rather, "a way of giving learners the opportunity to focus on their specific difficulties and receive

personalized feedback while becoming more autonomous in their learning process" (GOTTARDI; ALMEIDA; TUMOLO, 2022, p. 11). The next chapter will present the method that this research followed, giving special attention to the instruments designed during this study, which include practical suggestions on how to use ASR in L2 English classes.

3 METHOD

The objective of this section is to present the research design and the methodology of this study. First, the characterization of the research (section 3.1) presents the method design of this research, followed by the description of the participants (section 3.2) and the submission of this research to the ethics review board (section 3.3). Then, the instruments and procedures for data collection (section 3.4) are described, starting with the background questionnaire (section 3.4.1), followed by the survey (section 3.4.2), the workshop design (section 3.4.3) which presents the workshops' program and the criteria for the selection of the ASR program, and the seven ASR-based pronunciation activities (section 3.4.4). The last section (section 3.5) presents the procedures for data analyses.

3.1 CHARACTERIZATION OF THE RESEARCH

This research followed a mixed-method approach. This approach "involves collecting and integrating quantitative and qualitative data in a single project and therefore may result in a more comprehensive understanding of the phenomenon under investigation" (LEAVY, 2017, p. 164). Moreover, considering the chronological framework proposed by Cardoso (2022) for CALL-based speaking research (see Table 4), this research focused on the *exploration* (stage 2) of ASR as a teaching tool, that is, the examination of ASR's pedagogical affordances; and on the *assessing suitability* (stage 3), which investigates the acceptance, usability, and user's perceptions of this technology (from the teacher's perspective). Thus, to answer the two research questions guiding this study, it was necessary to combine the two above-mentioned stages of the four-stage framework presented below.

Table 4 – Summary of the chronological framework for CALL-based speaking research

Stage	Focus
1- Development	Development of a new resource taking into account insights from CALL, computer sciences, and second language acquisition.
2- Exploration	Examination of a novel technology's potential and its pedagogical affordances; tool's development follow-up; existing technology investigation.
3- Assessing suitability	Research on acceptance, usability, and learner's attitudes about a technology; pedagogical suitability investigation.
4- Assessing pedagogical effectiveness	Assessment of the pedagogical effectiveness of a target technology.

Source: author's summary based on Cardoso (2022, p. 307).

In order to answer the research questions, an online workshop session was designed and delivered by this researcher presenting an overview of the literature on pronunciation instruction and ASR, and seven original ASR-based pronunciation activities (presented on section 3.4.4). It was necessary to conduct a review of literature focusing on pronunciation teaching, CALL, and ASR-based pronunciation practice (*exploration* stage) to support the creation of the instruments presented

to the participants during the workshop, especially the ASR-based pronunciation activities and the guidelines for using the activities. In addition, a review of the literature allows the researcher to cover in-depth the investigated phenomena, analyzing previous scientific research and books (GIL, 2002). This was particularly important to understand the technology's architecture and limitations from a more technical point of view.

Furthermore, the background information of the participants was collected via an online background questionnaire[4] (presented on section 3.4.1). Participants' perceptions of the affordances and limitations of the ASR technology as well as their appraisal of the ASR-based pronunciation activities (*assessing suitability* stage) were collected via an online survey[5] (presented on section 3.4.2). Survey is a common research method to investigate quantitative research questions (CARDOSO, 2022). Hence, data from the background questionnaire and survey were mostly analyzed quantitatively except for one open-ended question from the survey. Moreover, a survey research design was chosen because "surveys are typically used for ascertaining individuals' attitudes, beliefs, opinions, or their reporting of their experiences and/or behaviors" (LEAVY, 2017, p. 269).

Regarding the instruments and procedures for data collection, a pilot study was conducted to assess the applicability of the instruments and the materials and activities designed for the workshop as well as the workshop duration. Some English teacher colleagues and PPGI (*Programa de Pós-Graduação em Inglês*) students were invited to be the participants of this pilot study and provide this researcher with feedback. After the pilot study, some changes were accommodated into the workshop design which will be described in section 3.4.3. Once the instruments for data collection did not differ significantly between the pilot study and the main study, participants' data from the pilot study was also used during the data analysis to enrich the discussion and support the answers to the research questions. The description of the participants from both groups, pilot study – from now on referred to as PS, and main study – from now on referred to as MS, are described in the following section. Finally, both workshop sessions were recorded and transcribed, and the chat log saved. These resources were used for qualitative analysis.

3.2 PARTICIPANTS

The participants were in-service English teachers from both public and private sectors. As requirements, they should be working as an English teacher by the moment of the workshop session and be a legal adult. Participants were teachers who work at public or private regular schools, language institutes, or work autonomously with groups and private students.

Participants were recruited online by sharing a poster containing the details of the workshop session and the link to the registration form. The poster was shared two weeks before the workshop session. Besides the background questionnaire, the registration form also contained an online term of consent for the participants as UFSC's Ethics Review Board requires (see section 3.3). The poster

[4] In this study, a *questionnaire* is "a list of questions that several people are asked so that information can be collected about something" (QUESTIONNAIRE, 2022). Thus, the background questionnaire was not designed to collect participant's opinions, only their personal information.

[5] In this study, a *survey* is "a set of questions people are asked to gather information or find out their opinions" (SURVEY, 2022). Therefore, the survey was designed to collect participant's information and their opinions and perceptions in relation to the workshop session and the pronunciation activities.

related to the PS was shared with PPGI's students via Moodle and sent to some English teachers known by this researcher that matched the participant's criteria. Participants from the MS were recruited via social media (Instagram, WhatsApp groups, LinkedIn, and Facebook) of this researcher, PPGI's Instagram and Facebook page, and via NUPFFALE's *(Núcleo de Pesquisa em Fonética e Fonologia Aplicada à Língua Estrangeira*) email group[6]. After answering the registration form and accepting the term of consent, participants received by e-mail (gathered by the background questionnaire) the link to the workshop session and further details on how to access the web conference platform.

Participants were invited to a single online workshop session. The background information of each participant was collected via a background questionnaire which was part of the registration form. Data from the two groups was combined for analysis, PS (*n=5*) and MS (*n=7*), resulting in a total of 12 participants (see Table 5). Both groups meet the same abovementioned criteria. Considering the PS, 6 teachers answered the registration form, but 5 of them joined the workshop session and answered the survey. For the MS, 38 teachers answered the registration form, but only 7 joined the workshop session and answered the survey. Therefore, only data from the participants who have joined the workshop and answered the survey was considered for analysis (*n=12*).

Table 5 – Participant's background information

Group	Code Name	Age	Gender	Teaching Experience*	Sector**
PS N = 5	P1	33	Male	11	Public
	P2	25	Female	8	Private
	P3	34	Female	15	Public
	P4	24	Female	7	Private
	P5	31	Male	12	Private
MS N = 7	P6	36	Female	15	Public
	P7	32	Female	14	Both
	P8	26	Male	6	Private
	P9	42	Male	4	Private
	P10	29	Male	10	Private
	P11	40	Female	16	Both
	P12	34	Female	10	Public

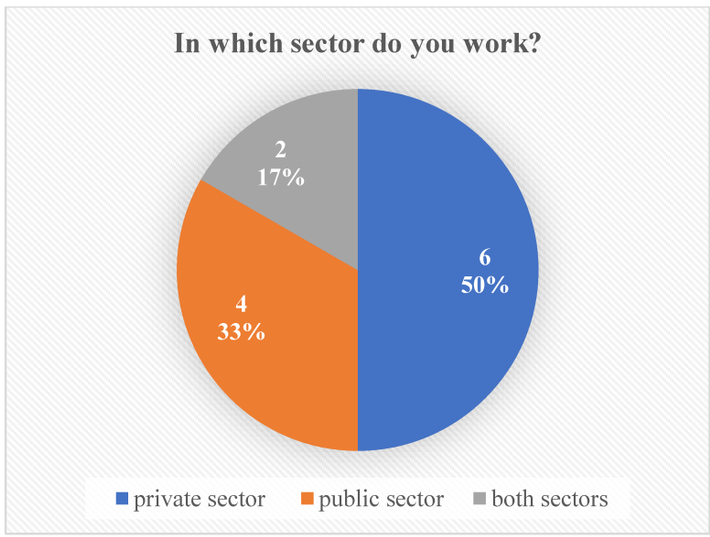
Source: author

*Years of experience teaching English. **Sector where the participant works (public, private, both).

Participants' age ranged from 24 to 42 (*M* = 32.2; *SD* = 5.6); 41.6% male (*n=5*) and 58.3% female (*n=7*). Regarding their experience teaching English, in years, the values ranged from 4 to 16 (*M* = 10.6; *SD* = 3.89). Considering the sector where the participants work, 50% work in the private sector (*n=6*); 33.3% in the public sector (*n=4*), and 16.6% in both sectors (*n=2*), as can be seen in Figure 1.

[6] nupffalenews@googlegroups.com

Figure 1 – Participants' work sector

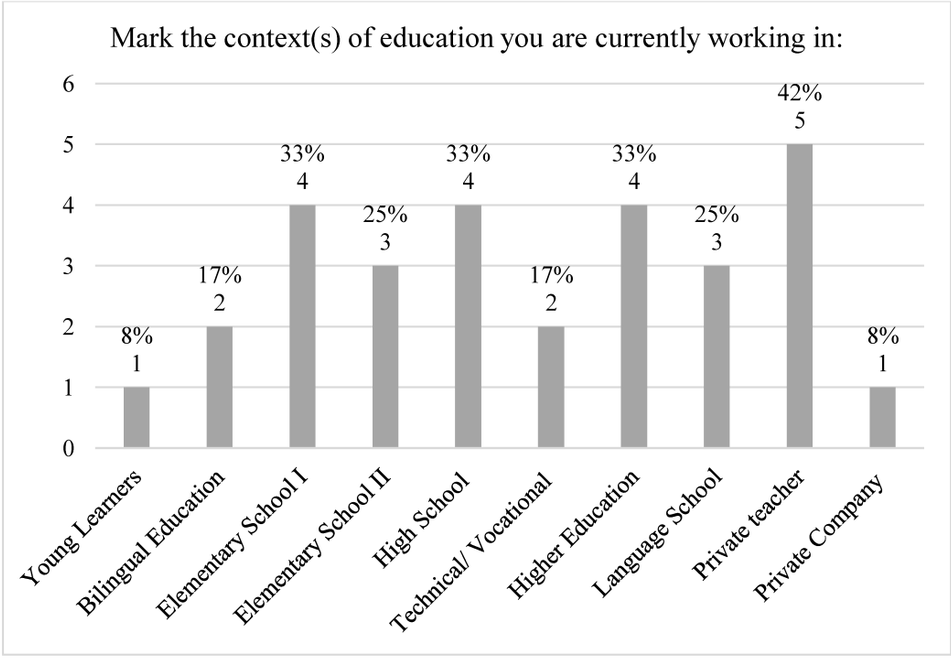


Source: author

All participants have an undergraduate degree; 83.3% (*n=10*) related to English teaching, and the other 16.6% (*n=2*) in a different area. In addition, the great majority of the participants have or are pursuing some sort of graduate degree in Linguistic, Literature or Education (specialization (*n=3*), master (*n=4*), and doctoral degree (*n=4*)). Considering the place where they currently live, most participants answered Brazil (*n=9*). The other countries cited were Italy (*n=1*), Poland (*n=1*), and Argentina (*n=1*).

When participants were asked to mark the context(s) of education they are currently working in, most of them work in more than one context of education, as seen in Figure 2. Special attention should be given to P6 and P10. P6 is currently working in 5 different contexts of education (Elementary School I, High School, Higher Education, Language School, Private teacher), and P10, in 6 different contexts of education (Elementary School I, Elementary School II, High School, Higher Education, Language School, Private teacher).

Figure 2 – Contexts of education participants are currently working in



Source: author

At least one participant marked one of the 10 possible alternatives, which makes this group of teachers heterogeneous. Table 6 displays the number and percentage of the options marked. The table is sorted from the largest to the smallest. The most common context of education was “Private teacher” (42%). The last common option was “Young Learners” and “Private Company (other)” (8%).

Table 6 - Summary of participant’s answers regarding the context of education they are currently working in

Item	**Total (*n*=12)**
Private teacher	5 (42%)
Elementary School I	4 (33%)
High School	4 (33%)
Higher Education	4 (33%)
Elementary School II	3 (25%)
Language School	3 (25%)
Technical/ Vocational	2 (17%)
Bilingual Education	2 (17%)
Young Learners	1 (8%)
Private Company (other)	1 (8%)

Source: author

All in all, all the participants who joined the workshop session and answered the background questionnaire and the survey matched the participant selection criteria. From the respondents, no missing data were found. Next section describes the ethical measures taken to protect participant’s personal data and integrity.

3.3 ETHICS REVIEW BOARD

Since this study involved human participants, this research was sent to the Institutional Review Board for Human Participants from the Federal University of Santa Catarina (CEPSH-UFSC) for approval[7]. After the research had been authorized, participants from the PS and the MS were recruited following the aforementioned procedures. In order to join the workshop session, participants had to sign a term of consent (Appendix A) which guarantees protection to the participants’ identity and privacy. The term of consent also asked participants for permission to record the workshop session. Thus, the recording could be used by the researcher to supplement the qualitative analyses. All the participants contributed to this study on a voluntary basis.

3.4 INSTRUMENTS AND PROCEDURES FOR DATA COLLECTION

In this section, all the instruments and procedures for data collection are described. First, the background questionnaire used for gathering the participants’ information and perceptions about

[7] The project was approved on September 29th, 2022, under the CAAE (*Certificado de Apresentação de Apreciação Ética*) 60910022.0.0000.0121.

digital technology and pronunciation teaching is presented. Then, the survey to collect the participants' perceptions after attending the workshop session is described. The following section presents the workshop design and the ASR program that was used during the workshop session. Finally, the last section shows the rationale behind the seven ASR-based pronunciation activities design.

3.4.1 Background questionnaire

A background questionnaire (see Appendix B) was shared online with the participants to support the data analyses and to guarantee that participants matched the criteria of this research. The questionnaire was designed to obtain *demographic characteristics* such as gender, age, residential location, occupation, education, and other background information (SUE; RITTER, 2007). There were also some behavioral questions to find out more about the participants' experience with digital technologies and pronunciation teaching. Such type of questions is used to understand who the respondents are, what they have done in the past, and what they are currently doing concerning the research topic (DÖRNYEI; TAGUCHI, 2009) - in this case, pronunciation teaching and the use of digital technology.

The questionnaire was created using Google Forms. The form contained the term of consent, and the background questionnaire. The form was available two weeks before each workshop session following the procedures described in section 3.2. In addition, the form allowed the participant to formally enroll in the workshop and, after its completion, the link to the online meeting was made available to the participants by e-mail. Thus, answering this questionnaire and accepting the term of consent were requirements for attending the workshop session and being part of the research. This first form was set up to stop accepting responses the day before the workshop session at 23:59 p.m.

The background questionnaire was divided into two sections. Participants were allowed to answer the questionnaire in Portuguese or in English. All fields were required. Before starting the first section, a welcome message was provided to the participants, containing further instruction on how to fill in the form and information related to the workshop session date. It also provided the participants with this researcher's e-mail address in case of any questions. The first section collected personal information. The second section contained behavioral questions to find out more about participant's experience with digital technologies and pronunciation teaching. All the items of the second section followed a 10-point Likert-scale. The items were designed to help understand the participant's behavior regarding this research topic and prepare better instructions during the workshop session - in case of further assistance with either digital resources to be used or pronunciation teaching practices suggested. The only difference from the PS background questionnaire in relation to the MS is the term "technological" which was changed to "digital" in order to improve clarity.

3.4.2 Survey

An online survey (see Appendix C) was shared with the participants after the workshop session in order to answer the research questions. This online survey was created using Google

Forms. The link to the survey was shared with the participants at the end of the workshop session after they had received further instructions on how to answer the form. The form was set up to stop accepting responses two weeks after each workshop session, PS and MS. All participants who joined the workshop session until the end answered the form. Participants were allowed to answer the survey in Portuguese or in English. The survey was divided into four sections. There was a welcome message and further instruction on how to complete the form before starting the first section. The survey was designed to be answered in around 30 minutes. At the end of the survey, there was a thank-you note which also contained this researcher's contact information.

No major changes happened in the survey administered to the PS the MS groups. Some words were changed to improve lexical coherence and clarity (e.g.: "technological" to "digital"). In addition, the PS survey contained one extra open question to collect feedback regarding the appropriateness of the instructions and the time to answer the survey properly. Participants from the PS answered that the survey contained proper instructions, the directions were clear, and that 30 minutes were enough to answer everything carefully. Therefore, no further changes were made, based on PS participants' feedback.

As Leavy (2017, p. 102) states, "survey items (questions in the questionnaire) are designed to help you test your hypotheses or answer your research questions". Thus, every survey item was carefully designed to address a particular research question. The items were divided into four sections, each of them aiming to answer a specific RQ or related construct. The items were designed to be "self-explanatory, easy to understand and answer, free of jargon, and visually appealing" (SUE; RITTER, 2007, p. 39).

Section 1, 2, and 3 of the survey address RQ1. Each section addresses one particular construct from RQ1: Teacher Development Needs (section 1), ASR Accessibility (section 2), ASR Affordances (section 3). These sections have the same names as the RQ1's constructs to make data analyses clearer and more straightforward. Section 4 addresses RQ2 solely. The number of the question resets in the beginning of each section. All Likert-scale questions followed a 10-point scale ranging from 0 (strongly disagree) to 10 (strongly agree).

The first section of the survey inquired about *Teacher Development Needs*. The questions addressed participants' needs for further education and professional development. Question 1 permitted selecting multiple items, including the option "Other" which allowed an open answer. Questions 2 and 3 included the 10-point Likert-scale. The questions from the second section, *ASR Accessibility*, asked about the suitability and accessibility of the presented ASR tool (GT) considering the participants' teaching context. Questions 1 and 2 followed the 10-point Likert-scale. Question 3 and 4 were a multiple-choice question, but only question 4 included the option "Other". The third section inquired about the *ASR Affordances*. The questions addressed the affordances of ASR technology for pronunciation teaching and learning, more specifically. Questions 1 to 5 included the 10-point Likert-scale. Question 6 permitted selecting multiple items, including the option "Other".

Section 4 collected information about RQ2. The questions of this section inquired about the design, suitability, and adaptability of the ASR-based pronunciation activities presented during the workshop session. A link to the Teacher's handout file, which contained the ASR-based activities, was made available to the participants on the description of this section. Participants were advised

to revisit the activities before answering the questions. The survey items of this section were designed based on Jamieson and Chapelle (2010) research. As the authors explain, the survey items were "logically linked to the six criteria for CALL evaluation" (p. 362) from Chapelle (2001). Therefore, for each one of the six Criteria for CALL Task Appropriateness, there were two items, except for the criterion "Practicality" which contained three items addressing it. This was necessary to disambiguate the terms "equipment" (e.g., microphone, earphone, computer, laptop, cellphone) and "infrastructure" (e.g., internet, wi-fi, computer lab) when participants were asked about how adaptable the activities were considering their teaching context.

Hence, a total of thirteen 10-point Likert-scale questions composed the fourth section. All these questions started with "the activities […]" to call attention to the set of activities instead of a single activity or the ASR technology itself. In addition, two more close-ended multiple-choice questions (question 14 and 15) inquired about which activities presented during the workshop the participant would use with their students (question 14), and among all the proposed activities which one was their favorite (question 15).

Finally, after the thank you-note, there was one final open-ended question that inquired about the workshop session itself (topic division, the way it was conducted, clarity, program). This was the only non-mandatory question of the survey, and there was no limit of characters or words for the answer. We now turn our focus on the workshop design. The next section shows it in detail followed by the description of the ASR-based pronunciation activities.

3.4.3 Workshop

There were two workshop sessions in total. The first session was delivered to the PS group on October 22, 2022. A total of 5 participants joined this session. The second session was delivered to the MS group two weeks after the PS group, on November 5, 2022. A total of 7 participants joined this session. Both sessions were offered online via a web conference platform, Zoom. They were delivered on a Saturday morning as a way of allowing teachers who work during the weekdays to attend the workshop session.

Some changes occurred from the PS session to the MS session. The main change was the duration of the session. The PS session was 2 hours long and the MS session was 3 hours long. These changes were made due to participants' feedback from the PS group who asked for more time for discussion and interaction. Nonetheless, the same materials were used for both groups and the same content was shared. Also, the PowerPoint slides used during the presentation were the same (Appendix G). Both sessions followed a specific program that can be observed in Table 7. The differences in duration between the sessions are underscored. For a more detailed version of the programs see Appendix D.

Table 7 – Workshop's Program Overview

Topic	**Duration (min) PS Group**	**Duration (min) MS Group**
Introduction	<u>10</u>	<u>15</u>
Review of the literature	15	15
Using Google Translate's ASR tool	10	10

Hands-on: ASR-based pronunciation activities	40	70
Coffee break	5	15
Brainstorming: ASR for pronunciation teaching	30	45
Discussion and final remarks	10	10
Total	**120 (2 hours)**	**180 (3 hours)**

Source: author

During the introduction, the research objectives and an overview of the workshop session were presented, and the Teacher Handout file was shared (see Appendix F) with the participants. This file contained important links (e.g., the link to the survey, websites suggestions, etc.), instructions on how to use Google Translate's ASR feature (from now on referred as GT), all the seven ASR-based pronunciation activities, directions on how to complete the survey after the workshop session, and further reading suggestions on the topics covered during the workshop session. This file was shared with the participants via Google Drive to help them follow the proposed activities during the session. Participants were invited to download it for personal use at their convenience.

The review of literature was identical in both PS and MS sessions. An overview of relevant topics was provided to the participants mainly related to pronunciation teaching, CALL, ASR technology, and the rationale behind the designed ASR-based pronunciation activities. Then, instructions on how to use GT's ASR feature were provided to the participants. They were able to use the tool on their own and to check if any technical problem occurred. For both groups, PS and MS, no technical problems happened, and participants confirmed that they could use GT.

After being able to use GT, participants were asked to perform the ASR-based pronunciation activities during the hands-on part. The activities were the focus of the workshop. Therefore, more time was invested in this part. Even so, the dedicated time to perform all the seven activities was not enough due to time constraint. For the PS session, activities 1, 2 and 3 were presented and performed by all participants. Activities 4 to 7 were only presented and discussed. Once the procedures for activity 3 (shadow reading) took a lot of time during the PS session, it was decided to make a change for the MS session. Thus, activities 1, 2, 4, and 5 were presented and performed by all participants while activities 3, 6 and 7 were only presented and discussed. Nevertheless, all the activities were presented to both groups, PS and MS, and all participants had access to the procedures and instructions on the Teachers Handout file. In regard to the hands-on part, the MS session was 30 minutes longer; hence, the MS workshop allowed more time to present all the activities at a slower pace with more time for discussion than in the PS workshop.

After the hands-on part, there was a time dedicated to a coffee break. Once the MS session was longer, the coffee break was longer as well. Then, participants were asked to brainstorm and discuss the content they had seen so far. First, they could debate with the whole group and share their perceptions. Second, they were divided into a breakout room so they could discuss among them only, without the researcher. The following questions were posed to facilitate the discussion on the breakout room: a) *Discuss the pros and cons of ASR for pronunciation teaching*; b) *Which one of the presented activities is more likely to be used? And less likely? Why?*; and c) *What pronunciation issues can be addressed using ASR?*. This procedure was planned to set a friendly environment for

discussion without the researcher's interference and possible biases. As both PS and MS groups had few participants, only one breakout room was needed. Finally, participants returned to the main room and shared their opinions with the researcher.

After the brainstorming, there were 10 minutes left to revisit the main topics covered throughout the workshop. Participants were able to ask questions and make comments. Then, instructions on how to complete the survey that would be shared with them right after the workshop session were provided. These last procedures were performed equally on both sessions.

Regarding the ASR tool used during the workshop, there were several reasons for choosing Google Translate (GT): 1) it is a free software available as a website or a mobile app, making this a highly accessible tool; 2) it is constantly improving and updating its speech features (VAN LIESHOUT; CARDOSO, 2022); 3) it has been on the market for more than a decade offering its services to more than five hundred million users. It supports more than one hundred languages and Brazil has used GT more than any other country (GOOGLE, 2016); and 4) Brazilian teachers might be more familiarized with Google products for English teaching and learning if compared to other digital tools (GOMES JUNIOR; SILVA; PAIVA, 2022).

All in all, both workshop sessions were recorded to support the data analysis and discussions. Participants were able to make comments in writing (via chat) or orally (via microphone) during the sessions. Their chat comments were also saved in a separate file, so they would be available for the qualitative analyses to help understand participants' perceptions. All things considered; we now turn our focus to the locus of the workshop: the pronunciation activities.

3.4.4 ASR-based pronunciation activities

Seven ASR-based pronunciation activities were designed by this researcher to be presented to the participants during the workshop (see Appendix E). The activities follow the *Criteria for CALL Task Appropriateness* from Chapelle (2001), from now on referred to as CCTA, and address some phases of the *Communicative Framework for Teaching Pronunciation* from Celce-Murcia, Brinton, and Goodwin (2010), from now on referred to as CFTP. Different teaching techniques were used in each activity to progress from controlled practice (3rd stage) to communicative practice (5th stage) (CELCE-MURCIA; BRINTON; GOODWIN, 2010). The pronunciation target features that each activity addresses were based on common intelligibility issues presented on L2 Brazilian learners' speech (GONÇALVES; SILVEIRA, 2015; SILVEIRA et al., 2017) and on different learning difficulties faced by L2 Brazilian learners' (ZIMMER; SILVEIRA; ALVES, 2009). All the activities follow the same structure, which was adapted from Stanley (2013) and can be seen on Table 8.

Table 8 – Activities' structure

Learning Focus	
Level	
Time	
Target Feature	
Communicative Framework Stage	

Dictation Passage	
Procedure [...] **Variation #** [...]	

Source: author

The “Learning Focus” row indicates the main learning goal of the activity. The “Level” row indicates the Common European Framework of Reference for Languages (CEFR) proficiency level for which the activity may be suitable. The “Target Feature” row is related to the pronunciation feature(s) addressed by the activity. The “Communicative Framework Stage” row indicates in which stage of the CFTP the activity is categorized into. The “Dictation Passage” row indicates the text, sentences, or worksheet to be used during the ASR dictation activity. The procedures that teachers should follow during the lesson are described step by step at the bottom of the table. Finally, each activity offers different variations that could be used by the teacher to adapt the activity according to his or her environment, context, or teaching objectives. Table 9 summarizes all the activities by presenting the expected proficiency level of the learners, the CFTP stage and the learning focus of the activity.

Table 9 – Activities summary

Activity Name	**CEFR Level**	**CFTP Stage**	**Learning Focus**
1 – Pronunciation Self-assessment	A1 – A2	Guided Practice	Pronunciation self-assessment
2 – Vowel Contrast	A2 and above	Controlled Practice	Raising awareness about acoustic features of each vowel and how to distinguish them in production
3 – Shadow Reading	B2 and above	Controlled Practice	Practicing speech fluency and accuracy
4 – Paragraph-reading Task	B1 – B2	Controlled Practice	Perceiving and producing the different -ed pronunciations of regular verbs in the simple past tense.
5 – Tongue Twisters	All levels	Controlled Practice	Pronunciation accuracy and fluency
6 – Monitoring Worksheet	All level	Controlled Practice	Pronunciation self-assessment and monitoring skills development
7 – Role-play Activity	B2 and above	Communicative Practice	Speech rehearsal for fluency and accuracy improvement

Source: author

As Table 9 demonstrates, most of the activities are categorized as “Controlled Practiced”. This decision was made as an attempt to overcome the limitations of the ASR technology as the literature suggests (ASHWELL; ELAM, 2017; GOTTARDI; ALMEIDA; TUMOLO, 2022; KIVISTÖ-DE SOUZA; GOTTARDI, 2022). However, the worksheet attached to Activity 1 (*Guided Practice* stage) can be personalized and adapted for different topics that require gathering learners’ personal information. In addition, activity 7 (*Communicative Practice* stage) suggests, as

a variation, the possibility of adapting varied role-play activities following the same procedures of the job interview.

In regard to the CCTA, the first criterion (Language Learning Potential) was prioritized while designing the activities, as suggested by Chapelle (2001). This criterion evaluates if the tasks "present sufficient opportunity for beneficial focus on form" (CHAPELLE, 2001, p.59). All the activities have a clear target pronunciation feature that might disrupt communication. Some of them focus on raising awareness about pronunciation patterns and acoustic features, while others, on improving the productions of segmental features (vowel contrast (/ ɪ / and / iː/), final clusters (-ed endings), and word-initial (/r/ and /h/)). Activity 3 focuses on suprasegmental features and the connected speech phenomena. In addition, procedures for giving explicit instruction and feedback were provided.

The second criterion (Learner Fit) evaluates the difficulty level of the activity and how appropriate the activity is, considering learner's characteristics (CHAPELLE, 2001). Every activity was designed to aim at a specific CEFR proficiency level range. Some of them (activity 5 and 6) can be used for all levels by adapting the target dictation passage accordingly, and some activities allow students to adapt the target dictation passage to their own difficulties. Also, one advantage of ASR is that learners can practice at their own pace with instant orthographic feedback; hence, students are allowed to increase or decrease their speech rate during most of the activities. Finally, all activities present variations that allow teachers to adapt the procedures according to their environment and learners' individual characteristics.

Moving to the third criterion (Meaning Focus), it evaluates if "learner's primary attention is directed toward the meaning of the language that is required to accomplish the task" (CHAPELLE, 2001, p. 56). Although the meaning might not be the primary focus of some activities, all activities were designed to improve speech intelligibility aiming at successful communication. Furthermore, CFTR is a 5-staged framework. Thus, progression is needed in order to achieve effective communication, starting from stage one to five (CELCE-MURCIA; BRINTON; GOODWIN, 2010). For example, some activities focus on delivering the intended oral message considering an ambiguous textual context (activity 2), fluency and appropriate intonation (activity 3), and a message in a specific tense (activity 4). Others ask learners to give personal information (activity 1) and to use it in a job interview (activity 7). All in all, most activities provide learners with opportunities to improve the articulatory movement needed for oral communication, speech correction strategies that can help overall communication, and the extended practice of words and phrases that may cause intelligibility issues and hinder communication.

The fourth criterion (Authenticity) evaluates how learners could profit from the activities outside the classroom; that is, if the CALL activity is connected to what learner would expect to see out of out of the classroom. First, all the activities offer extended opportunity for output practice; hence, more opportunity to use the language itself. In addition, the activities' procedures suggest pair work or group work, offering an extra interactive element. As Chapelle (2001, p. 56) states, "the criterion of authenticity indicates the need to develop learners' willingness to communicate, but it also extends beyond the conditions believed important for acquisition"; therefore, activities should offer opportunity to interact with other language users and to generate more spoken output. The activities focus on the practice of common words and sentences of everyday use, including

swearwords, which is a socially significant pronunciation feature once it may create social difficulties (LEVIS, 2018). Moreover, being intelligible while reading or speaking any text passage or sentence out loud is what one is expected to be during out of the classroom spoken interactions. Finally, activity 7 is a role-play activity that simulates a real-life interaction between two language users. This last activity also provides learners with the opportunity to create strategies for their own language development that can be applied to other tasks, which is connected to the fifth criterion (Positive Impact), which we now turn our focus on.

The fifth criterion (Positive Impact) evaluates the activity's learning outcomes regarding the target language itself and learning strategies. It also evaluates if both teacher and learners might have a positive learning experience during the activity (CHAPELLE, 2001). First, regarding the learning experience, the activities were designed to offer a degree of adaptability for each teacher according to his/her context and preferences. Also, variations were offered in each activity as suggestions for the teachers and learners in order to offer a more appropriate learning experience. Additionally, most activities provide learners with the opportunity to create and adapt learning strategies to use beyond the classroom (activities 3), raise awareness about their own pronunciation (activities 2, 4, and 5), monitor their improvement or lack thereof (activities 1 and 6), and develop strategies on how to overcome communication breakdowns (activity 7). Furthermore, most activities suggest the use of other digital resources to support their learning process such as Text-to-Speech on GT, online dictionaries, and other websites. The activities also encourage students to take notes regarding GT's transcriptions and the procedures they have followed so they can adjust the human-machine interactions according to their own needs. All in all, the activities' procedures were designed to offer different strategies on how to use GT's ASR feature, allowing learners to be autonomous users of this technology.

The last criterion (Practicality) evaluates how adequate the resources used during the activities are. It encompasses software, hardware, and personnel resources required by the activity so it can be successfully implemented (CHAPELLE, 2001). All activities share the same infrastructure needed, except for activity 3, which requires the use of a headset to be performed (see activities' technical requirement). As previously stated, GT is a common tool to be used and it demands basic information technology (IT) infrastructure. Nonetheless, all activities require stable internet connection and an input device (e.g., microphone, headphone, cellphone's microphone). Hence, the activities have an IT requirement which may impede teachers and learners from performing them.

In summary, the rationale behind the ASR-based pronunciation activities was provided. The rationale to design the activities came from two relevant frameworks: the CCTA (CHAPELLE, 2001) and the CFTP (CELCE-MURCIA ET AL., 2010). Intelligibility issues presented on L2 Brazilian learners' speech were prioritized during the activities design as well as Brazilian learners' pronunciation learning difficulties. The next section presents the procedures for data analyses.

3.5 PROCEDURES FOR DATA ANALYSES

In order to answer both RQs, data were gathered from the background questionnaire (Appendix H), and the online survey (Appendix I), and then, analyzed quantitatively. The open-

ended question from the online survey was analyzed qualitatively. Data from the background questionnaire helped enrich the discussion and interpret the online survey's outcomes. Data collected from the instruments were input into Microsoft Excel's spreadsheets for analysis.

For the quantitative analysis, descriptive statistics were run to "summarize findings by describing general tendencies in the data and the overall spread of the scores" (DÖRNYEI, 2007, p. 213). The standard deviation (*SD*) and the means (*M*) were used to show useful information about the results (CRESWELL; GUETTERMAN, 2019). Most survey items were closed-ended questions following a 10-point Likert-scale; therefore, these questions were analyzed quantitatively. For these items, *summary statistics* were computed. According to Sue and Ritter (2007, p. 109), "summary statistics are designed to provide concise descriptions of the distributions of answers to survey questions". As the authors explain, *mean* and *standard deviation* are two common approaches to describe these distributions. Both measures were used during the data analyses. For the closed-ended multiple-choice questions, which were also analyzed quantitatively, data were scored, and *frequencies* were reported as percentages and presented on tables or graphs. This type of descriptive statistics "count the number of occurrences of a category" (LEAVY, 2017, p. 111).

Regarding the qualitative analysis, it was necessary to categorize all the data during a systematic content analysis, "whereby the pool of diverse responses is reduced to a handful of key issues in a reliable manner" (DÖRNYEI; TAGUCHI, 2009, p. 99) so the data could be properly coded and then analyzed. Data were gathered from the open-ended question of the online survey, the transcriptions of the workshop session recordings[8], and the saved chat logs. Regarding the transcriptions, two different files were hand-analyzed and coded. PS's transcription resulted in a 34-page document while the MS's transcription resulted in a 25-page document. In regard to the chat log files, the PS's log resulted in a 5-page document while the MS, 4-page document. Considering all the transcriptions and the chat logs files, a total of 34.036 words were analyzed. This combination of multiple sources allowed a multiple perspective data analysis; that is, an analysis that "provide several viewpoints from different individuals and sources of data as evidence for a theme" (CRESWELL; GUETTERMAN, 2019, p. 251). All data that did not provide information to answer the RQs were disregarded.

After the data were cleaned and combined, data were coded into four different categories (or themes): 1) Teacher Development Needs; 2) ASR Accessibility; 3) ASR Affordances; and 4) Activities Appraisal. Each of these categories matches the RQ1's constructs (category 1, 2, and 3, respectively). In addition, comments regarding the workshop design and its instruments were coded as Category 1. Category 4, *Activities Appraisal*, encompasses data to support RQ2's answer solely. As Creswell and Guetterman (2019, p. 245) explains, *categories* are "similar codes aggregated together to form a major idea in the database". The authors further state that the reason behind this small number of categories is that "it is best to write a qualitative report providing detailed information about a few themes, rather than general information about many themes" (CRESWELL; GUETTERMAN, 2019, p. 245).

[8] Microsoft Word Online's "transcribe feature" was used to convert the workshop recording audio file to a text transcript separating each speaker individually. This feature is only available on Word for the web and required a Microsoft 365 subscriptions (MICROSOFT WORD ONLINE, 2023).

Finally, a table with the coded data are presented in Appendix J. The data were manually sorted, cleaned, annotated, and then organized into the categories. The annotations consist of contextualizing or clarifying the speech passage. All annotated text is presented between square brackets. The source of speech, which is presented at the end of each speech passage, can be found between curly brackets, and varied from three sources: chat (from the chat log files), orally (from the transcriptions of the workshop session recordings), or survey (the open-ended item of the online survey). The result of the qualitative analysis is reported by the category that addresses each RQ or RQ's construct as can be seen in the next section. The results of this mixed-method approach are presented and discussed next.

4 RESULTS AND DISCUSSION

This chapter reports and explains the results in detail in order to answer the RQs. The theory and conceptual frameworks that guided this research are revisited to support the discussion. Section 4.1 answers RQ1 - *What are in-service English teachers' perceptions of ASR for pronunciation teaching after attending a workshop on how to use this speech technology regarding the following constructs (Teacher Development Needs, ASR Accessibility, and ASR Affordances)?*. There is one sub-section for each RQ1's construct to improve readability. These sections also match each survey's section designed for answering each construct. Then, Section 4.2 answers RQ2 - *How do in-service teachers appraise the ASR-based pronunciation activities designed to be implemented in L2 English classes*. In each section, first, the results from the quantitative analysis are reported and discussed, followed by the qualitative analysis' results, and summarized findings.

4.1 TEACHERS' PERCEPTIONS OF ASR FOR PRONUNCIATION TEACHING

4.1.1 Teacher development needs

In order to understand the English teachers' perceptions of ASR for pronunciation teaching regarding their development needs, a specific section of the survey was designed. Notwithstanding, data regarding the participant's experience with digital technologies and pronunciation teaching was gathered by the background questionnaire (see section 3.4.1). Table 10 shows the summary of the answers. Answers ranged from 0 (not much) to 10 (very much) Most participants feel fairly confident in teaching pronunciation (M = 7.9; SD = 1.2), and they answered that they often teach pronunciation in their classes (M = 8.0; SD = 2.0). Regarding digital resources, they feel confident in using them (M = 8.3; SD = 2.1), and they frequently use them in their classes (M = 8.9; SD = 1.4). Concerning the program Google Translate, most participants do not use it in their classes (M = 2.8; SD = 3.2). Finally, most of the participants do not use any ASR tool as a resource for teaching pronunciation (M = 2.0; SD = 3.0).

Table 10 - Summary of the participant's teaching practices responses

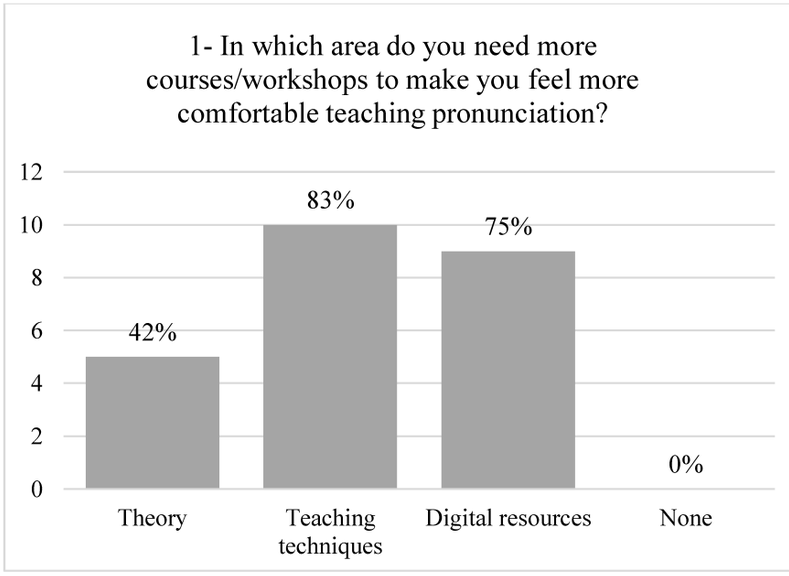
Background Questionnaire Question	M	SD
How confident are you in teaching pronunciation?	7.9	1.2
How often do you teach pronunciation in your classes?	8.0	2.0
How confident are you in using digital resources (computer, cell phone, projector, websites, apps) in your pedagogical practices?	8.3	2.1
How often do you use digital resources in your classes?	8.9	1.4
How often do you use Google Translate in your classes?	2.8	3.2
How often do you use any Automatic Speech Recognition (ASR) tool as a resource for teaching pronunciation?	2.0	3.0

Source: author

Bearing in mind the background of the participants, the online survey aimed to elicit information on their development needs related to technology and pronunciation teaching. The survey contained three questions. Question 1 asked participants in which area they feel the need for

more courses to make them feel more comfortable teaching pronunciation. This question listed 4 different options. The options and the results of the answers are displayed in Figure 3. Additionally, the option "other", which enabled participants to add specific information about their needs, was available to the participants. Only one participant used the "other" option to add a different answer[9]. However, his answer was considered as part of the "Digital resources" option, once it was considered that this option encompasses giving information on how to use the resources for a teaching purpose instead of only presenting the resource per se.

Figure 3 – Participant-teachers' development needs



Source: author

As shown in Figure 3, all participants felt the need for some development opportunity once no one marked the option "none". Considering the background of the 12 participants of this study (experienced teachers, most of them with higher education degree on English teaching and related area, and fairly confident in teaching pronunciation and using digital resources), 83% of them are willing to have more courses/ workshops regarding pronunciation "teaching techniques". The second most common answer was "digital resources" (75%). The least common option was "theory" (42%) regarding pronunciation teaching. Such results indicate that, even considering the background of the participants of this study, they still feel the need for more development opportunities regarding teaching techniques and digital resources. In addition, these results suggest that teachers would benefit more from practical suggestions than theory on teaching pronunciation. The results also imply a need for more education on how to integrate the digital resource with a proper teaching technique once most participants marked both "teaching techniques" and "digital resources" option. Therefore, digital resources without the proper technique to use them might not be so relevant to teachers.

Questions 2 and 3 inquired about the ASR use by the participant-teachers. As Table 11 displays, most participants feel comfortable using ASR to teach pronunciation after attending the workshop (M = 8.6; SD = SD=1.4). Their answers ranged from 7 to 10. In addition, most of them would use ASR to improve their own pronunciation (M = 8.7; SD = 1.6), and their answers ranged

[9] P4 answered "*Practical ideas on how to use the resources*" in the "other" option.

from 5 to 10. All things considered, we can speculate that a single workshop session might provide teachers with relevant insights on how to use ASR for both teaching and learning.

Table 11 – ASR use for teaching and learning

Survey's Item	***M* (*SD*)**
2- I feel comfortable to use ASR to teach pronunciation after attending the workshop.	8.6 (1.4)
3- I would use ASR to improve my own pronunciation.	8.7 (1.6)

Source: author

The previous above-mentioned statement can also be supported by the qualitative analysis. As can be seen in Table 12 below, most comments are related to the workshop design, and participants demonstrated to show an overall positive opinion about it. Not all participants made comments on this category; therefore, they were excluded from the table (P3, P4, and P6). The most common positive comments were in relation to the teacher's handout file, the hands-on part, good instructions on how to use the digital resource and the activities, and how useful the workshop was. On the other hand, the most common negative comments were related to the duration of the workshop (too short), and a wish for more time dedicated for participants to express themselves/ share their experience.

Regarding the duration of the workshop, most negative comments came from the PS participants. This was the reason why one extra hour for the MS session was added. Apparently, this modification was welcomed by the participants, as P7 states, "even though a 3h online workshop may seem a long time initially, the time really went by fast, and I don't see how it could have been shorter" (survey). Hence, a 3-hour online workshop may be more suitable than a 2-hour session, taking into account time for the literature review, the hands-on part, and time for discussion. In addition, a short literature review, and having more time dedicated to the hands-on part was also seen as a more appropriate choice, as P5 mentions, "the review of the literature was well summarized, was quite short, was, uh, great!" (orally, during the "Final Remarks" part). This comment is in accordance with the answers to Question 1 when participants mention more need for practical use than theory regarding pronunciation teaching.

Table 12 – Participants' comments on teacher development needs

P	**1st Category - Teacher Development Needs**
P1	I think the hands-on part is the rich[est] part of your workshop. {orally}
P2	I loved your workshop. I think that you could add 1h to have more time to discuss the topics and provide feedback to people that are participating. {survey}
P5	I really enjoyed the workshop, it was clear, with a nice flow and an interesting topic as well. I just wish we were given more time to express ourselves. {survey} [Final Remarks] The review of the literature was well summarized, was quite short, was, uh, great. [...] I just think we should have more time mainly in the beginning … for us to have some time to talk about our experience with Google Translate. {orally}
P7	Even though a 3h online workshop may seem a long time initially, the time really went by fast, and I don't see how it could have been shorter! The workshop was always interactive, with moments of explanation and hands on.... all the activities proposed were well explained and I believe this made all the difference in showing how we can take advantage of the tool. {survey}
P8	I really enjoyed the workshop. {survey}

P9	It [the workshop] was very clearly and logically presented. {survey}
P10	[Final Remarks] Uh, first of all, I'd like to thank you because it was a good opportunity to know how we can use this technology… like in a webinar. {orally}
P11	A very useful workshop and a great handout - seriously! {survey}
P12	The workshop was amazing and highly useful. {survey}

Source: author
Annotated text is presented between square brackets.
The source of speech is presented between curly brackets (chat, orally, or survey).

All in all, participants demonstrate their wish for more development opportunities related to teaching techniques and digital resources for pronunciation teaching. They also reported they feel comfortable using ASR for both teaching and learning pronunciation after attending the workshop. The results indicate that despite their educational background, they need more practical suggestions on how to teach pronunciation, and how to integrate the digital resource into pronunciation teaching. These results are in accordance with Silveira, Zanchet, and Pereira (2022). It is hypothesized, therefore, that this was the reason why participants demonstrated a positive attitude regarding the teachers' handout file, which contained the seven ASR-based pronunciation activities suggesting procedures to use the digital resource (ASR) from a pedagogical point of view (an activity plan).

Moreover, participant-teachers of this study seem to be confident in teaching pronunciation as the background questionnaire results indicate, which is also in accordance with previous studies (BUSS, 2016; SILVEIRA; ZANCHET; PEREIRA, 2022). Also, participant-teachers answered that they commonly use digital resources for pronunciation teaching; therefore, they might benefit from more development opportunities related to this area once it may be relevant to their practices.

Overall, participants seemed to have a positive opinion in relation to a 3-hour workshop session format. In addition, they were willing to use ASR to improve their own pronunciation after the workshop. Thus, this might be an alternative to remediate a lack of proper instruction on how to teach pronunciation during their undergraduate program (BUSS, 2013; COSTA, 2016), and to help those who did not have access to practice their own pronunciation (ALBINI; KLUGE, 2013).

4.1.2 ASR accessibility

Aiming at gathering participants' perceptions of ASR for pronunciation teaching regarding its accessibility, another section of the survey was designed, containing 4 questions. Question 1 inquired about how easy it was for the participants to use GT's voice recognition feature. Question 2 asked if participants and students would have access to Google Translate to practice their pronunciation. The means and the standard deviations are presented in Table 13.

Table 13 – Participants' perceptions of Google Translate's ASR feature

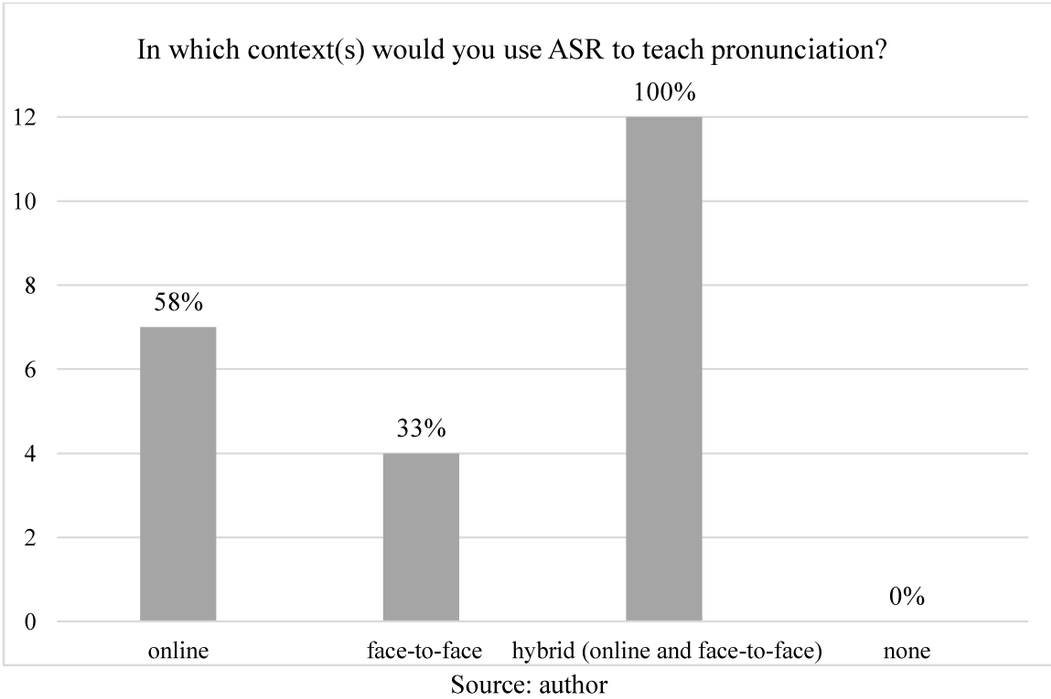
Survey's Item	***M* (*SD*)**
1- It is easy to use Google Translate's voice recognition feature.	9.3 (1.1)
2- My students have access to Google Translate to practice pronunciation.	9.3 (1.5)

Source: author

As it can be seen in Table 13, nearly all participants think it is easy to use GT's ASR feature (M = 9.3/ SD = SD=1.1). Also, most reported that their students would have access to GT to practice pronunciation (M = 9.3/ SD = SD=1.5). Their answers to Question 1 ranged from 7 to 10, while to Question 2, from 5 to 10. P12 was the only participant who assigned a rate value below 8. This participant lives in Argentina and teaches in the public sector (high school and higher education). Regarding the overall results, it is possible to conclude that GT's ASR is an accessible and easy to use digital resource for the participants of this study. However, the number of participants of this research is rather limited (*n*=12); hence, the result should not be generalized. Only 75% of the participants were currently living in Brazil, and participants' background varied considerably. Bearing in mind these particularities, no further implications can be drawn from these results.

Question 3 asked participants about the most appropriate teaching context (online, face-to-face, or hybrid) to use ASR for pronunciation teaching. Figure 4 displays the percentage of responses. A hybrid environment was unanimously chosen as appropriate by the participants (100%), followed by online (58%), and then, lastly, face-to-face (33%). Notice, however, that no participant marked the option "none". Thus, according to the participants of this study, a hybrid environment, that is, online and face-to-face classes, represents the most appropriate teaching context to use ASR for pronunciation teaching. This might be due to the possibility to use the ASR-based activities online and to provide further guidance during the face-to-face classes. Therefore, taking advantage of both environments and circumventing the possible limitations. Another explanation for the most common answers (online and hybrid) is related to the changes and challenges that the Covid-19 pandemic has brought to educators worldwide such as the implementation of remote learning (SILVEIRA; ZANCHET; PEREIRA, 2022). As many teachers had to deal with an online environment, such teaching context might have allowed them to see the advantages of it. Lastly, it is possible to deduce that a face-to-face environment was not marked by many participants due to the limitations that this context imposes on the teacher, mainly referring to the technological infrastructure. This topic was approached by the next question.

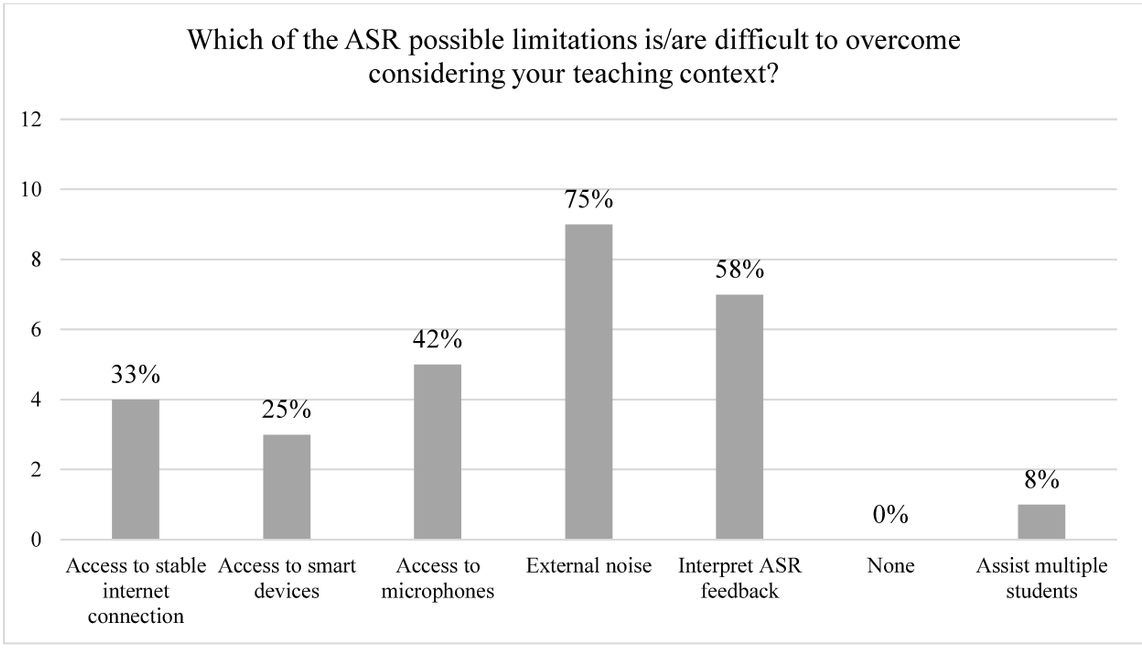
Figure 4 – Participants' preferred teaching context for ASR use



Source: author

Question 4 inquired about the possible limitations and difficulties that the participant-teachers would have to overcome to use ASR to teach pronunciation considering their teaching context. The percentages are presented in the following Figure (5).

Figure 5 – Participants' perceptions of ASR possible limitations



Source: author

The participant-teachers perceived the "external noise" as the most difficult possible limitation to overcome (75%). As speech recognition is easier with a proper input device, such as head-mounted microphones, and in a quiet place (JURAFSKY; MARTIN, 2021), the task of speech recognition is affected by the environment and equipment used. Therefore, if it is not possible to guarantee a quiet place (e.g., schools, or language centers), speech recognition will probably be affected. Likewise, "access to microphones" (42%), which was the third most common answer, might be an ASR possible limitation hard to overcome (JURAFSKY; MARTIN, 2021).

The second most common limitation was "helping students interpret the orthographic feedback provided by the ASR" (58%). Although ASR is a promising technology that offers immediate feedback on students' pronunciation (LIAKINA; LIAKIN, 2022), research shows that many participant-learners felt frustrated when they were not able to comprehend why the program could not understand them and how they could correct their speech (LIAKIN; CARDOSO; LIAKINA, 2017). Thus, a considerable percentage of participant-teachers believe that the difficulty in interpreting ASR feedback is a possible drawback. Participants also mentioned as a possible ASR limitation to overcome "the access to stable Internet connection" (33%), "access to smart devices" (25%), and "difficulty in assisting multiple students" (8). This last option was an added answer to the "other" option (P4).

In relation to the qualitative analysis, P7, P8, P9, P10, and P11 made comments regarding ASR accessibility. Table 14 shows these participants' comments. The most common positive comment was that the workshop provided ways to circumvent some of the ASR limitations. On the other hand, the negative comments were more common. There were comments regarding external noise (P7), limited cellphone and Internet connection (P8, P11), limited time for learners (private

students) to be able to use the digital resource properly (P9), difficulties handling different environments/ realities (private versus public sector) and students' proficiency level (P10).

Table 14 – Participants' comments on ASR accessibility

P	2nd Category - ASR Accessibility
P7	[Answer to the question "Do you think your students would enjoy the activities?] Sorry, I said maybe because I don't know. It's so personal. Maybe I need to learn where my students take the class, because if there is a noise, maybe you know, because when they say something on the microphone, I listen to the TV or other people speaking, so maybe this could be one reason for them not to, and so that's why I don't say yes, definitely. I need to do a research first. [...] Yeah, because maybe they're taking the class in the library too, and that they can't speak, you know, so maybe could be frustrating. Depending on these issues. {orally}
P8	So basically, what we were discussing [regarding the discussion on the breakout room] like my cons. The problem, you know, about this ASR, uhm, is accessibility, you know, 'cause I remember working with these types of things, especially I try to use, uhm, my phone, because… like… there were some lessons, almost no one had phones, you know, but they had but with no Internet, so I had to share my Internet with them. So, in public schools, I think this is the issue, but as a teacher we can try to find ways to work and bring different things like this. Once you showed us so. Yep, just really nice this workshop. {orally}
P9	[Regarding the discussion on the breakout room] OK, yeah, obviously different group of students [this participant teaches private students only] so the accessibility is not usually an issue, it's just time because everyone is so busy doing, they say or whatever else it is. But yeah, but I think definitely could be useful for the demographic (population) I teach usually. {orally}
P10	[Final Remarks] When I was talking in the in the other room [breakout room discussion] I was always saying that there are a lot of different realities, especially in Brazil according to the region. For example, I'm from Maranhão, so I'm a teacher here in Maranhão and it's really difficult sometimes to teach pronunciation for the students, especially when you have different realities. We have the public school realities. Uh, the English courses reality - I teach in English courses too, the university, and it's really difficult according to this reality, especially when these students they don't have like this background from the language. Sometimes they are really "raw" (not experienced) in the classrooms like starting, uh, to study English for their first time, and for example, it's something that it demands time. It demands to do things in advance. We need to set an environment for them in order to make them understand the concept and understand the tool and how they can use it, but it's really, really interesting. {orally}
P11	[Final Remarks] I just wanted to say that you know, for my own research also, I can share that even if we are planning to use technology in the classroom, we need to take into consideration the fact that we need to devote some time and it's not going to be 5 minutes that maybe half an hour or even an hour or two. It was to kind of create the habit of using some technology in the classroom or just demonstrate to our students how to access this technology. Ensure that everyone has access because I've learned that out of 20 people, there's usually two or three who will have problems with downloading a certain app or with access to the Internet, and you know that they won't be able to use a particular technology and you will have to stay after class depending on how you teach the course and help them out. And so, it's really good to have like a pre-started session which you actually check the technicalities and see whether people can use the tool or the activity, right? And so, I would very much advise everyone to have that before they introduce any technology in the classroom. {orally}

Source: author
Annotated text is presented between square brackets.
The source of speech is presented between curly brackets (chat, orally, or survey).

In short, most comments are in accordance with participants' answers to question 4. In addition, P4 mentioned that "it's really good to have like a pre-started session which you actually check the technicalities and see whether people can use the tool or the activity […]" (orally). This can be a recommendation for teachers who want to start using ASR with a student or in a group. Moreover, as P7 stated, "I need to learn where my students take the class" (orally), knowing the learning environment of the students first can also be a good strategy before introducing a new digital resource to them. P10 pointed out that "we need to set an environment for them [students] in order to make them understand the concept and understand the tool and how they can use it" (orally). All things considered, teacher's guidance may play an important role for the success, or lack thereof, of ASR for pronunciation teaching. Such findings are in accordance with previous studies in the area (GOTTARDI; ALMEIDA; TUMOLO, 2022; KIVISTÖ-DE SOUZA; GOTTARDI, 2022).

4.1.3 ASR affordances

As an attempt to comprehend the English teachers' perceptions of ASR for pronunciation teaching in relation to the ASR affordances, a third section of the survey was created, containing six questions. The means and standard deviations of close-ended Likert scale questions number 1 to question number 5 can be seen in Table 15.

Table 15 – Participants' perceptions of the ASR Affordances

Survey's Item	***M* (*SD*)**
1- ASR facilitates the teaching of pronunciation.	9.0 (1.0)
2- My students would be interested in using ASR to improve their own pronunciation.	7.7 (1.9)
3- I would use ASR as a complementary tool for teaching pronunciation in my classes.	9.1 (1.2)
4- ASR can encourage/motivate learners to produce more output outside the classroom.	8.7 (1.6)
5- ASR transcription (orthographic feedback) can be beneficial to the development of learner's pronunciation.	8.8 (1.1)

Source: author

Question 1 asked participants whether ASR facilitates the teaching of pronunciation. Participants' answers ranged from 8 to 10 (M = 9.0/ SD = 1.0). Most participants agree that ASR indeed facilitates the teaching of pronunciation. However, when they were asked if their students would be interested in using ASR to improve their pronunciation (Question 2), the results varied considerably (M = 7.7/ SD = 1.9). The answers ranged from 5 to 10, and three participants rated the minimum value (P2, P4, and P9). The level of education these participants[10] work in are fairly similar (both work with bilingual education and teach in elementary school). It is possible to conjecture that ASR may not be so attractive to younger learners. In addition, it is known that the speech recognition task for young children is usually harder to be performed by the ASR systems

[10] P9 works with private students and, unfortunately, there was no further questions asking about the age of the participants' students.

(JURAFSKY; MARTIN, 2021). Hence, although ASR seems to facilitate the teaching of pronunciation, ASR might be used with caution with young learners.

Previous findings, however, indicate that learners show high interest on ASR technology for pronunciation learning (CHEN, 2011; INCEOGLU; LIM; CHEN, 2020; KIM, 2006; MROZ, 2018). Nevertheless, these studies are usually conducted with higher education students. As the teaching environment of the participants of this study was considerably varied, more research should be conducted with the goal of understanding for which context(s) of education ASR technology is more appropriate for.

Question 3 inquired whether participants would use ASR as a complementary tool for teaching pronunciation in their classes. The answers to this question ranged from 7 to 10 (M = 9.1/ SD = 1.2). Most reported, thus, that ASR can be a complementary resource for pronunciation teaching, and that the participant-teachers hold a positive attitude towards using ASR for pronunciation teaching. These results are in agreement with previous studies that indicate ASR as a possible classroom complement for this purpose (GOTTARDI; ALMEIDA; TUMOLO, 2022; KIVISTÖ-DE SOUZA; GOTTARDI, 2022).

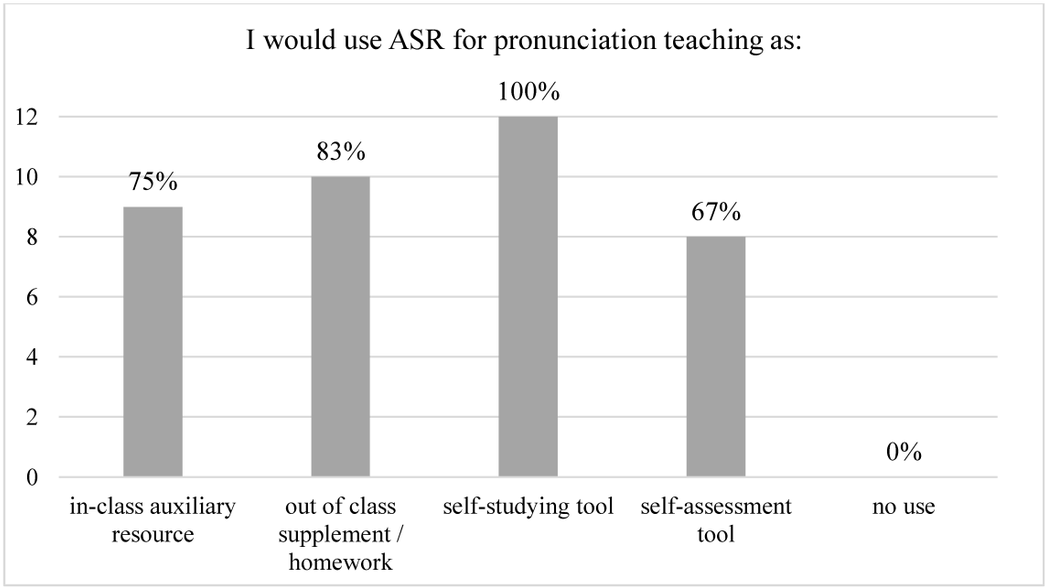
Participants were also asked whether they believe that ASR can encourage/ motivate learners to produce more output outside the classroom (Question 4). Participants' answers to this question ranged from 5 – 10 (M = 8.7/ SD = 1.6). P2 was the only one to rate below 7. It is hypothesized that this rating happened due to the level of education P2 is currently working in (young learners), and, therefore, P2's students may have lesser autonomy for using ASR on their own. Also, for young learners, some digital resources may demand teacher's mediation and may be distracting for them (SILVEIRA; ZANCHET; PEREIRA, 2022). Notwithstanding, the overall perceptions of the participant-teachers indicates that ASR can help by encouraging learners to produce more output outside the classroom. Such statement is supported by previous studies in the area that investigated learners' attitudes towards pronunciation learning aided by ASR practice (KIM, 2006; LIAKIN; CARDOSO; LIAKINA, 2015, 2017; MCCROCKLIN, 2018, 2014, 2016; MROZ, 2018; INCEOGLU; LIM; CHEN, 2020).

Question 5 asked participants whether ASR's orthographic feedback can be beneficial to the development of learner's pronunciation. Their answers ranged from 6 to 10. Even though, in the previous section, participants mentioned that interpreting ASR's orthographic feedback might be a difficulty to surpass while using ASR for pronunciation teaching, most participants (M = 8.8/ SD = 1.1) believe that ASR transcription can be beneficial for the learners' pronunciation development. As Kivistö-de Souza and Gottardi (2022, p. 778) state, "for some learners, the feedback offered by them [ASR-based pronunciation tools] might be less face-threatening than feedback from other language users". This might be one of the reasons why participants demonstrated a positive attitude towards ASR's orthographic feedback. In addition, it is possible to conjecture that this result might have a relation to the extended ASR-aided output production outside the classroom, as participants demonstrated positive attitude towards this matter as the previous question (4) inquired. On the other hand, we must have caution once some digital resources may provide limited feedback, which can hinder learners' motivation (SILVEIRA; ZANCHET; PEREIRA, 2022).

Question 6 elicited information on how participants would use ASR for pronunciation teaching. It was a multiple-choice question with an option "other" to allow participants to add new

entries. However, no one marked this option, nor have they selected the option “I would not use it”. As displayed in Figure 6, participants would unanimously use ASR as a learning tool for self-studying (100%). The perceptions of the participant-teachers are, hence, in agreement with previous studies that indicate ASR as an appropriate autonomous learning tool (INCEOGLU; LIM; CHEN, 2020; LIAKIN; CARDOSO; LIAKINA, 2017; MCCROCKLIN, 2016; MROZ, 2018; VAN LIESHOUT; CARDOSO, 2022). In fact, the literature suggests this as one of the most important ASR affordances. Thus, it is possible to hypothesize that, disregarding the level of education that the participant-teachers are currently teaching, ASR for pronunciation teaching might be an appropriate learning tool suggestion for self-study, and therefore, autonomous learning.

Figure 6 – Participants’ perceptions of ASR as a teaching resource



Source: author

Furthermore, many participants reported that they would also use ASR as an out of class supplement/ homework (83%). This result is in accordance with the aforementioned answers from the sections II and III of the survey as most the participants believe that ASR can motivate learners to produce more output *outside* the classroom, and that students can beneficiate from ASR’s orthographic feedback; thus, an interesting out of class supplement indeed. In addition, a considerable number of participants reported that they would use ASR as an auxiliary resource for the regular classes (75%), and also as a pronunciation self-assessment tool suggestion (67%). Considering the results from the previous sections, it can be deduced that ASR for pronunciation teaching can be adequate as an axillary resource for regular classes delivered, especially in a hybrid environment.

Finally, more than half of the participants marked all the available options, except for the option indicating that they would not use ASR for pronunciation teaching. These options were designed considering previous research in the area. Consequently, it is possible to state that the participant-teachers’ perceptions are in accordance with previous findings (KIM, 2006; LIAKIN; CARDOSO; LIAKINA, 2017; MCCROCKLIN, 2016; MROZ, 2018). However, these studies were conducted to gather learners’ perspective on ASR pedagogical use. Despite the limited number of

participants in this study, it is possible to hypothesize that teachers' perspective agrees with the one from learners based on what the literature suggests.

Considering all the close-ended Likert scale questions of the survey that addressed RQ1's constructs, it was possible to notice that P10 answered the maximum rating option (10) for all the questions. P3 demonstrated similar satisfaction but rated Question 2 of the first section of the survey as 9. Conversely, P9 (M = 7.4; SD = 1.7) and P12 (M = 7.2; SD = 1.5) were more judgmental regarding ASR for pronunciation teaching. Nevertheless, these participants were not as judgmental in regard to the ASR-based pronunciation activities presented to them during the workshop session, as we will see in the next section (4.2).

Moving to the qualitative analysis (see Table 16), it was not computed any comments from participants P2, P4, P6 regarding ASR affordances for pronunciation teaching. Most comments were related to Google Translate and how to use it for pronunciation teaching. For example, P3 stated that "it [the workshop session] was a good moment to rethink about google translate" (survey). Moreover, P8 stated that "I have never thought I could use Google Translate like that. I know for sure it helped me" (survey). Lastly, P10 explained that "I usually use Google Translate in my classrooms, but not in this perspective […] this perspective is really interesting in your research. […] It's another vision (perspective), another way [of using Google Translate], and probably I will use it in the future with my students" (orally). Such statements suggest that Brazilian teachers might be familiarized with Google products for English teaching and learning (GOMES JUNIOR; SILVA; PAIVA, 2022), but they might not know how to use it for pronunciation teaching to its full potential.

Table 16 – Participants' comments on ASR affordances

P	3rd Category - ASR Affordances
P1	With short questions the intonation is ok-ish, maybe my previous question was too long... It's certainly a good resource, though. {survey}
P3	It [the workshop session] was a good moment to rethink about google translate. {survey}
P5	[After the hands-on part] You should take Google translator with a grain of salt for sure. So, it [Google Translate] is important. It is a nice app, but you should take it with a grain of salt so students should use other resources as well. Well, we should use other resources as well as teachers. We shouldn't just, uh, trust what Google Translate says. We should use other resources as well. {orally} [Final Remarks] When it comes to Google Translate, I'm going to start doing some things with my students. Which is basically teaching them that there's more to Google Translate than just typing the word, 'cause, I mean, let's be clear, everybody uses Google Translate and it came a long way since, uh, it started. I would say… Google Translate, It's now in elementary school, whereas Alexa is in kindergarten. 'cause when you talk to Alexa it doesn't get what you're saying. Even though you're speaking in Portuguese or in English, whereas when you work with Google Translate it came a long way from when it, uh, started, so I think it's improved a lot and I'm going to start teaching my students not to just type the word, but let's imagine they heard… they listen to a song, and they got a word, but they don't know how to spell that word. So, OK, they can pronounce it. So, I think it's important for us, teachers, not only, uh, to provide them with these apps or, I don't know, things that they can use, but also explain the other features they could use. {orally}
P7	[While practicing tongue-twisters with Google Translate's ASR feature] For me it worked well, but I think I spoke very slowly […] now I spoke faster and it worked too! {chat}

P8	I have been using Siri too in some of my lessons, however I have never thought I could use Google Translate like that. I know for sure it helped me. {survey}
P9	Pronunciation is often completely neglected in language instruction, and this is a tool that can really help. {survey}
P10	[While practicing tongue-twisters with Google Translate's ASR feature] it worked pretty well!{chat} [Final Remarks] I'm an English teacher and I usually use the Google Translate in my classrooms, but not in this perspective, yes? so this perspective is really interesting in your research. It is really interesting too. Congratulations and I like it so much. [...] It's another vision, another way [of using Google Translate], and probably I will use it in the future with my students. {orally}
P11	[While exploring Google Translate] Just a shame GT can't save translations longer than 300 characters... {chat}
P12	[While exploring Google Translate's ASR feature] The app recognized my speech very fast! {chat}

Source: author
Annotated text is presented between square brackets.
The source of speech is presented between curly brackets (chat, orally, or survey).

As it can be seen in the Table 16, Google Translate ability to recognize speech rapidly and accurately was also mentioned by the participants. P7, for example, was surprised with GT's ASR feature while practicing some tongue-twisters during the workshop: "For me it worked well, but I think I spoke very slowly. [some moments later] Now I spoke faster and it worked too!" (P7, chat). Likewise, P10 made a comment during the same tongue-twisters practice: "it worked pretty well!" (P10, chat). In a similar vein, P12 exclaimed while exploring Google Translate's ASR feature, at the begging of the workshop, "the app recognized my speech very fast!" (P12, chat). Somehow, these participants seemed surprised with the results from GT's transcription or how it performed the task. It is hypothesized, thus, that although Google Translate is a commonly used tool for Brazilian users (GOOGLE, 2016), some teachers might not have even tried its ASR feature before, or at least, explored it with longer or more complex sentences (e.g., tongue-twisters).

Furthermore, it was interesting to observe that P5 was skeptical in relation to GT's ASR feature at first, but during the final remarks (workshop's last part), P5 demonstrated a positive attitude towards it. The participant commented during the "hands-on part" of the workshop that GT "is a nice app, but you should take it with a grain of salt so students should use other resources as well. […] We [teachers] shouldn't just, uh, trust what Google Translate says. We should use other resources as well." (orally). Nonetheless, during the final remarks, P5 stated that "that there's more to Google Translate than just typing the word" (orally). This participant further explained that GT "[…] is important for us, teachers, not only, uh, to provide them [students] with these apps or, I don't know, things that they can use, but also explain the other features they could use [referring to GT's speech features" (P5, orally). Possibly, the discussion with other participants in the breakout room made this participant (P5) reevaluate their perceptions in regard to this digital resource. This demonstrates the importance of allowing participants to discuss among themselves without the interference of the workshop facilitator; that is, speak freely in an environment where the participants are not being observed/evaluated by the facilitator or researcher. These comments also may indicate a possible negative bias regarding GT by the English teachers. Although GT is a translation tool, GT offers, with its speech features, much more than only translation. Also, much

improvement has been made during the past years regarding its speech technology (VAN LIESHOUT; CARDOSO, 2022).

On the other hand, P11 mentioned a frustrating limitation while exploring Google Translate, before the "hands-on" part. The participant wrote that it is "just a shame [that] GT can't save translations longer than 300 characters" (P11, chat). This limitation can be impactful if the teacher intends to use a longer text with their students and save this passage for later. This limitation, however, can be easily circumvented by saving the passage on a different file on Google Docs, for instance. Furthermore, P1 commented that GT is "certainly a good resource" (survey) but complained that GT performs well when transcribing short questions, rather than long ones.

Overall, there were much more positive comments than negative ones. As a final comment, P9 stated that "pronunciation is often completely neglected in language instruction, and this is a tool that can really help" (survey). As the qualitative data shows, most participants had a positive experience while using GT's ASR feature, and most of them seemed optimistic about its use for pronunciation teaching.

All in all, section 4.1 has summarized the participant-teachers' perceptions of ASR for pronunciation teaching after attending an online workshop session about this speech technology practical use. The participants of this study demonstrated to be aware of the limitations and affordances of ASR for pronunciation teaching. Similar results were found in a previous study (SILVEIRA; ZANCHET; PEREIRA, 2022) where most participant-teachers' demonstrated awareness regarding the limitations and affordances of digital technology aiding pronunciation teaching. Moreover, this section examined the pedagogical potential and affordances of ASR for pronunciation teaching from the perspective of English teachers and, as such, it aimed to understand teachers' attitudes towards this speech technology. These two objectives represent the *exploration* and the *assessing suitability* level of the chronological framework for CALL-based speaking research proposed by Cardoso (2022). The next section summarizes and discusses the participant-teachers' appraisal of the seven ASR-based pronunciation activities presented to the participants during the workshop session.

4.2 TEACHERS APPRAISAL OF THE ASR-BASED PRONUNCIATION ACTIVITIES

Aiming at gathering participant-teachers' appraisal of the ASR-based pronunciation activities designed to be implemented in L2 English classes, a fourth section of the online survey was designed containing 15 questions. Question 1 to 13 addressed the six Criteria for CALL Task Appropriateness (CCTA) from Chapelle (2001). Table 17 shows the means and standard variations for each question addressing the six CCTA. The overall mean for each criterion is also displayed. There were two questions addressing each criterion, with the exception of the last criterion that contained three questions. All the answers ranged from 3 to 10. At the end of the table, the overall mean and standard deviation considering all questions for all the six criteria is presented.

Table 17 – Summary of participants' answers for the questions addressing the six CCTA from Chapelle (2001)

Chapelle's Criteria	Survey Item	M (SD)	Overall Mean per Criterion
Language learning potential	1- The activities have a clear and reasonable learning focus.	9.6 (.8)	9.5
	2- The activities have addressed linguistic features related to pronunciation (ex. minimal pairs, vowel quality, segments, prosody, accuracy, fluency).	9.5 (.9)	
Learner fit	3- The activities have an appropriate degree of difficulty for the learners considering the target level indicated by the researcher during the workshop.	8.8 (1.5)	9.1
	4- The activities can be adapted according to learners' individual characteristics (e.g., age, computer experience, learning style).	9.4 (1.0)	
Meaning focus	5- The activities allow learners to use the language for interpreting and constructing meaning.	8.7 (1.7)	8.2
	6- The activities direct learner's attention primarily toward the meaning of the language.	7.8 (2.5)	
Authenticity	7- The activities show a strong correspondence between the tasks and what learners would expect to see outside the classroom.	7.4 (2.1)	8.0
	8- The activities address real pronunciation difficulties faced by Brazilian learners of English.	8.5 (1.8)	
Positive impact	9- The activities can help learners develop autonomous learning strategies.	9.1 (1.1)	9.0
	10- The activities can offer a positive teaching/learning experience with the digital technology used (Google Translate).	8.8 (1.7)	
Practicality	11- The activities can be implemented considering the particular constraints of my class and language program.	8.3 (1.5)	8.9
	12- The activities are adaptable to my teaching context in relation to the **equipment** used (e.g., microphone, earphone, computer, laptop, cellphone).	9.4 (.7)	
	13- The activities are adaptable to my teaching context in relation to the **infrastructure** required (e.g., internet, wi-fi, computer lab).	8.8 (1.2)	
		Overall *M*:	**8.8**
		Overall *SD*:	**0.7**

Source: author

Question 1 and 2 addressed the "Language learning potential" criterion. For both questions the standard deviation was low (SD < 1), and the overall mean for this criterion was high (M = 9.5). Among all the six criteria, this one obtained the best rates (highest means and lowest standard deviations). These results indicate that the participant-teachers could clearly see the language learning potential (beneficial focus on form) of the ASR-based pronunciation activities. As

suggested by Chappelle (2001), this criterion should be prioritized while designing CALL activities or else the introduced technology might not be pedagogically effective. Thus, according to the great majority of the participants of this study, the activities present a clear and reasonable learning focus and they have clearly addressed linguistic features related to pronunciation.

The second criterion (learner fit) was addressed by questions 3 and 4. This criterion evaluates how appropriate the activity is according to learner's characteristics, and its difficulty level (CHAPELLE, 2001). The results suggest that the overall evaluation of the participants was positive (M = 9.1). This criterion received the second best rates. Nonetheless, while participants could clearly see a high adaptability of the activities according to learners' individual characteristics (Question 4, M = 9.4; SD = 1.0), there was a small divergence in the appraisal of the activities' degree of difficulties for the learners (Question 3, M = 8.8; SD = 1.5). This indicates a need to review the proposed CEFR proficiency level of the activities.

Questions 5 and 6 were related to the third criterion (meaning focus). The overall mean was low if compared to the other criteria (M = 8.2). This criterion evaluates whether learner's attention is focused on the meaning of the language behind the activity's purpose. Most participants reported that the activities allow learners to use the language for constructing and interpreting meaning (Question 5, M = 8.7; SD = 1.7). However, participants' opinions seem not to be so congruent when they were inquired whether the learner's attention was primarily directed toward the meaning of the language during the activities (Question 6, M = 7.8; SD = 2.5). Even though all activities were designed to improve learners' speech intelligibility aiming at successful communication, it seems that the intelligibility issues that might disrupt communication addressed by each activity were not clearly observed as a focus on meaning (successful communication). Such a result might imply some modifications in the activities related to how to make it clear to the teachers the intelligibility issues addressed by each activity and how these issues may disrupt communication.

The fourth criterion is related to the "authenticity" of the CALL activities. Questions 7 and 8 elicited information on this matter. This criterion received the lowest rates (M = 8.0). Questions 7 presented the lowest rates among all the other questions (M = 7.4/ SD = 2.1). It inquired about whether the activities showed a strong correspondence between the activity and what learners would see outside the classroom. It is possible to hypothesize that minimal pair activities and tongue-twisters might not represent authentic language use. Nevertheless, such activities supply learners with speech correction strategies that aid successful communication. This criterion evaluates to what extend learners might profit from the activities outside the classroom (CHAPELLE, 2001). Hence, it is emphasized that a clearer link between the intelligibility issues addressed by the activities and its implications for successful communication should be drawn. Also, the activities suggest the practice of words and sentences of everyday use, embracing swearwords, which may result in social difficulties (LEVIS, 2018). All in all, it is possible to speculate that the procedures proposed by the activities to practice those words and sentences were seen as artificial and not authentic by the participants of this study to some extent. This result is in accordance with a previous study that indicated negative comments from teachers in relation to how some digital resources for pronunciation teaching may differ from real language use (SILVEIRA; ZANCHET; PEREIRA, 2022).

Regarding question 8, participants fairly agree that the activities addressed real pronunciation difficulties faced by Brazilian learners of English (M = 8.5). It is important to mention that not all participants lived in Brazil, although their native country was never inquired in the background questionnaire. It is speculated that, because of this reason, the standard deviation was high (SD = 1.8). Another possible explanation for this result is the varied background of the participants regarding the teaching experience and the current level of education they are teaching, which may result in a different view of learners' difficulties.

Moving to the fifth criterion (positive impact), Questions 9 and 10 endeavored to evaluate whether the activities can provide learners with the possibility to develop autonomous learning strategies (Question 9, M = 9.1; SD = 1.1), and if the activities can offer a positive teaching/ learning experience with the digital resource used; that is, Google Translate (Question 10, M = 8.8; SD = 1.7). This criterion evaluates the activity's positive impact on both teaching and learning as well as the possibility to develop student's ability to learn to learn (CHAPELLE, 2001). As the overall mean of this criterion was high (M = 9.0), it is deduced that the activities accomplished this objective successfully.

Finally, the last criterion (practically) endeavored to evaluate whether the resource used during the activities (GT's ASR feature) was adequate, encompassing hardware, software, and personnel resources (CHAPELLE, 2001). The overall mean was considerably high (M = 8.9). Question 11 asked whether the activities can be implemented considering the constraints of the participants' class and language program. This question obtained the lowest rates of this criterion (M = 8.3; SD = 1.5), although it might not be considered a low rate. According to the diverse background of the participants of this research, this result is not surprising. Question 12 asked whether the activities can be adaptable to participants' teaching context in relation to the *equipment* used (M = 9.4; SD = 0.7) while Question 13 asked in relation to the *infrastructure* required (M = 8.8; SD = 1.2). It seems that participants had more concerns about the infrastructure required than the equipment used. According to these results, no major difficulties were found by the participants as an impediment to implementing or adapting the activities considering their teaching contexts.

All in all, the overall mean considering all the questions that address the six CCTA from Chapelle (2001) was considerably high (M = 8.8; SD = 0.7). It is concluded, therefore, that the activities suit most of the participant-teachers' realities and that the participants demonstrate an overall positive attitude towards the presented ASR-based pronunciation activities. In addition, some improvement suggestions can be made about the activities design from the results of these questions as previously discussed. In relation to the rigorousness of participants' appraisal to those questions, it is possible to affirm that 83% (*n*=10) of the participants have a rate mean of 8 or above, as Table 18 displays.

Table 18 – Participants' rating variance to the questions addressing the six CCTA

P*	**M (SD)**	**Answer's Range**
P1	7.7 (2.8)	3 - 10
P2	8.6 (1.6)	5 - 10
P3	9.8 (.4)	9 - 10
P4	8.3 (1.7)	5 - 10
P5	8.2 (1.0)	6 - 9

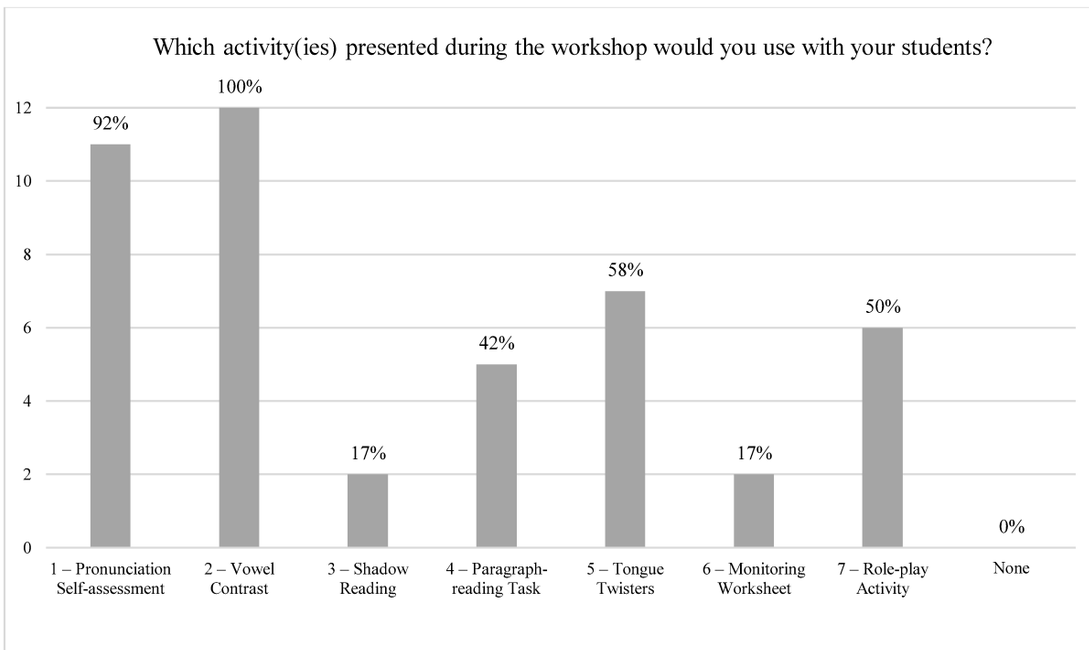
P6	9.0 (.4)	8 - 10
P7	9.6 (1.0)	7 - 10
P8	9.6 (.7)	8 - 10
P9	9.4 (1.0)	7 - 10
P10	9.1 (1.9)	5 - 10
P11	7.9 (2.2)	3 - 10
P12	8.2 (.8)	6 - 9

Source: author
*Participant code.

Table 18 displays the means, standard deviations, and the range of participants' answers to Question 1 to 13. It can be observed that only P1 (M = 7.7; SD = 2.8) and P11 (M = 7.9; SD = 2.2) have a rate mean below 8. Also, they were the only participants who assigned the lowest rate value (3). On the other extreme, P3 (M = 9.8; SD = 0.4), P7 (M = 9.6; SD = 1.0) and P8 (M = 9.6; SD = .7) were the participants with a rate mean higher than 9.5. Interestingly, the most rigorous participants regarding the pronunciation activities (P1, and P11) were not the same participants with the lowest rate mean considering the previous section, ASR affordances (P9, and P12). It is hypothesized, hence, that activities designed considering both affordances and limitations of the digital resource can circumvent some of the resource's limitations to some extent by offering appropriate instructions.

After the questions that addressed the six CCTA, Question 14 asked participant-teachers which one(s) of the seven ASR-based pronunciation activities presented during the workshop they would use with their students. Figure 7 illustrates the distribution of the responses to this multiple-choice question.

Figure 7 – Activities participant-teachers would use



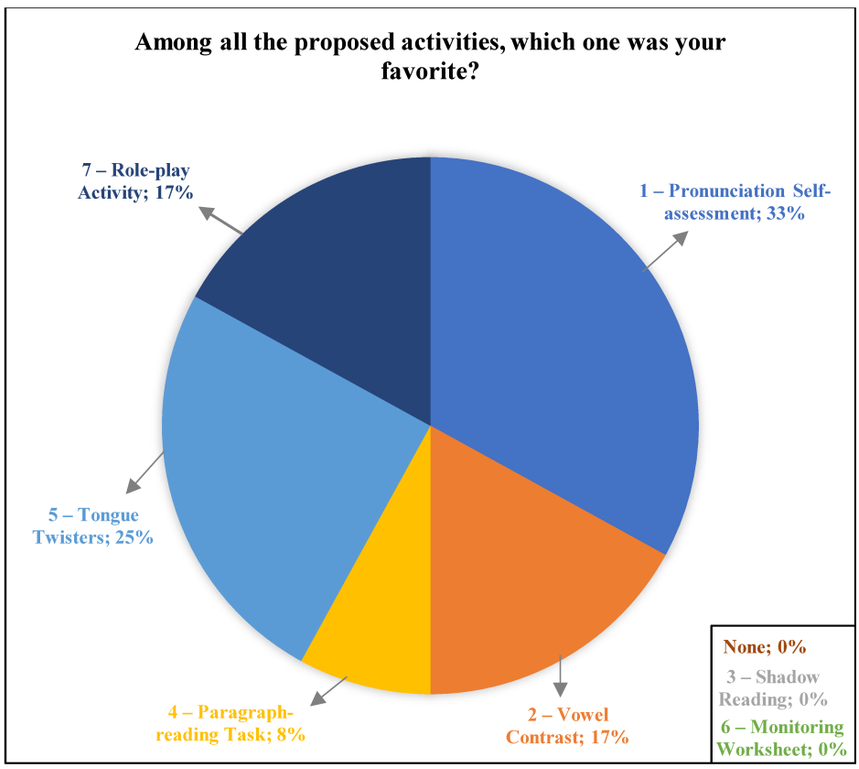
Source: author

As can be seen in Figure 7, all participants marked Activity 2 (Vowel Contrast). The second most marked option (92%) was Activity 1 (Pronunciation self-assessment). Activity 5 (Tongue-twisters), 58%, and Activity 7 (Role-play Activity), 50%, were the other two activities that at least

half of the participants marked as a viable option for their students. Therefore, it can be concluded that these activities fit the participants' teaching needs more properly. On the other hand, regarding Activity 3 (Shadow Reading), it is possible to speculate that participants did not mark this option as frequently because this was the only activity that requires headphones. In addition, this activity is more difficult for teachers to instruct the students (there were more steps on this activity, and they are also more complex), and for students to follow the procedures (students need to access GT in two different tabs to follow the shadow reading activity, besides the headphones requirement). Lastly, participants might have avoided this activity because it deals with suprasegmentals, which may be seem as a more difficult feature to teach, and less time is dedicated to teaching it if compared to segmental aspects of speech (BUSS, 2016).

The last question of this section (Question 15) asked which of the presented activities was the participant's favorite. Figure 8 displays the percentage of participants' answers. The most common marked option (33%) was Activity 1 (Pronunciation self-assessment). The second most common answer was Activity 5 (Tongue-twisters), with 25% of the answers, followed by Activity 2 (Vowel Contrast) and Activity 7 (Role-play Activity), with 17% of the answers each. The last favorite option was Activity 4 (paragraph reading task), with only 8% of the answers. This result is in agreement with the previous questions, demonstrating participants' low interest in Activities 3, 4, and 6. Intriguingly, Activity 2 was not selected as a favorite by many participants even though all of them would use it with their students.

Figure 8 – Participants' favorite activity



Source: author

Moving to the qualitative data analysis, many participants did not make any comments regarding the activities. Table 19 displays the comments provided by P1, P4, P5, P6, P7, and P10. Some comments were compliments on the activities and the tips provided during the workshop

session (P1 P6, and P10, survey). These participants demonstrated positive attitudes towards the activities and the way they were presented to them.

Table 19 – Participants' comments on the ASR-based pronunciation activities

P	4th Category - Activities Appraisal
P1	The activities were really good! (*Translated speech from Portuguese*) {survey}
P4	[Answer to the question "Do you think your students would enjoy the activities?] I think it will depend on the profile of your students. So, for example, I'm working with teenagers, and they don't like anything… like my group of students at the moment. They don't like anything, so I know that this type of activities wouldn't work with them, but on the other hand I have other groups... but yeah, it would work for sure. So, I think it's part of your job to know your students well and then you can decide which activities would work according to the profile of your students. {orally} I think this [activity 1] could be also like a nice way to use the technology if you are trying to raise awareness to the intonation. {orally}
P5	[After the hands-on part] I just wanna comment on the on the second activity. I really liked the second activity and even though it is intended for beginners. I wouldn't say it's just beginners 'cause most of our learners, they have questions about "oh is it cheap? Is it cheap? Is it …?" Yep, so I really liked the activity even though I work with private students, I think for people that work with other classroom settings it would be nice as a variation for them to be like working in pairs and they would come up with some sentences and then the partner would have to guess what is the word that the person is trying to say. But it's also very important for the teacher to, yeah, monitor and check if the word is being is not being mispronounced, But I'm totally going to use the second activity and the last one you gave. It's more like for B1, B2 [learners], I don't know… a bit more advance, but I also really liked it. […] But yeah, just that I'm totally going to use the second activity. I think it's really important for them to know the difference between these words and the example of "fill" and "feel"… I, uh, I gave two options: [speaking to Google Translate] I said "can you fill?" Not "[fill] it" I just said "can you fill" without the "it" in the end, so "can you fill" and then the Google Translate transcribed sentiu [feel] and then I said "can you fill my glass", and then it [Google Translate] put like "fill" encher, so sometimes it works by the context as everybody was saying before so. {orally} [After returning from the breakout room] So, as we [the participants] were discussing [on the breakout room], we thought the activities 1 and activity 2 would be the ones that we would use the most. Most mainly because of the fact that the third one, the shadow reading, would be a bit tricky, not just for the teacher but also for the students to be able to perform it well even though it is a nice activity, it would demand a lot of time to be able to do it right? 'cause you don't want to do a pronunciation activity where it's not effective. So you don't want to do it like "oh, let me do it in 5 minutes", just 'cause I think it's nice... it [the pronunciation activity] needs to be, uh, effective, right? So, uh, the tongue-twister [activity 5], the second one about the words that have similar pronunciation, and the first one - to complete with their own information - would be the ones we [participants] would use the most. {orally} [Final Remarks] I'm gonna for sure use some of the activities you provided into my practice. {orally}
P6	I loved the tips! (*Translated speech from Portuguese*) {survey}

P7	[the workshop provided] a wide range of activities that suit all groups of students (and may also be adapted). {survey} I loved the activities you created. I wanna use it with my Portuguese students as well {chat} [Answer to the question "Do you think your students would enjoy the activities?] By the way, the activities, uh, that you created? Yes, I would definitely say yes, but I need to learn about the context that they're taking the classes, that's why, uh, maybe, sometimes like homework is a better idea… that's why, like, they all [the activities] offer alternatives… like, homework is always an alternative. Uh, if you teach private students is one thing… a class is a completely different thing, right? You have to say if it's possible or not. {orally}
P10	I really liked the research and the activities! {survey}

Source: author
Annotated text is presented between square brackets.
The source of speech is presented between curly brackets (chat, orally, or survey).

There were two comments indicating the need for the teacher to understand their students in order to make a proper decision on which activities to use and how to use them. P4 stated that “it's part of your job [as a teacher] to know your students well and then you can decide which activities would work according to the profile of your students” (orally). Likewise, P7 commented that it is necessary to learn more about the context in which the teacher’s students are taking classes “that's why, uh, maybe, sometimes like homework is a better idea… that's why, like, they all [the activities] offer alternatives… like, homework is always an alternative” (P7, orally). This participant further stated that the workshop offered “a wide range of activities that suit all groups of students (and may also be adapted)” (P7, survey), and that he/she would use the activities “with my Portuguese students as well” (P7, chat). Therefore, it can be implied that, although the activities were designed for English teaching, the activities can be adapted to teach other foreign languages.

Moreover, there were comments addressing specific activities. For example, P4 mentioned that Activity 1 (Pronunciation self-assessment) "could be also a nice way to use the technology if you are trying to raise awareness to the intonation” (orally). This participant’s suggestion combines two recommendations from Buss (2016): 1) Brazilian EFL learners might benefit from a more frequent use of awareness-raising techniques; and 2) more professional training regarding some suprasegmental features, such as intonation, that are not frequently taught (BUSS, 2016). Thus, it might be an interesting ASR-based pronunciation activity, or variation for Activity 1, to focus on raising awareness to suprasegmentals such as the intonation.

Participant 5 commented right after the “hands-on part”; that it, after the presentation of all activities, that he/she would definitely use Activity 2 (Vowel Contrast) and Activity 7 (Role-play Activity) with his/her private students (P5, orally). Nevertheless, after the brainstorming part of the workshop involving all the participants, P5 informed that “we [participant-teachers] thought activities 1 and activity 2 would be the ones that we would use the most” (orally). This comment is in agreement with the findings from Question 14 results. This participant further added Activity 5 (Tongue-twisters) as one of the activities participant-teachers discussed during the brainstorming part that they would use the most (P5, orally).

Furthermore, P5 further commented, after the brainstorming part, regarding the difficulties that Activity 3 (Shadow Reading) would possibly pose, indicating that this activity "would be a bit tricky, not just for the teacher but also for the students to be able to perform it well even though it is a nice activity, it would demand a lot of time to be able to do it right?" (P5, orally). Hence, time constraint is one of the reasons why participant-teachers did not select Activity 3 when answering Question 15 and only a few of them marked it for Question 14. Notwithstanding, the only comment about time constraint being a problem for the participant-teachers was in relation to Activity 3. All things considered; it is hypothesized that all the other activities predict a reasonable amount of time to be delivered by the participant-teachers once no one made any negative comments on that matter.

All in all, the qualitative analysis is in agreement with the findings from the quantitative analysis. Overall, participant-teachers indicated that Activity 1 (Pronunciation self-assessment), Activity 2 (Vowel Contrast), Activity 5 (Tongue-twisters), and Activity 7 (Role-play Activity) were the activities they would use the most. Activity 3 (Shadow Reading) was disregarded as a viable option mainly due to time constraints and the complexity involving its procedures. Participants reported holding positive attitudes towards the activities and how they were presented during the workshop session. The next session summarizes the main findings of this study, provides some pedagogical implications, reports the limitations of this research, and offers suggestions for further research in the field.

5 CONCLUSIONS

This study sought to explore the affordances of ASR for pronunciation teaching from the perspective of in-service English teachers and investigate teachers' appraisal of ASR-based pronunciation activities designed to be implemented in L2 English classes. For this purpose, a workshop session was delivered to two different groups. A total of 12 participant-teachers joined two different workshop sessions and followed the procedures for data collection. To answer the researcher questions, this research followed a mixed-method approach, combining the quantitative analysis of the data gathered by an online background questionnaire and an online survey, and the qualitative analysis of the data from the open-ended question of the online survey, the transcriptions of the workshop session recordings, and the saved chat logs. This final chapter presents the summary of the main findings of this research with the pedagogical implications (section 5.1), followed by the limitations of this study and suggestions for further research (section 5.2).

5.1 MAIN FINDINGS AND PEDAGOGICAL IMPLICATIONS

In relation to the first research question (RQ1) - *What are in-service English teachers' perceptions of ASR for pronunciation teaching after attending a workshop on how to use this speech technology regarding the following constructs: Teacher Development Needs, ASR Accessibility, and ASR Affordances?* – the findings can be divided by construct. First, regarding the teachers' development needs, participant-teachers indicated a need for more development opportunities regarding teaching techniques and digital resources. The results of this study further suggest that teachers would benefit more from practical suggestions than theory on teaching pronunciation. Similar to Silveira, Zanchet, and Pereira's (2022) findings, the results imply that, despite their educational background, they need more practical suggestions on how to teach pronunciation, and how to integrate the digital resource into pronunciation teaching.

Regarding the workshop session delivered to the participant-teachers, it was concluded that a single, brief workshop session was able to provide them with relevant insights on how to use ASR for both teaching and learning. In addition, participant-teachers demonstrated a positive attitude towards the use of ASR to improve their own pronunciation. Therefore, ASR-based pronunciation practice can be an alternative to help teachers who did not have access to practice their own pronunciation (see Albini and Kluge (2013)).

Moving to the second construct (ASR accessibility), the participant-teachers of this study indicated that GT's ASR feature is an accessible and "easy to use" digital resource. Furthermore, results suggest a hybrid environment (online and face-to-face classes) as the most appropriate teaching context to use ASR for pronunciation teaching. Finally, regarding the ASR limitations, participant-teachers perceived "external noise" as the most difficult limitation to overcome while using ASR for pronunciation teaching. Other commonly mentioned drawbacks were not having "access to microphones", "limited Internet connection" and "difficulty in interpreting ASR feedback". Participant's perceptions are in accordance with the literature (JURAFSKY; MARTIN, 2021; LIAKIN; CARDOSO; LIAKINA, 2015).

On the other hand, participants mentioned that the workshop session provided them with means to circumvent some of these limitations. Also, results indicate that teacher's guidance may play an important role for the success of ASR for pronunciation teaching, mainly regarding the preparation of the teaching environment. The importance of teacher's role for successful ASR use for pronunciation teaching is also cited in previous studies in the area (GOTTARDI; ALMEIDA; TUMOLO, 2022; KIVISTÖ-DE SOUZA; GOTTARDI, 2022; LIAKINA; LIAKIN, 2022; WALLACE, 2016).

Finally, regarding the last construct (ASR affordances), the results imply that, although ASR seems to facilitate the teaching of pronunciation, ASR might be used with caution with young learners. In addition, the results further indicated a possible negative bias regarding GT from the participant-teachers at first. The most expressive negative comments were in relation to GT's limitation to interpret the correct intonation in longer questions, and the impossibility to save translations longer than 300 characters. After the workshop session, however, most participants reported having a positive experience using GT's ASR feature, and most of them seemed optimistic about its use for pronunciation teaching. Therefore, results indicate that an online translation can be suitable for pronunciation practice, which agrees with previous studies (HE; CARDOSO, 2021; VAN LIESHOUT; CARDOSO, 2022).

In general, participants of this research demonstrated to be aware of both affordances and limitations of ASR for pronunciation teaching. These results are in agreement with a previous study (SILVEIRA; ZANCHET; PEREIRA, 2022) where most participant-teachers' demonstrated to be aware of the affordances and limitations of digital technology aiding pronunciation teaching. All in all, the overall perceptions of the participant-teachers indicate that ASR can 1) be used as a learning tool suggestion for self-studying; 2) help encourage learners to produce more output outside the classroom; 3) provide students with relevant orthographic feedback; 4) be used as an out of class supplement; and 5) be an adequate auxiliary resource for pronunciation teaching in regular classes, especially, in a hybrid environment. These results are in agreement with other studies that investigated ASR for pronunciation teaching (GOTTARDI; ALMEIDA; TUMOLO, 2022; INCEOGLU; LIM; CHEN, 2020; KIVISTÖ-DE SOUZA; GOTTARDI, 2022; LIAKIN; CARDOSO; LIAKINA, 2015, 2017; MCCROCKLIN; EDALATISHAMS, 2020; MCCROCKLIN; HUMAIDAN; EDALATISHAMS, 2019; MROZ, 2018; VAN LIESHOUT; CARDOSO, 2022).

In relation to the last research question (RQ2) - *How do in-service teachers appraise the ASR-based pronunciation activities designed to be implemented in L2 English classes*. – the findings indicated that the activities which fit the participants' teaching needs more properly and, therefore, they would use the most, were Activity 2 (Vowel Contrast), Activity 1 (Pronunciation self-assessment), Activity 5 (Tongue-twisters), and Activity 7 (Role-play Activity). On the other hand, participant-teachers demonstrated low interest for Activities 3 (Shadow Reading), Activity 4 (Paragraph-reading Task), and Activity 6 (Monitoring Worksheet), even though participants have answered that they would use all of them with their students. It was identified that time constraint and the complexity involving the procedures were the reasons why participant-teachers demonstrated low interest for Activity 3 (Shadow Reading).

Overall, participants reported positive attitudes towards the ASR-based pronunciation activities and how they were presented during the workshop session. In addition, results imply that

the design of the activity is of paramount importance when using ASR for pronunciation teaching. Participant-teachers' perceptions indicated that, by considering both affordances and limitations of the digital resource, and offering appropriate instructions, it is possible to circumvent some of the ASR's limitations. Finally, participants mentioned that, although the activities were designed for English teaching, the activities can be adapted to teach other foreign languages.

It was also possible to conclude that this study has presented activities that provided teachers with procedures and techniques to use ASR for pronunciation teaching in a practical way. In addition, the activities presented in this study have shown that ASR-based pronunciation practice can be implemented for phases three to five (controlled practice, guided practice, and communicative practice) of the Communicative Framework for Teaching Pronunciation proposed by Celce-Murcia et al (2010). Also, ASR's transcription can provide learner with orthographic feedback aiding pronunciation improvement during the three abovementioned phases.

Moreover, participant-teachers evaluated the ASR-based pronunciation activities according to the six criteria for CALL task appropriateness proposed by Chapelle (2001). According to the great majority of the participants of this study, the activities present a clear and reasonable learning focus and they have clearly addressed linguistic features related to pronunciation (*Language Learning Potential*). Also, participants could clearly see a high adaptability of the activities according to learners' individual characteristics (*Learner Fit*). Nonetheless, there was a small divergence in the appraisal of the activities' degree of difficulty for the learners, indicating a need for revising the proposed CEFR proficiency level of the activities.

Regarding the third criterion (*Meaning Focus*), most participants reported that the activities allow learners to use the language for constructing and interpreting meaning. However, participants' opinions were not congruent when they were inquired whether the learner's attention was primarily directed toward the meaning of the language during the activities. Thus, these results imply a need for clarifying the intelligibility issues addressed by each activity and how these issues may disrupt communication.

In addition, participants fairly agree that the presented activities addressed real pronunciation difficulties faced by Brazilian learners of English (*Authenticity*). However, results demonstrate that the procedures proposed by the activities to practice some text passages were seen as artificial and not authentical by the participant-teachers to some extent. This result is in agreement with a previous study that indicated negative comments from participant-teachers on how some digital resources aiding pronunciation teaching might differ from real language use (SILVEIRA; ZANCHET; PEREIRA, 2022).

In relation to the fifth criterion (*Positive Impact*), results indicate that the activities can provide learners with the possibility to develop autonomous learning strategies, and the activities can offer a positive teaching/ learning experience with the digital resource used (GT). Finally, regarding the last criterion (*Practicality*), participant-teachers' indicated that the activities suit most of their realities and that the participants demonstrated an overall positive attitude towards the presented ASR-based pronunciation activities. Participants also indicated that no major difficulties were found to implement or adapt the activities considering their teaching contexts.

In summary, participant-teachers demonstrated an overall positive attitude towards ASR for pronunciation teaching and towards the seven ASR-based pronunciation activities. Although the

activities can be improved as aforementioned stated, they still can provide teachers with practical ideas on how to integrate ASR technology into pronunciation teaching. In short, the affordances of ASR for pronunciation teaching are numerous. Yet, teachers' guidance is of paramount importance for an optimal result. Thus, the presented findings and pedagogical implications may help teachers to circumvent some of the ASR limitations and benefit from ASR affordances.

5.2 LIMITATIONS OF THE STUDY AND SUGGESTIONS FOR FURTHER RESEARCH

It is important to stress that the number of participants in this research was quite limited (n=12); hence, the result should not be generalized. Due to time constraint, it was not possible to follow the procedures for recruiting more participants and delivering another workshop session once data collection and analysis, especially the qualitative data, demanded a lot of time. On the other hand, the mixed-method approach allowed to analyze in depth the perceptions of the participant-teachers. Thus, more research following a similar approach should be conducted in order to offer more robust results regarding ASR for pronunciation teaching. Also, further studies should investigate the use of other ASR-based dictation tools to offer more possibilities for teachers.

Another limitation of this study was the participant recruitment method. We can wonder whether the participant-teachers who joined the workshop session and answered the forms were teachers who are highly interested in pronunciation teaching or educational digital technologies, maybe both. Therefore, the results regarding teachers' attitudes towards ASR and the ASR-based activities should be taken with caution. In addition, the backgrounds of the participants of this study do not represent the majority of the English teachers in Brazil, nor worldwide. Therefore, further research with a more heterogeneous group of teachers, considering their background education and teaching interests, should be conducted in order to comprehend the perspective of a different population.

Finally, as the teaching contexts of the participants of this study were considerably varied (e.g., Young Learners, Elementary School, High School, Higher Education, Language School, and Private teacher), more research should be conducted with the endeavor of understanding which context(s) of education ASR technology is more appropriate for. Also, it is noteworthy that the participants of this study appraised the activities based on the workshop presentation and their teaching experience. Therefore, classroom observation could shed some light regarding teachers' difficulties on following the activities' procedures and handling possible ASR limitations. Lastly, since there was a small divergence in the participant-teachers' appraisal of the activities' degree of difficulty for the learners, revising the proposed CEFR proficiency level of the activities is therefore advisable.

All things considered, the results reported in this research contribute to the field of pronunciation teaching by offering practical suggestions on how to implement ASR technology in L2 English classes. Moreover, this study explored the affordances of ASR for pronunciation teaching from the perspective of in-service English teachers. Thus, the results can contribute to the field of applied linguistics by offering insights on how to use ASR technology for pronunciation teaching and what further support teachers need in order to use this technology confidently in their L2 English classes.

REFERENCES

ALBINI, A. B.; KLUGE, D. Professores de inglês da rede pública paranaense e o ensino da pronúncia. **Revista de Letras.** v. 14, 2013.

ASHWELL, T.; ELAM, J. R. How accurately can the google web speech API recognize and transcribe Japanese L2 English learners' oral production? **JALT CALL Journal**, v. 13, n. 1, p. 59–76, 2017.

BAKER, A. Exploring Teachers' Knowledge of Second Language Pronunciation Techniques: Teacher Cognitions, Observed Classroom Practices, and Student Perceptions. **TESOL Quarterly**, v. 48, n. 1, p. 136–163, 2014.

BALDISSERA, L. G.; TUMOLO, C. H. S. Apps for developing pronunciation in English as an L2. **Revista X**, v. 16, n. 5, p. 1355, 2021.

BRETT, D.; GONZÁLEZ-LLORET, M. Technology-Enhanced Materials. **The handbook of language teaching**, p. 351–369, 2009.

BRINTON, D. M. Innovations in pronunciation teaching. Em: **The Routledge handbook of contemporary English pronunciation**. Routledge, 2017. p. 448–461.

BURNS, A.; SEIDLHOFER, B. Speaking and pronunciation. Em: **An introduction to applied linguistics**. Routledge, 2020. v. 3rd Ed, p. 240–258.

BUSS, L. Pronunciation from the perspective of pre-service EFL teachers: An analysis of internship reports. **Pronunciation in Second Language Learning and Teaching Proceedings** 4. 2013.

BUSS, L. Beliefs and practices of Brazilian EFL teachers regarding pronunciation. **Language Teaching Research**, v. 20, n. 5, p. 619–637, 1 set. 2016.

BUSS, L. The role of training in shaping pre-service teacher cognition related to L2 pronunciation. **Ilha do Desterro**, v. 70, n. 3, p. 201–226, 2017.

CARDOSO, W. English syllable structure. Em: KANG, O.; THOMSON, R.; MURPHY, J. (Eds.). **The Routledge Handbook of Contemporary English Pronunciation**. 1st. ed. New York: Routledge, 2017. p. 122–136.

CARDOSO, W. Technology for speaking development. Em: DERWING, T. M.; MUNRO, M. J.; THOMSON, R. I. (Eds.). **The Routledge Handbook of Second Language Acquisition and Speaking**. 1st. ed. Routledge, 2022. p. 299–313.

CARLET, A.; KIVISTÖ-DE SOUZA, H. Improving l2 pronunciation inside and outside the classroom: Perception, production and autonomous learning of l2 vowels. **Ilha do Desterro**, v. 71, n. 3, p. 99–123, 2018.

CELCE-MURCIA, M.; BRINTON, D. M.; GOODWIN, J. M. **Teaching pronunciation: A reference for teachers of English to speakers of other languages**. Cambridge University Press, 1996.

CELCE-MURCIA, M.; BRINTON, D. M.; GOODWIN, J. M. **Teaching pronunciation hardback with audio CDs (2): A course book and reference guide**. 2. ed. Cambridge: Cambridge University Press, 2010.

CELCE-MURCIA, M.; BRINTON, D.; SNOW, M. A. **Teaching English as a second or foreign language**. Heinle ELT, 2014.

CHAPELLE, C. **English language learning and technology**. John Benjamins Publishing Company, 2003.

CHAPELLE, C. A. **Computer applications in second language acquisition**. Cambridge University Press, 2001.

CHAPELLE, C. A. Computer-Assisted Language Learning Effectiveness Research. Em: CHAPELLE, C. A. (Ed.). **The Encyclopedia of Applied Linguistics**. New York: Wiley Online Library, 2013.

CHEN, H. H. J. Developing and evaluating an oral skills training website supported by automatic speech recognition technology. **ReCALL**, v. 23, n. 1, p. 59–78, 2011.

CHUN, D. M. Computer-assisted pronunciation teaching. Em: **The Concise Encyclopedia of Applied Linguistics**. Wiley Online Library, 2020. p. 213–223.

CHUN, D.; SMITH, B.; KERN, R. Technology in Language Use, Language Teaching, and Language Learning. **Modern Language Journal**, v. 100, p. 64–80, 2016.

COLLINS, L.; MUÑOZ, C. The Foreign Language Classroom: Current Perspectives and Future Considerations. **Modern Language Journal**, v. 100, p. 133–147, 1 Feb. 2016.

COSTA, B. C. A. **Pronunciation Teaching Is Not a One-Size-Fits-All Endeavor: EFL Teachers' Beliefs and Classroom Practices**. Dissertation (master's) – Universidade Federal de Santa Catarina – Florianópolis. 2016.

CRESWELL, J. W.; GUETTERMAN, T. C. **Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research**. Pearson, 2019.

CUCCHIARINI, C.; NERI, A.; STRIK, H. Oral proficiency training in Dutch L2: The contribution of ASR-based corrective feedback. **Speech Communication**, v. 51, n. 10, p. 853–863, 2009.

CUCCHIARINI, C.; STRIK, H. Automatic Speech Recognition for second language pronunciation training. Em: **The Routledge handbook of contemporary English pronunciation**. Routledge, 2018. p. 556–569.

DAVIES, G. Computer-Assisted Language Education. Em: **Concise encyclopedia of applied linguistics**. Elsevier, 2006. p. 261–271.

DAVIES, G.; OTTO, S. E. K.; RÜSCHOFF, B. Historical perspectives on CALL. **Contemporary computer-assisted language learning**, p. 19–38, 2013.

DELATORRE, F.; GONCALVES, R.; SILVEIRA, R. The intelligibility of English verbs in the simple past tense: native and non-native speakers and Brazilian listeners. **Veredas**, v. 21, n. 2, p. 57–78, 2017.

DERWING, T. M. Utopian Goals for Pronunciation Teaching. **Proceedings from the 1st Conference of Pronunciation in Second Language Learning and Teaching**. October 2010, p. 24–37, 2010.

DIB, M. **Automatic Speech Recognition of Arabic Phonemes with Neural Networks: A Contrastive Study of Arabic and English**. Springer, 2019.

DIZON, G.; TANG, D. Intelligent personal assistants for autonomous second language learning: An investigation of Alexa. **JALT CALL Journal**, v. 16, n. 2, p. 107–120, 2020.

DÖRNYEI, Z.; TAGUCHI, T. **Questionnaires in second language research: Construction, administration, and processing**. Routledge, 2009.

DUCATE, L.; ARNOLD, N. CALL: Where are we and where do we go from here. Em: **Calling on CALL: From theory and research to new directions in foreign language teaching**. v. 5p. 1–20. San Marcos (Texas): Calico, 2006.

EVERLY, P. Expanding pronunciation instructional time beyond the classroom: Microsoft Office 2016 OneNote Class Notebook as an interactive delivery platform. **TESOL journal**, v. 10, n. 2, p. e00421- n/a, 2019.

FOOTE, J. A.; MCDONOUGH, K. Using shadowing with mobile technology to improve L2 pronunciation. **Journal of Second Language Pronunciation**, v. 3, n. 1, p. 34–56, 2017.

FOUZ-GONZÁLEZ, J. Trends and directions in computer-assisted pronunciation training. **Investigating English pronunciation: Trends and directions**, p. 314–342, 2015.

GIL, A. C. **Como Elaborar Projetos de Pesquisa**. 4 eds. São Paulo: Atlas, 2002.

GOLONKA, E. M. et al. Technologies for foreign language learning: A review of technology types and their effectiveness. **Computer Assisted Language Learning**, v. 27, n. 1, p. 70–105, 2014.

GOMES JUNIOR, R. C.; SILVA, L. DE O.; PAIVA, V. L. M. DE O. E. Tecnologias digitais para aprender e ensinar inglês no Brasil. **Texto Livre**, v. 15, p. e38008, 2022.

GONÇALVES, A. R.; SILVEIRA, R. Intelligibility research in Brazil: empirical findings and methodological issues. **Revista Horizontes de Linguística Aplicada**, v. 14, n. 1, p. 51–81, 2015.

GOOGLE. **Ten years of Google Translate**. Disponível em: <https://blog.google/products/translate/ten-years-of-google-translate/>. Acesso em: 26 abr. 2022.

GOTTARDI, W.; ALMEIDA, J. F. DE; TUMOLO, C. H. S. Automatic speech recognition and text-to-speech technologies for L2 pronunciation improvement: reflections on their affordances. **Texto livre**, v. 15, 2022.

HARMER, J. **The practice of English language teaching**. 5th edition. Pearson Longman, 2015.

HE, Y.; CARDOSO, W. Can online translators and their speech capabilities help English learners improve their pronunciation? Em: **CALL and professionalisation: short papers from EUROCALL 2021**. Research-publishing.net, 2021. p. 126–131.

HUBBARD, P. Evaluating CALL software. **Calling on CALL: From theory and research to new directions in foreign language teaching**, p. 313–338, 2006.

INCEOGLU, S.; LIM, H.; CHEN, W. H. Asr for EFL pronunciation practice: Segmental development and learners' beliefs. **Journal of Asia TEFL**, v. 17, n. 3, p. 824–840, 2020.

JAMIESON, J.; CHAPELLE, C. A. Evaluating CALL use across multiple contexts. **System (Linköping)**, v. 38, n. 3, p. 357–369, 2010.

JORDÃO, C. M. ILA-ILF-ILE-ILG: Quem dá conta? EAL-ELF-EFL-EGL: Same Difference? **Revista brasileira de linguística aplicada**. V. 14, p. 13-40, 2014.

JURAFSKY, D.; MARTIN, J. H. **Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition**: Unpublished manuscript. USA: 2021. Disponível em: <https://web.stanford.edu/~jurafsky/slp3/ed3book_sep212021.pdf>.

KIM, I. S. Automatic speech recognition: Reliability and pedagogical implications for teaching pronunciation. **Educational Technology and Society**, v. 9, n. 1, p. 322–334, 2006.

KIVISTÖ-DE SOUZA, H.; GOTTARDI, W. How well can ASR technology understand foreign-accented speech?. **Trabalhos em Linguística Aplicada**, v. 61, n. 3, p. 764–781, Dez. 2022.

LAI, C. **Autonomous language learning with technology: Beyond the classroom**. Bloomsbury Publishing, 2018.

LANE, L.; BROWN, H. D. **Tips for teaching pronunciation: A practical approach**. Pearson Longman, 2010.

LEAVY, P. **Research design: Quantitative, qualitative, mixed methods, arts-based, and community-based participatory research approaches**. Guilford Publications, 2017.

LEVIS, J. Revisiting the Intelligibility and Nativeness Principles. **Journal of Second Language Pronunciation**, v. 6, n. 3, p. 310–328, 2020.

LEVIS, J. M. **Intelligibility, Oral Communication, and the Teaching of Pronunciation**. Cambridge University Press, 2018.

LEVIS, J.; SUVOROV, R. Automatic Speech Recognition. Em: CHAPELLE, C. A. (Ed.). **The encyclopedia of applied linguistics**. New York: Wiley-Blackwell, 2013. p. 316–323.

LI, J. et al. **Robust automatic speech recognition: a bridge to practical applications**. Academic Press, 2016.

LIAKIN, D.; CARDOSO, W.; LIAKINA, N. Learning L2 pronunciation with a mobile speech recognizer: French/y/. **CALICO Journal**, v. 32, n. 1, p. 1–25, 2015.

LIAKIN, D.; CARDOSO, W.; LIAKINA, N. Mobilizing Instruction in a Second-Language Context: Learners' Perceptions of Two Speech Technologies. **Languages**, v. 2, n. 3, p. 11, 2017.

LIAKINA, N.; LIAKIN, D. Speech technologies and pronunciation training: What is the potential for efficient corrective feedback? Em: **Second Language Pronunciation**. De Gruyter Mouton, 2022. p. 287–312.

MARTINS, C. B. M. J.; MOREIRA, H. O campo CALL (Computer Assisted Language Learning): definições, escopo e abrangência. **Calidoscópio**, v. 10, n. 3, p. 247–255, 2012.

MCCROCKLIN, S. Automatic Speech Recognition: Making It Work for Your. **Pronunciation in Second Language Learning and Teaching Proceedings 6**. 2018.

MCCROCKLIN, S. Dictation programs for second language pronunciation learning: Perceptions of the transcript, strategy use and improvement. **Konińskie Studia Językowe**. v. 7, n. 2, p. 137–157, 2019.

MCCROCKLIN, S.; EDALATISHAMS, I. Revisiting Popular Speech Recognition Software for ESL Speech. **TESOL Quarterly**, v. 54, n. 4, p. 1086–1097, 2020.

MCCROCKLIN, S.; HUMAIDAN, A.; EDALATISHAMS, I. ASR dictation program accuracy: Have current programs improved? **Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference**, n. June, p. 191–200, 2019.

MCCROCKLIN, S. M. Dictation programs for pronunciation learner empowerment. **Proceedings of the 5th pronunciation in second language learning and teaching conference.**, n. September, p. 30–39, 2014.

MCCROCKLIN, S. M. Pronunciation learner autonomy: The potential of Automatic Speech Recognition. **System**, v. 57, n. April 2016, p. 25–42, 2016.

MICROSOFT WORD ONLINE. **Transcribe your recordings**. Disponível em: < https://support.microsoft.com/en-us/office/transcribe-your-recordings-7fc2efec-245e-45f0-b053-2a97531ecf57/>. Acesso em: 22 fev. 2023.

MOUSSALLI, S.; CARDOSO, W. Intelligent personal assistants: can they understand and be understood by accented L2 learners? **Computer Assisted Language Learning**, v. 33, n. 8, p. 865–890, 2020.

MROZ, A. Seeing how people hear you: French learners experiencing intelligibility through automatic speech recognition. **Foreign Language Annals**, v. 51, n. 3, p. 617–637, 2018.

MUNRO, M. J. Foreign accent and speech intelligibility. **Phonology and second language acquisition**, v. 5, p. 193–218, 2008.

MUNRO, M. J.; DERWING, T. M. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. **Language learning**, v. 45, n. 1, p. 73–97, 1995.

MUNRO, M. J.; DERWING, T. M. Intelligibility in research and practice: Teaching priorities. Em: REED, M.; LEVIS, J. M. (Eds.). **The Handbook of English Pronunciation**. Wiley Online Library, 2015. p. 375–396.

NERI, A.; CUCCHIARINI, C.; STRIK, H. ASR-based corrective feedback on pronunciation: Does it really work? **INTERSPEECH 2006 - ICSLP**, v. 4, n. May 2014, p. 1982–1985, 2006.

NERI, A.; CUCCHIARINI, C.; STRIK, H. The effectiveness of computer-based speech corrective feedback for improving segmental quality in L2 Dutch. **ReCALL**, v. 20, n. 2, p. 225–243, 2008.

O'BRIEN, M. G. et al. Directions for the future of technology in pronunciation research and teaching. **Journal of Second Language Pronunciation**, v. 4, n. 2, p. 182–207, 2018.

PENNINGTON, M. C.; ROGERSON-REVELL, P. **English pronunciation teaching and research**. v. 10, p. 978–988, Londres: Palgrave Macmillan, 2019.

PERNA, C. B. L.; DELGADO, H. O. K.; SILVA, A. D. C. Successful digital resources to enhance English lessons. **Ilha do Desterro**, v. 74, p. 445–461, 2022.

QUESTIONNAIRE. **Dicionário Online Cambridge**, 28 dez. 2022. Cambridge: Cambridge University Press. Disponível em: <https://dictionary.cambridge.org/dictionary/english/questionnaire>. Acesso em: 28 dez. 2022.

REVELL-ROGERSON, P. M. Computer-Assisted Pronunciation Training (CAPT): Current Issues and Future Directions. **RELC Journal**, v. 52, n. 1, p. 189–205, 2021.

ROCCAMO, A. Effective pronunciation instruction in basic language classrooms: A modular approach. Em: LEVIS, J. M.; MCCROCKLIN, S. M. (Eds.). **Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference**. IA: Iowa State University, 2014.

ROGERSON-REVELL, P. **English phonology and pronunciation teaching**. Bloomsbury Publishing, 2011.

SICOLA, L.; DARCY, I. Integrating Pronunciation into the Language Classroom. Em: **The handbook of English pronunciation**. Wiley Online Library, 2015. p. 471.

SILVEIRA, R. et al. Percepção, produção e inteligibilidade do inglês falado por usuários brasileiros. **Perspectivas atuais de aprendizagem e ensino de línguas. Florianópolis: LLE/CCE/UFSC**, p. 237–283, 2017.

SILVEIRA, R.; ZANCHET, C. E.; PEREIRA, M. H. Affordances of digital technology for English pronunciation teaching: The perspective of Brazilian teachers. **Veredas-Revista de Estudos Linguísticos**. n. 2, v. 26. 2022.

SLABAKOVA, R. **Second language acquisition**. Oxford University Press, 2016.

SOLEIMANI, H. **Computer Assisted Language Learning: Theory and Practice**. Payame Noor University, 2021.

STANLEY, G. **Language learning with technology: Ideas for integrating technology in the classroom**. Cambridge University Press, 2013.

SUE, V. M.; RITTER, L. A. **Conducting the surveys**. **Conducting Online Surveys**. USA: SAGE Publications, 2007.

SURVEY. **Dicionário Online Cambridge**, 28 dez. 2022. Cambridge: Cambridge University Press. Disponível em: <https://dictionary.cambridge.org/dictionary/english/survey>. Acesso em: 28 dez. 2022.

SUVOROV, R.; CHAPELLE, C. A. Computer-Assisted Language Learning Effectiveness Research. Em: **The Concise Encyclopedia of Applied Linguistics**. Wiley Online Library, 2020. p. 208–213.

TUMOLO, C. Digital resources and the learning of English as a foreign language. **Ilha do Desterro**, n. 66, p. 203–238, 2014.

TUMOLO, C. H. S.; FINARDI, K. R. DIGITAL RESOURCES IN ENGLISH AS L2: DESIGNS AND AFFORDANCES. **Ilha do Desterro**, n. 3, v. 74, 2021.

VAN LIESHOUT, C. VAN; CARDOSO, W. Google Translate as a Tool for Self-Directed Language Learning. **Language Learning & Technology**. v. 26, n. 1, p. 1–19, 2022.

WALESIAK, B. Mobile apps for pronunciation training. Em: **English Pronunciation Instruction: Research-based insights**. v. 19, p. 357. John Benjamins Publishing Company, 2021.

WALLACE, L. Using Google Web speech as a springboard for identifying personal pronunciation problems. **Proceedings of the 7th annual Pronunciation in Second Language Learning and Teaching Conference**, v. 2015 p. 180–186, 2016.

YOSHIDA, M. T. Choosing technology tools to meet pronunciation teaching and learning goals. **The CATESOL Journal**, v. 30, n. 1, p. 195–212, 2018.

YU, D.; DENG, L. **Automatic Speech Recognition A Deep Learning Approach**. London: Springer, 2015.

ZIMMER, M.; SILVEIRA, R.; ALVES, U. K. **Pronunciation Instruction for Brazilians: Student's Book**. Cambridge Scholars Publishing, 2009.

APPENDIX A – Term of consent

UFSC - Universidade Federal de Santa Cataria
CCE - Centro de Comunicação e Expressão
PPGI - Programa de Pós-Graduação em Inglês: Estudos Linguísticos e Literários

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Prezado(a) participante,

Meu nome é **William Gottardi** e sou mestrando no Programa de Pós-Graduação em Inglês (PPGI) da Universidade Federal de Santa Catarina (UFSC). Junto com a minha orientadora, Prof. Dra. Rosane Silveira (DLLE/CCE), iremos realizar uma pesquisa científica intitulada: *Reconhecimento Automático da Fala como recurso no ensino de pronúncia: quais as percepções dos professores sobre esta tecnologia*?

Este documento tem como objetivo esclarecer como a pesquisa será realizada para que o(a) participante possa decidir se deseja contribuir voluntariamente com a coleta de dados, além de apresentar todas as informações legais e necessárias. Como a pesquisa envolve etapas virtuais, seus procedimentos obedecem às orientações dispostas no Ofício Circular No 2/2021/CONEP/SECNS/MS. Caso o(a) participante aceite o convite livremente e autorize sua participação na pesquisa, faz-se necessário a leitura cuidadosa desse documento e a sua assinatura no final.

1) OBJETIVOS DA PESQUISA:

O principal objetivo dessa pesquisa é investigar o uso da tecnologia de Reconhecimento Automático da Fala para o ensino de pronúncia pela perspectiva de professores de inglês.

2) PROCEDIMENTO DE COLETA DE DADOS

Antes de terem acesso aos instrumentos de coleta de dados, os participantes assinarão um termo de consentimento (adaptado em formato de formulário digital) no qual afirmam que são **maiores de 18 anos de idade**. É importante ressaltar que apenas participantes maiores de 18 poderão participar da pesquisa. A coleta de dados se fará por meio dos seguintes instrumentos:

a) um questionário em formato *Google Forms* contendo este termo de consentimento e indicando os procedimentos para se registrar à oficina online ofertada por esta pesquisa;
b) uma oficina online de sessão única, por meio da plataforma Zoom, sobre o uso da tecnologia de Reconhecimento Automático de Fala para o ensino de pronúncia com duração de **duas horas**. A oficina será apresentada em um sábado de manhã. Serão esperados pelo menos vinte participantes e um máximo de cinquenta. A oficina será gravada e o registro do bate-papo salvo. Os arquivos ficarão guardados sob responsabilidade dos pesquisadores e em total sigilo, sendo somente utilizados para conferência futura de acordo com os objetivos desta pesquisa.
c) um segundo questionário em formato *Google Forms* que será disponibilizado aos participantes logo após a oficina. O questionário incluirá perguntas que podem ajudar a desvendar as possibilidades de uso da tecnologia de Reconhecimento Automático de Fala para o ensino de pronúncia e como aplicá-la em diferentes contextos escolares. Além disso, este questionário busca compreender as experiências docentes dos participantes e o contexto escolar onde eles estão inseridos, bem como suas impressões sobre a oficina online.

3) POSSÍVEIS RISCOS E DESCONFORTOS DA PESQUISA

Ainda que os procedimentos de coleta não utilizem nenhuma instrumentação invasiva, o estudo pode oferecer o risco de causar ansiedade e desconforto pelo cansaço físico e mental; constrangimento ou aborrecimento ao participar das tarefas propostas ao longo da oficina e ao responder ao questionário após o fim da oficina. Para minimizar essa situação, você poderá optar por fazer pequenas pausas durante o preenchimento do questionário e sinalizar quaisquer desconfortos durante as atividades da oficina. Além disso, como as tarefas serão feitas de forma on-line, o(a) participante estará exposto(a) aos riscos característicos do contato remoto por meio eletrônicos e ambientes virtuais e precisará, porventura, lidar com limitações tecnológicas.

4) POSSÍVEIS BENEFÍCIOS, GANHOS E RESULTADOS DA PESQUISA

A sua participação é voluntária e não trará qualquer benefício direto, mas proporcionará um conhecimento mais amplo sobre o ensino de pronúncia e o uso da tecnologia de Reconhecimento Automático de Fala para o ensino de pronúncia. Os pesquisadores oferecerão ao participante uma oficina sobre ensino de pronúncia com o uso da tecnologia de Reconhecimento Automático de Fala além de compartilhar com os participantes os planos de atividades utilizando tal tecnologia, o que poderá auxiliá-lo nas suas atividades docentes futuras.

5) COMPENSAÇÃO FINANCEIRA

O participante não será remunerado pela sua participação na pesquisa. Ainda que a pesquisa não preveja gastos com aquisição de equipamento, deslocamento ou alimentação, o participante tem direto a ressarcimento e indenização caso haja algum dano material ou imaterial ocasionado pela pesquisa, e esse seja devidamente comprovado. Este documento garante o reparo ao dano que deve ser pago de acordo com a Resolução 510/16.

6) DESISTÊNCIA DA PARTICIPAÇÃO NA PESQUISA:

Caso o(a) participante não queira continuar a participação na pesquisa ou que os dados coletados não sejam usados, não há nenhum problema. A desistência pode ocorrer a qualquer momento, sem qualquer prejuízo para o participante ou para a pesquisa. Basta entrar em contado com os pesquisadores através dos números de telefone ou e-mail informados no **item 8** desse documento.

7) CONFIDENCIALIDADE

Este documento garante a confidencialidade da identidade e informações privadas do(a) participante. Ou seja, a garantia de que as informações privadas estão protegidas e confiadas aos pesquisadores, que tomarão todas as providências necessárias para manter o sigilo. Tais dados não serão revelados sem as devidas autorizações. Porém, sempre existe a possibilidade da quebra de sigilo, mesmo que não intencional ou/e involuntária, cujas consequências serão tratadas nos termos da lei. Durante a intervenção do estudo, que será uma oficina online ministrada ao vivo, os participantes terão a opção de manter as suas câmeras desligadas e usar um pseudônimo a fim de não terem suas identidades reveladas ao interagir com os demais participantes. É importante ressaltar que os resultados deste estudo poderão ser publicados em revistas científicas ou apresentados em congressos científicos, sem que a identidade do participante seja revelada. Os pesquisadores do estudo declaram ainda conhecer e cumprir os requisitos da Lei Geral de Proteção de Dados (Lei No 13.709, de 14 de agosto de 2018) quanto ao tratamento de dados pessoais e dados pessoais sensíveis que serão utilizados para a execução da presente pesquisa.

8) ASSISTÊNCIA, CONTATOS E ENDEREÇO DOS PESQUISADORES:

Ao longo da pesquisa, o(a) participante receberá o acompanhamento e assistência necessários caso haja alguma dúvida ou problema. Informamos abaixo os contatos e endereço dos pesquisadores em caso de dúvidas e para mais informações.

Pesquisador: William Gottardi
Celular/ WhatsApp: (47) 99278-3966
E-mail: william.gottardi@posgrad.ufsc.br
Endereço do Programa de Pós-Graduação em Inglês (PPGI)

Centro de Comunicação e Expressão – CCE "B" Sala 313
Campus Universitário – Trindade – Florianópolis – SC
CEP: 88.040-900
Orientadora: Dra. Rosane Silveira
Celular/ WhatsApp: (48) 9615-9978
E-mail: rosane@cce.ufsc.br

9) CEPSH – UFSC E RESOLUÇÃO 510/16:

De acordo com o trecho disponível no site da CEPSH (Comitê de Ética em Pesquisa com Seres Humanos), o comitê "é um órgão colegiado interdisciplinar, deliberativo, consultivo e educativo, vinculado à Universidade Federal de Santa Catarina, mas independente na tomada de decisões, criado para defender os interesses dentro de padrões éticos". A CEPSH-UFSC se encontra no Prédio Reitoria II, 4º andar, sala 701, localizado na Rua Desembargador Vitor Lima, nº 222, Trindade, Florianópolis. Telefone para contato: 3721 – 6094.

Além disso, declaro que conduzirei a pesquisa de acordo com o que preconiza a Resolução 510/16, que dispõe sobre as normas aplicáveis a pesquisas em Ciências Humanas e Sociais, e se encontra no site da CEPSH – UFSC (http://cep.ufsc.br/).

Ao concordar com este documento de assentimento esclarecido e livre por meio deste formulário eletrônico, você está aceitando participar da pesquisa. Se você estiver de acordo, assinale a opção **SIM** abaixo para prosseguir com o registro e coleta de informações pessoais.

Eu declaro que li este documento e obtive dos pesquisadores todas as informações que julguei necessárias para me sentir esclarecido e livre em participar da pesquisa **Reconhecimento Automático da Fala como recurso no ensino de pronúncia: quais as percepções dos professores sobre esta tecnologia?**.

Você concorda com o termo acima?

( ) sim

( ) não

APPENDIX B – Background questionnaire

UFSC - Universidade Federal de Santa Cataria
CCE - Centro de Comunicação e Expressão
PPGI - Programa de Pós-Graduação em Inglês: Estudos Linguísticos e Literários

Background Questionnaire

Hello everyone, and thanks for being part of this study!

Please, **always read the descriptions** carefully to answer the questions appropriately.

First, you will need to accept the consent form in order to be part of this study. After agreeing to it, you will answer some personal questions to collect demographic information for this study. Then, you will be eligible for the workshop session on **November 5th (Saturday) at 9am (**Brasília Time - UTC/GMT -3 hours**)**. You will receive a confirmation of your registration by email. **Remember**, always check your junk e-mail!

Please, add this e-mail address to your contact list so it will not be considered as a spam e-mail that easily: **william.gottardi@posgrad.ufsc.br**

You can answer this form in Portuguese if you prefer!

Thank you!

Section I - Personal Information:
You will share some **demographic characteristics** such as gender, age, residential location, occupation, and other relevant background information for this study.

Name: _____.

E-Mail: _____.

Age: _____.

Gender: ( ) Female / ( ) Male / () Other: _____.

City where you live: _____.

State where you live: _____.

Country where you live: _____.

Which undergraduate course(s) have you taken/are you taking (write the name of the course(s), the institution(s), and the (expected) year of completion)? _____.

Have you taken/are you taking any postgraduate course (write the name of the course, the institution and the (expected) year of completion)? _____.

Years of experience teaching English: _____.

In what sector do you work? ( ) Public / ( ) Private / ( ) Both

Mark the level(s) of education you are currently working in:

Young Learners (Educação infantil)
Bilagual Education (Educação bilíngue)
Elementary School I (Ensino Fundamental I)
Elementary School II (Ensino Fundamental II)
High School (Ensino médio)
Technical/ Vocational (Curso técnico ou profissionalizante)
Higher Education (Ensino Superior)
Language School (Curso de idiomas)
Private teacher (Professor particular)
Other

Section II - Information about your Teaching Practices:

You will answer some **behavioral questions** to find out more about your experience with **digital technologies and pronunciation teaching**.

How confident are you in teaching pronunciation?

(not much) 0 – 10 (very much)

How often do you teach pronunciation in your classes?

(rarely) 0 – 10 (very frequently)

How confident are you in using digital resources (computer, cell phone, projector, websites, apps) in your pedagogical practices?

(not much) 0 – 10 (very much)

How often do you use digital resources in your classes?

(rarely) 0 – 10 (very frequently)

How often do you use Google Translate in your classes?

(rarely) 0 – 10 (very frequently)

How often do you use any Automatic Speech Recognition (ASR) tool as a resource for teaching pronunciation?

(rarely) 0 – 10 (very frequently)

Thank you!

Thanks for answering this questionnaire! Information regarding the **workshop** session and the **Zoom link** to access it will be sent to **your e-mail** as soon as the registration form stops receiving responses. Keep an eye on your spam folder, just in case. You can contact me via e-mail in case of any question. Here it is: william.gottardi@posgrad.ufsc.br

APPENDIX C – Survey

UFSC - Universidade Federal de Santa Cataria
CCE - Centro de Comunicação e Expressão
PPGI - Programa de Pós-Graduação em Inglês: Estudos Linguísticos e Literários

Survey

Hello Everyone!

Thanks for joining the workshop and contributing to this research!

This survey was designed to collect data related to the workshop you have joined and its instruments. It is divided into 4 sections. It should not take longer than 30 minutes to answer all of them carefully. Please, take your time and reflect upon each question so the data is consistent and reliable.

Remember: you should submit your answers by **November 14th.**

Please, always **read the descriptions** carefully to answer the questions appropriately.
You can answer this form in **Portuguese** if you prefer!

Thank you again!

Full Name: _____.

Confirm your e-mail: _____.

Section I - Teacher Development Needs:
The questions address your needs for **further education** and **professional development**.

1- In which area do you need more courses/workshops to make you feel more comfortable teaching pronunciation? (you can check more than one answer)

( ) Theory
( ) Teaching techniques
( ) Digital resources
( ) None
( ) Others: _____

2- I feel comfortable to use ASR to teach pronunciation after attending the workshop.
(Strongly disagree) 0 – 10 (Strongly agree)

3- I would use ASR to improve my own pronunciation.
(Strongly disagree) 0 – 10 (Strongly agree)

Section II - ASR Accessibility:
The questions address the **suitability** and **accessibility** of the presented ASR tool considering your teaching context.

1- It is easy to use Google Translate's voice recognition feature.
(Strongly disagree) 0 – 10 (Strongly agree)

2- My students have access to Google Translate to practice pronunciation.
(Strongly disagree) 0 – 10 (Strongly agree)

3- In which context(s) would you use ASR to teach pronunciation? (you can check more than one answer)
( ) online
( ) face-to-face
( ) hybrid (online and face-to-face)
( ) none

4- Which of the ASR possible limitations is difficult to overcome considering your teaching context? (you can check more than one answer)
( ) Access to stable internet connection
( ) Access to a personal computer (PC) or mobile devices (cell phones, tablets, and laptops)
( ) Access to microphones/ headphones with microphone
( ) External noise.
( ) Helping students interpret the orthographic feedback provided by the ASR
( ) None
( ) Other: _____

Section III - ASR Affordances:
The questions address the **affordances** of ASR technology for **pronunciation teaching and learning**.

1- ASR facilitates the teaching of pronunciation.
(Strongly disagree) 0 – 10 (Strongly agree)

2- My students would be interested in using ASR to improve their own pronunciation.
(Strongly disagree) 0 – 10 (Strongly agree)

3- I would use ASR as a complementary tool for teaching pronunciation in my classes.
(Strongly disagree) 0 – 10 (Strongly agree)

4- ASR can encourage/motivate learners to produce more output outside the classroom.
(Strongly disagree) 0 – 10 (Strongly agree)

5- ASR transcription (orthographic feedback) can be beneficial to the development of learner's pronunciation.
(Strongly disagree) 0 – 10 (Strongly agree)

6- I would use ASR for pronunciation teaching as: (you can check more than one answer)

( ) an auxiliary resource for the regular classes
( ) an out of class supplement/ homework
( ) a learning tool suggestion for self-studying
( ) a pronunciation self-assessment tool suggestion
( ) I would not use it
( ) Other: _____

Section IV - Activities Appraisal:
The questions address the **design, suitability, and adaptability** of the activities presented during the workshop. You can check the activities on the **Teacher's handout** file (*page 2*) available here: https://bit.ly/LINK.

1- The activities have a clear and reasonable learning focus.
(Strongly disagree) 0 – 10 (Strongly agree)

2- The activities have addressed linguistic features related to pronunciation (ex. minimal pairs, vowel quality, segments, prosody, accuracy, fluency).
(Strongly disagree) 0 – 10 (Strongly agree)

3- The activities have an appropriate degree of difficulty for the learners considering the target level indicated by the researcher during the workshop.
(Strongly disagree) 0 – 10 (Strongly agree)

4- The activities can be adapted according to learners' individual characteristics (e.g., age, computer experience, learning style).
(Strongly disagree) 0 – 10 (Strongly agree)

5- The activities allow learners to use the language for interpreting and constructing meaning.
(Strongly disagree) 0 – 10 (Strongly agree)

6- The activities direct learner's attention primarily toward the meaning of the language.
(Strongly disagree) 0 – 10 (Strongly agree)

7- The activities show a strong correspondence between the tasks and what learners would expect to see outside the classroom.
(Strongly disagree) 0 – 10 (Strongly agree)

8- The activities address real pronunciation difficulties faced by Brazilian learners of English.
(Strongly disagree) 0 – 10 (Strongly agree)

9- The activities can help learners develop autonomous learning strategies.
(Strongly disagree) 0 – 10 (Strongly agree)

10- The activities can offer a positive teaching/learning experience with the digital technology used (Google Translate).
(Strongly disagree) 0 – 10 (Strongly agree)

11- The activities can be implemented considering the particular constraints of my class and language program.

(Strongly disagree) 0 – 10 (Strongly agree)

12- The activities are adaptable to my teaching context in relation to the **equipment** used (e.g., microphone, earphone, computer, laptop, cellphone).

(Strongly disagree) 0 – 10 (Strongly agree)

13- The activities are adaptable to my teaching context in relation to the **infrastructure** required (e.g., internet, wi-fi, computer lab).

(Strongly disagree) 0 – 10 (Strongly agree)

14- Which activity(ies) presented during the workshop would you use with your students? If you want to recall any of them, just access them through this link: *activities_workshop*. (you can check more than one answer)

( ) Activity 1 – Pronunciation Self-assessment
( ) Activity 2 – Vowel Contrast
( ) Activity 3 – Shadow Reading
( ) Activity 4 – Paragraph-reading Task
( ) Activity 5 – Tongue Twisters
( ) Activity 6 – Monitoring Worksheet
( ) Activity 7 – Role-play Activity
( ) None

15- Among all the proposed activities, which one was your **favorite**? (dropdown list / only one answer is possible)

( ) Activity 1 – Pronunciation Self-assessment
( ) Activity 2 – Vowel Contrast
( ) Activity 3 – Shadow Reading
( ) Activity 4 – Paragraph-reading Task
( ) Activity 5 – Tongue Twisters
( ) Activity 6 – Monitoring Worksheet
( ) Activity 7 – Role-play Activity
( ) I did not like any of them

Thank you!

Thanks for answering this survey! You can contact me via e-mail in case of any question or suggestion. Here it is: william.gottardi@posgrad.ufsc.br

Could you please provide some feedback regarding the workshop itself (topic division, the way it was conducted, clarity, program, etc.)?

(Long answer text / Paragraph)

APPENDIX D - Workshop's detailed program

Topic	Description	Duration* PS	Duration* MS
Introduction	Introduce myself, present my project overview and objectives. Share the handout file with the participants.	10	15
Review of the literature	• Pronunciation teaching overview. • The communicative framework for teaching English pronunciation (CELCE-MURCIA; BRINTON; GOODWIN, 2010). • Criteria for CALL task appropriateness (CHAPELLE, 2001). • ASR Technology: ASR brief definition, limitations, and possible use. Present some studies related to ASR-based pronunciation practice. • ASR and Pronunciation Teaching: speech intelligibility improvement, orthographic visual feedback, and the role of output for second language acquisition. • Rationale behind the designed ASR Activities that will be presented to the participants. (GONÇALVES; SILVEIRA, 2015; MUNRO; DERWING, 2015; SILVEIRA et al., 2017; ZIMMER; SILVEIRA; ALVES, 2009).	15	15
Using Google Translate's ASR tool	• Explain how to use Google Translate's ASR tool. • Address some possible technical problems and system requirements (allow microphone use, set-up the headset, etc.) • Show participants examples on how to dictate using Google Translate. • Allow participants to experience how to use Google Translate's ASR tool by their own.	10	10
Hands-on: ASR-based Pronunciation Activities	• Participants will have the opportunity to use the ASR tool following the pronunciation activities suggested in the handout. • Participants can ask questions orally or comment in writing (using the chat) during the workshop.	40	70
Coffee break		5	15
Brainstorming: ASR for pronunciation teaching	• Participants are encouraged to debate the affordances of ASR for pronunciation teaching and ask further questions to enrich the discussion.	30	45

	• Participants can share their opinions orally or comment in writing (using the chat). • Participants will be divided into different breakout rooms to have more opportunities for discussion and interaction.		
Discussion and final remarks	• The main points covered throughout the workshop will be revisited. • Further reading suggestions will be provided. • Instruction on how to fill out the survey will be provided. • Participants can ask further questions and make comments orally or in writing (using the chat).	10	10
Total		**120 (2 hours)**	**180 (3 hours)**

Source: author
*Time expressed in minutes.

APPENDIX E – ASR-based pronunciation activities

Activity 1 – Pronunciation Self-assessment

Learning Focus	Pronunciation self-assessment
Level	A1 – A2
Time	30-60 minutes
Target Feature	Raising awareness about pronunciation patterns that might disrupt communication
Communicative Framework Stage	Guided Practice (4) with an information-gap exercise
Dictation Passage	Appendix A

Procedure

1. Instruct students to access Google Translate (GT) and give them the *Personal Information Sheet* (Appendix A).
2. Tell students to fill in the gaps with their personal information. A hint word is beside each gap to help students understand the kind of information that is required. Motivate students to ask questions regarding vocabulary at this step so they can focus on pronunciation next.
3. Then, ask students to voice type each sentence from the Personal Information Sheet at a time and write down the resulting transcription provided by GT in the second column of the sheet. They can repeat this procedure if some interference has happened during the voice typing.
4. Encourage students to underline and take notes of any words and phrases that GT transcribed differently from their original sentence.

    ⇒ During this process, students may require assistance in improving their pronunciation to be understood by the application. GT's text-to-speech feature or an online dictionary (www. dictionary.cambridge.org) can serve students as the target model for the words that were not accurately transcribed by GT.

5. After completing all the rows, students can practice the most difficult sentences or create new sentences about their personal information using the two last rows.

Variation: this activity can be used as *homework*. However, it may be necessary to suggest the use of GT's text-to-speech feature or an online dictionary to serve students as a model for difficult words.

Appendix A- Personal Information Sheet:

Student: _____ Date: _____ **Tell me about you!**	
Sentence	**Google Translate Transcription**
I am _____ [number] years old.	
I have _____ [number] brother(s). / I don't have a brother. I have _____ [number] sister(s). / I don't have a sister.	
My house is _____ [color].	
My favorite color is _____ [color].	
I have a pet. My pet is a(n) _____ [animal]. / I don't have a pet.	
My favorite food is _____ [food].	
My favorite drink is _____ [drink].	
?	
?	

Activity 2 – Vowel Contrast

Learning Focus	Raising awareness about acoustic features of each vowel and how to distinguish them in production.
Level	A2 and above
Time	30 – 45 minutes
Target Feature	Producing the vowel contrast / ɪ / and / iː/
Communicative Framework Stage	Controlled Practice (3) with sentence-level minimal pairs
Dictation Passage	Look at that **sheep**! Look at that **ship**! I may **live**. I may **leave**. It is a **sin**. It is a **scene**. Can you see the **beach**? Can you see the **bitch**? How can I spell "**feel**"? How can I spell "**fill**"? What is the definition of "**sit**"? What is the definition of "**seat**"?

Procedure

1. Instruct students to access Google Translate (GT) and show them the dictation passage.
2. Ask students to voice type each sentence from the dictation passage and pay close attention to the words in bold (minimal pairs). Explicit instruction on how the two vowels differ acoustically and articulatorily may be necessary at this point to help students raise awareness regarding these contrasting sounds.
3. Encourage students to repeat the voice typing if GT transcribes the word in bold differently from the original sentence.

4. Tell students they can listen to the target words (in bold) using GT's text-to-speech feature or an online dictionary and try to imitate the sounds during their production.

Variation 1: this activity can be used as *homework*. A video can be used to give explicit instructions regarding the vowel contrast. Example: https://youtu.be/FYI6Vt3uq7s.

Variation 2: different vowel contrast can be practiced using these procedures. However, the sentences should be carefully tested using GT before suggesting them to the students.

Activity 3 – Shadow Reading

Learning Focus	Practicing speech fluency and accuracy
Level	B2 and above
Time	30 – 45 minutes
Target Feature	Suprasegmental features and the connected speech phenomena
Communicative Framework Stage	Controlled Practice (3) with a shadow reading activity
Dictation Passage	Any text passage.
Technical requirements	Earphones/ headphones are **indispensable** for this activity

Procedure

1. Instruct students to access GT on **two** different tabs on their browser. Students **must wear earphones/ headphones** during the voice typing step.
2. On the first tab, students will hand type or paste a previously copied text into GT's text field, right below the suggested dictation passage. On the second tab, students will voice type the text passage.

⇒ The target text passage should preferably have between 200 to 500 characters so it will be easier to annotate the transcription later. Note that Google Translate (GT) counts the characters just below the text field (see appendix A).

3. Before beginning the voice typing step, explain how the *shadow reading* technique works. This technique consists of the student repeating either along with or slightly after someone's speech (input) (CELCE-MURCIA; BRINTON; GOODWIN, 2010). In this activity, the input will be provided by GT's text-to-speech feature (TTS) from the first tab and students should repeat the passage slightly after it.
4. First, students should go to the second tab and click on the "translate by voice" button (see appendix A) to start the voice typing feature. Next, they should go to the first tab and click on the "listen" button to reproduce the audio input. Finally, they can read the text passage along with the TTS.
5. After finishing voice typing the passage, tell students to write down the words and phrases transcribed by GT that differ from the original passage on the first tab.
6. Then, students should repeat each word/ phrase at a time, trying to be intelligible to GT. During this process, students may require assistance in improving their pronunciation to be understood by the application; therefore, explicit instruction might be necessary.

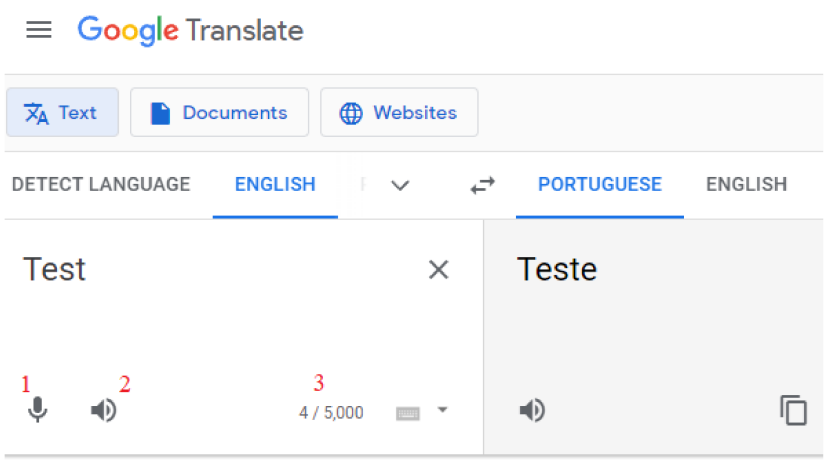
Variation 1: this activity can be used as *homework*. Students can take notes regarding the unintelligible parts to GT and ask for explicit instructions from the teacher during the next class.

Variation 2: a non-synthetic voice can be used instead of GT's TTS. For example, students can select a text passage from https://learningenglish.voanews.com/. This website contains pieces of news for intermediate-level students that have an embedded audio player. Therefore, it can be used as the audio input for the shadow reading activity.

References

CELCE-MURCIA, M.; BRINTON, D. M.; GOODWIN, J. M. Teaching pronunciation hardback with audio CDs (2): A course book and reference guide. 2. ed. Cambridge: Cambridge University Press, 2010.

Appendix A- Google Translate Tools:



1- "translate by voice" button.
2- "listen" button.
3- character counter.

Activity 4 – Paragraph-reading Task

Learning Focus	Perceiving and producing the different -ed pronunciations of regular verbs in the simple past tense.
Level	B1 – B2
Time	40 - 60 minutes
Target Feature	Final clusters: -ed endings
Communicative Framework Stage	Controlled Practice (3) of -ed endings in a paragraph-reading task.
Dictation Passage	When I was 20 years old, I **worked** at a hotel. All I **wanted** was to be the manager. I **helped** many customers, and I **started** offering travel tips. I even **saved** a cat once! Some years ago, I **decided** to change my profession. I **traveled** a lot, and I **watched** many documentaries looking for insights. Finally, I **reached** a final decision: I want to be a tour guide!

Procedure

1. Instruct students to access Google Translate (GT) and show them the dictation passage.
2. Give students explicit instruction on how the pronunciation of regular verbs (-ed ending) in the simple past tense may differ. The dictation passage can be read aloud by the teacher and the different -ed endings can be stressed before the student's production.
3. Then, ask students to voice type the passage, nonstop, until the end of it.
4. Tell students to write down the resulting transcription provided by GT and underline every word and phrase that differs from their original passage paying close attention to the words in bold (-ed endings).
5. Next, encourage students to take notes and try to interpret the reasons why GT mistranscribed the passage. They can work in pairs or groups.
6. Finally, students can voice type the underlined words and phrases again, trying to be understood by GT.

Variation 1: this activity can be used as *homework*. A video lesson can be used as explicit instruction during step 2. Example: https://youtu.be/vv7cBMCBUdk.

Variation 2: this activity can be designed as a guided practice (4). You can ask students to tell a story that happened when they were younger instead of using the dictation passage presented above.

Variation 3: GT's text-to-speech feature can be used as the model for this dictation passage to help students notice the production of the -ed endings before they start voice typing and during the practice of the underlined words.

Activity 5 – Tongue Twisters

Learning Focus	Pronunciation accuracy and fluency
Level	All levels
Time	20 minutes
Target Feature	Oral production of difficult pronunciation patterns that might disrupt communication
Communicative Framework Stage	Controlled Practice (3) with tongue twisters
Dictation Passage	Appendix A

Procedure

1. Explain what a *tongue twister* is and provide an example. Challenge the class to say it, showing that it might be difficult to pronounce all the tricky sounds without practice.
2. Show students a list of tongue twisters (appendix A) and ask them to choose one from this list.
3. Then, instruct students to access Google Translate (GT) and ask them to practice the chosen tongue twister by voice typing it.
4. Tell them that they can speak faster only if GT is transcribing the tongue twister correctly. Explicit instruction on how to produce the difficult pronunciation patterns may be necessary at this point if students have not been intelligible to GT even if they are speaking slowly.
5. Encourage students to increase their speech rate and then challenge their classmates.
6. If time allows, they can choose another tongue twister from the list (or even create their own) and repeat the procedures.

Variation 1: this activity can be used as *homework*. Students can practice the tongue twisters at home and challenge their classmates during the next class.

Variation 2: a different tongue twister can be designed according to the level and the difficulties of the students. In addition, students can search for different tongue twisters online and use their lists to practice and challenge their classmates.

Appendix A- Tongue Twister List:

a) She sells seashells by the seashore.
b) I slit the sheet, the sheet I slit, and on the slitted sheet, I sit.
c) I was thinking of thanking you.
d) Rural rat (say it 3 times).
e) Yes sir, sure sir! (say it 3 times).

Activity 6 – Monitoring Worksheet

Learning Focus	Pronunciation self-assessment and monitoring skills development
Level	All level
Time	30 - 45 minutes
Target Feature	Oral production of word-initial /r/ and /h/ which may hinder intelligibility
Communicative Framework Stage	Controlled Practice (3) with minimal pair practice and a monitoring worksheet
Dictation Passage	Appendix A.

Procedure

1. Instruct students to access Google Translate (GT) and give them the *Monitoring Worksheet* (appendix A).
2. Provide explicit instruction on how the word-initial minimal pair /r/ and /h/ may hinder intelligibility and how this minimal pair differs acoustically and articulatory.

3. Tell students to practice the pronunciation of every word on the worksheet using the voice typing feature of GT. Students can use GT's text-to-speech feature or an online dictionary to listen to the target words before they start producing them.
4. Ask students to write down the resulting transcription after every attempt from the "1st attempt" column to the "5th attempt" column. If external noise or any other technical problems might have affected the attempt, students can disregard it and try to pronounce the word once more.
5. Instruct students to calculate the number of attempts that GT correctly transcribed their speech for each word.
6. Encourage students to add some comments on the last column, specifying some limitations of the program, difficulties that they have faced during the activity, and pronunciation strategies that are beneficial to their development.
7. After finishing the eight target words, tell students to think about one more word starting with /h/ and one starting with /r/ and follow the same procedure.
8. Finally, with the whole worksheet filled in, ask students to calculate the final score by summing up all the totals from the penultimate column and writing it down on the first gap of the last row.
9. Then, they can multiply this result by two, representing the percentage of correctly transcribed words by GT. Students can use this as a self-assessment metric.

Variation 1: this activity can be used as *homework*.

Variation 2: different pronunciation features can be practiced using these procedures. However, it is necessary to provide explicit instruction according to the target feature and select the minimal pairs in advance, considering homophones.

Variation 3: students can create their own monitoring worksheet containing difficult words for them. They can practice the worksheet throughout the semester, working as a pronunciation diary.

Appendix A- Monitoring Worksheet:

Student: _____ **Date:** _____							
Target Words	**1st attempt**	**2nd attempt**	**3rd attempt**	**4th attempt**	**5th attempt**	**Total of correct attempts**	**Comments**
rose							
hose/ hoes							
hug							
rug							
hail							
rail							
rat							
hat							
?							
?							
					Final score:	_____ * 2 = _____ %	

Activity 7 – Role-play Activity

Learning Focus	Speech rehearsal for fluency and accuracy improvement
Level	B2 and above
Time	1 hour – 2 hours
Target Feature	Raising awareness about pronunciation patterns that might disrupt communication in a real-life situation
Communicative Framework Stage	Communicative Practice (5) with a role-play activity

Dictation Passage	Appendix A

Procedure

1. Introduce the topic of the activity by asking students to think about their dream job.
2. Tell them that they will perform a role-play activity in pairs. One student will be the candidate for a job application and the other will be the interviewer.
3. Divide the group into pairs and ask one student of the pair to access Google Translate (GT).
4. Then, give each pair the *Job Interview Questions* (appendix A).
5. The pairs should practice each question at a time. Before the interviewer asks the question, he/she should click on the GT's voice typing feature. The interviewer should not show the screen to the candidate since it may disrupt their attention to the activity.
6. After GT finish transcribing the answer, the interviewer should copy the resulting transcription and paste it to another document or send it to the candidate via an instant messaging program so students can check the transcription when the activity is over.
7. Students should repeat the previous step until all the questions have been answered.
8. Finally, each pair will analyze GT's transcriptions and take notes on the unintelligible part. Encourage the interviewer to give honest feedback to the candidate on whether he/she agrees with GT's mistranscribed passages (according to the candidate's communicative intention) or not.
9. For the unintelligible parts of the speech, the candidate can ask the teacher for explicit instruction on how to overcome the communication breakdown.
10. Finally, the candidate can practice the question that he/she found more difficult to answer.
11. Students can exchange roles and start the activity again.
12. Once students finish rehearsing, they can present the dialog to the class and discuss their answers with their classmates.

Variation 1: this activity can be used as *homework* individually. Students can practice their own answers to the job interview questions and take notes about any difficult words or unintelligible parts of their speech to GT. Then, students should share their notes with the teacher so he/ she can help them overcome possible communication breakdowns.

Variation 2: this activity can be used as *homework* in pairs using a video conference program. Students should take notes and share them with the teacher so he/ she can help them overcome the communication breakdowns that have happened.

Variation 3: this role-play activity can be used for different contexts. The questions provided by the teacher can simulate a conversation between an employee and a customer, or two people asking/giving directions, for example.

Appendix A- Job Interview Questions:

1) Could you tell me a little about yourself?
2) What is your greatest *personal* achievement?
3) What about your *professional* achievement?
4) What qualities make a reliable leader?
5) What are your biggest weaknesses?
6) What are your biggest strengths?
7) What happened last time a customer or co-worker got angry with you?
8) What would be your dream job?
9) What do you know about our industry?
10) Why should we hire you?

APPENDIX F – Teacher's handout

Teacher's Handout

1- Workshop's Program:

Topic	Duration (min)
Introduction	15
Review of the literature	15
Using Google Translate's ASR tool	10
Hands-on: ASR-based pronunciation activities	70
Coffee break	**15**
Brainstorming: ASR for pronunciation teaching	45
Discussion and final remarks	10
Total	**180 (3 hours)**

2- Google Translate's ASR feature

- Tutorial on how to **translate by speech** and **troubleshoot error** messages: https://bit.ly/3fHr4Vd

3- ASR-based Pronunciation Activities

(Same activities presented on Appendix E)

4- Answering the Survey and Further Information

After attending the workshop, you will have access to the survey (see link below) that collects your perceptions of the workshop, the pronunciation activities, and the ASR affordances. Please, read the instructions in each survey section and fill in the form accordingly. Further information will be presented at the end of the survey. You should submit your answers by **November 14th**. If you have questions, comments, or suggestions, you can contact me by email: **william.gottardi@posgrad.ufsc.br**.

Survey's link:

https://forms.gle/D9EBFMNVcPB2XpZA9

5- Further Reading

- **Books:**

CELCE-MURCIA, M.; BRINTON, D. M.; GOODWIN, J. M. **Teaching pronunciation hardback with audio CDs (2): A course book and reference guide**. 2. ed. Cambridge: Cambridge University Press, 2010.

CHAPELLE, C. A. **Computer applications in second language acquisition**. Cambridge University Press, 2001.

PENNINGTON, M. C.; ROGERSON-REVELL, P. English pronunciation teaching and research. **London: Palgrave Macmillan**, v. 10, p. 978–988, 2019.

SICOLA, L.; DARCY, I. Integrating Pronunciation into the Language Classroom. Em: **The handbook of English pronunciation**. Wiley Online Library, 2015. p. 471.

ZIMMER, M.; SILVEIRA, R.; ALVES, U. K. **Pronunciation Instruction for Brazilians: Student's Book**. Cambridge Scholars Publishing, 2009.

- **Articles:**

GOLONKA, E. M. et al. Technologies for foreign language learning: A review of technology types and their effectiveness. **Computer Assisted Language Learning**, v. 27, n. 1, p. 70–105, 2014.

GOMES JUNIOR, R. C.; SILVA, L. DE O.; PAIVA, V. L. M. DE O. E. Tecnologias digitais para aprender e ensinar inglês no Brasil. **Texto Livre**, v. 15, p. e38008, 2022.

GONÇALVES, A. R.; SILVEIRA, R. Intelligibility research in Brazil: empirical findings and methodological issues. **Revista Horizontes de Linguistica Aplicada**, v. 14, n. 1, p. 51–81, 2015.

GOTTARDI, W.; ALMEIDA, J. F. DE; TUMOLO, C. H. S. Automatic speech recognition and text-to-speech technologies for L2 pronunciation improvement: reflections on their affordances. **Texto livre**, v. 15, 2022.

LIAKIN, D.; CARDOSO, W.; LIAKINA, N. Mobilizing Instruction in a Second-Language Context: Learners' Perceptions of Two Speech Technologies. **Languages**, v. 2, n. 3, p. 11, 2017.

MCCROCKLIN, S.; EDALATISHAMS, I. Revisiting Popular Speech Recognition Software for ESL Speech. **TESOL Quarterly**, v. 54, n. 4, p. 1086–1097, 2020.

MUNRO, M. J.; DERWING, T. M. Intelligibility in research and practice: Teaching priorities. Em: REED, M.; LEVIS, J. M. (Eds.). **The Handbook of English Pronunciation**. Wiley Online Library, 2015. p. 375–396.

SILVEIRA, R. et al. Percepção, produção e inteligibilidade do inglês falado por usuários brasileiros. **Perspectivas atuais de aprendizagem e ensino de línguas. Florianópolis: LLE/CCE/UFSC**, p. 237–283, 2017.

SOUZA, H. K., & GOTTARDI, W. (2022). HOW WELL CAN ASR TECHNOLOGY UNDERSTAND FOREIGN-ACCENTED SPEECH?. In SciELO Preprints. https://doi.org/10.1590/010318138668782v61n32022

YOSHIDA, M. T. Choosing technology tools to meet pronunciation teaching and learning goals. **The CATESOL Journal**, v. 30, n. 1, p. 195–212, 2018.

APPENDIX G – Workshop slides

1

2

01. Introduction and Research Overview

3

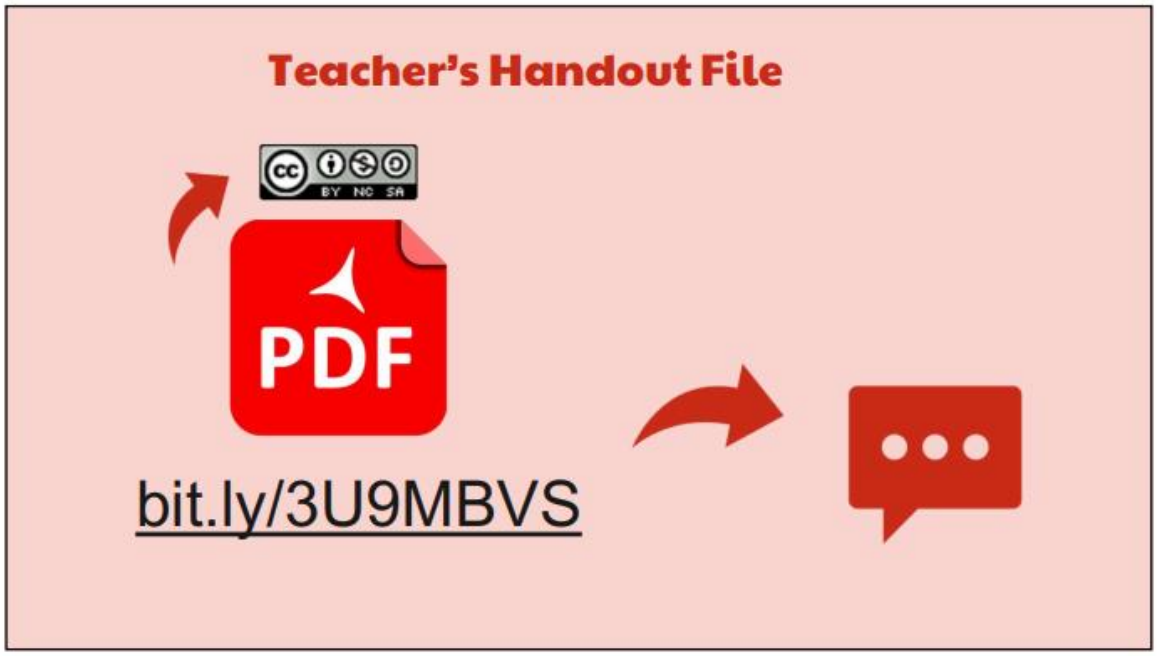
What is this research about?

Automatic Speech Recognition as a **Pronunciation Teaching** Resource: What are the in-service Teachers' Perceptions of this Speech Technology?

4


About this Workshop

5


Teacher's Handout File
BY NC SA
PDF
bit.ly/3U9MBVS

6

Workshop' Program

p. 1

Topic	Duration (min)
Introduction	15
Review of the literature	15
Using Google Translate's ASR tool	10
Hands-on: ASR-based pronunciation activities	70
Coffee break	**15**
Brainstorming: ASR for pronunciation teaching	45
Discussion and Final remarks	10
Total	**180 (3 hours)**

7

02. Review of the Literature

8

Pronunciation Teaching

- "Pronunciation is a term used to capture all aspects of how we employ speech sounds for **communicating**" (BURNS; SEIDLHOFER, 2020, p. 247).
- Pronunciation is a key element aiming at **successful spoken communication** (PENNINGTON; ROGERSON-REVELL, 2019).
- Pronunciation is a commonly **neglected** component in many language programs (DERWING, 2010).
- **Pronunciation teaching** is important to help raise students' awareness regarding **different sounds and sound features** and to improve their overall **speaking skills** (HARMER, 2015).

9

Pronunciation Teaching

Focus on **intelligibility**

"the extent to which a speaker's message is **actually understood** by a listener".

(MUNRO; DERWING, 1995, p. 289)

- **Intelligibility** issues presented on L2 Brazilian learners' speech (GONÇALVES; SILVEIRA, 2015; SILVEIRA et al., 2017).
- Different L2 Brazilian learners' difficulties (ZIMMER et al., 2009).

10

Communicative Framework for Teaching Pronunciation

(CELCE-MURCIA et al., 2010, p. 45)

1. Description and analysis
2. Listening discrimination
3. Controlled practice
4. Guided practice
5. Communicative practice

11

Automatic Speech Recognition (ASR)

Transcribes speech
based on oral input, also known as STT
(Speech-to-Text)

(MOUSSALLI; CARDOSO, 2020)

12

How can I use ASR?

web browsers

voice typing

smart assistants

mobile apps

13

Research on ASR-supported pronunciation teaching

- Increased learner autonomy in & out of the classroom.
  (LEVIS; SUVOROV, 2013; LIAKIN et al., 2017; MCCROCKLIN, 2014, 2016; BASHORI et al., 2020, 2021; KIVISTÖ-DE SOUZA; GOTTARDI, in press)
- More willingness to communicate & reduce foreign language anxiety.
  (INCEOGLU et al., 2020; KIM, 2006)
- Oral production improvement.
  (NERI et al., 2006, 2008)
- Provide relevant **feedback** during speaking activities.
  (LEVIS; SUVOROV, 2013)

14

Research on ASR-supported pronunciation teaching

- Limitless opportunities to practice speech (more output). (DIZON; TANG, 2020; WALESIAK, 2021)
- Facilitate the evaluation of pronunciation. (FOUZ-GONZÁLEZ, 2015)
- Positive influence on students' learning process. (MROZ, 2018; WALESIAK, 2021)

15

ASR Dictation Accuracy Rate

- Google's accuracy for non-native speech were **above 88%** (MCCROCKLIN et al., 2019).
- Google's accuracy in a sentence-reading task was **over 90%** (MCCROCKLIN; EDALATISHAMS, 2020).
- **Brazilian Portuguese** (80%) and Spanish speakers (89%) were highly intelligible to VoiceNotebook (Google's technology) and Microsoft Word Dictation (KIVISTÖ-DE SOUZA; GOTTARDI, in press).

16

ASR Dictation: Limitations

- Use predictable vocabulary: avoid proper nouns, some combinations of words (ASHWELL; ELAM, 2017; KIVISTÖ-DE SOUZA; GOTTARDI, in press).
- Avoid side talks or multitalker speech (YU; DENG, 2015).
- Some speaker characteristics may affect ASR accuracy (e.g., child speech) (JURAFSKY; MARTIN, 2021).
- **Fluent** and **proficient speech** may represent an extra barrier for the ASR programs due to the speech rate (KIVISTÖ-DE SOUZA; GOTTARDI, in press).

17

ASR Dictation - Pedagogical Implications

To achieve better results, ensure access to:

- a good-quality microphone
- stable internet access
- quiet environment
- an appropriate text passage
- teacher's guidance and extra feedback

(GOTTARDI et al., 2022; KIVISTÖ-DE SOUZA; GOTTARDI, in press)

18


03.
Google Translate's ASR tool

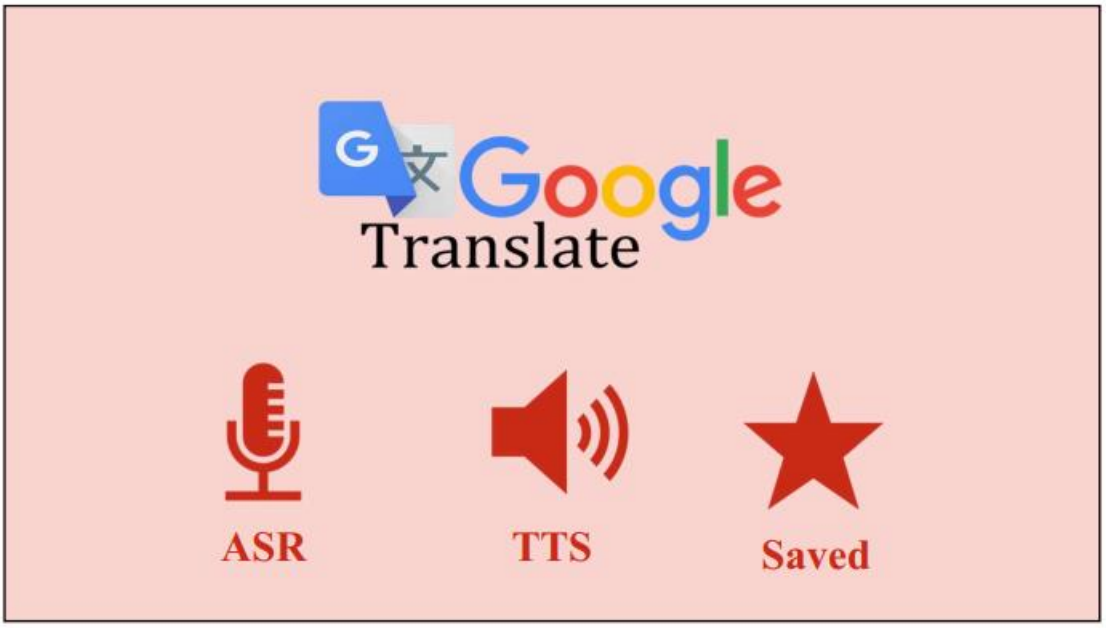
19


Google Translate
PDF
p. 1
Need permission to use microphone
Voice input isn't supported on this browser

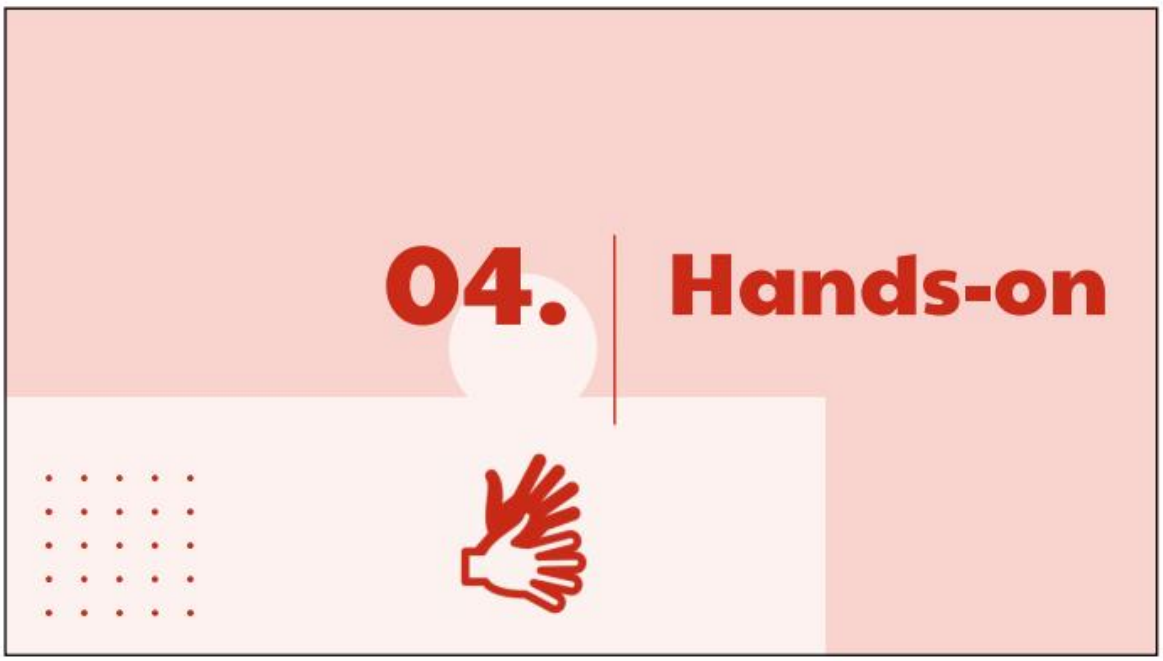
20


Google Translate
We're having trouble hearing you
Configurações...
Sistema...
Som...
Entrada...

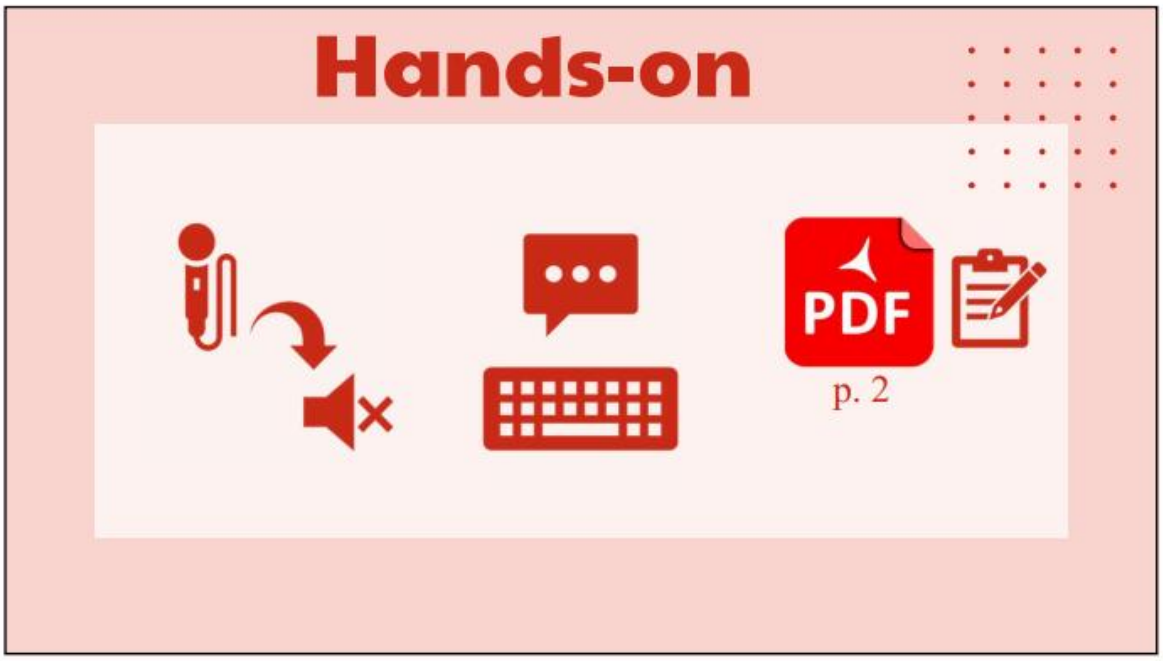
21


Google Translate
ASR
TTS
Saved

22



23



24

Hands-on

Activities:
1 - 2 - 4 - 5.

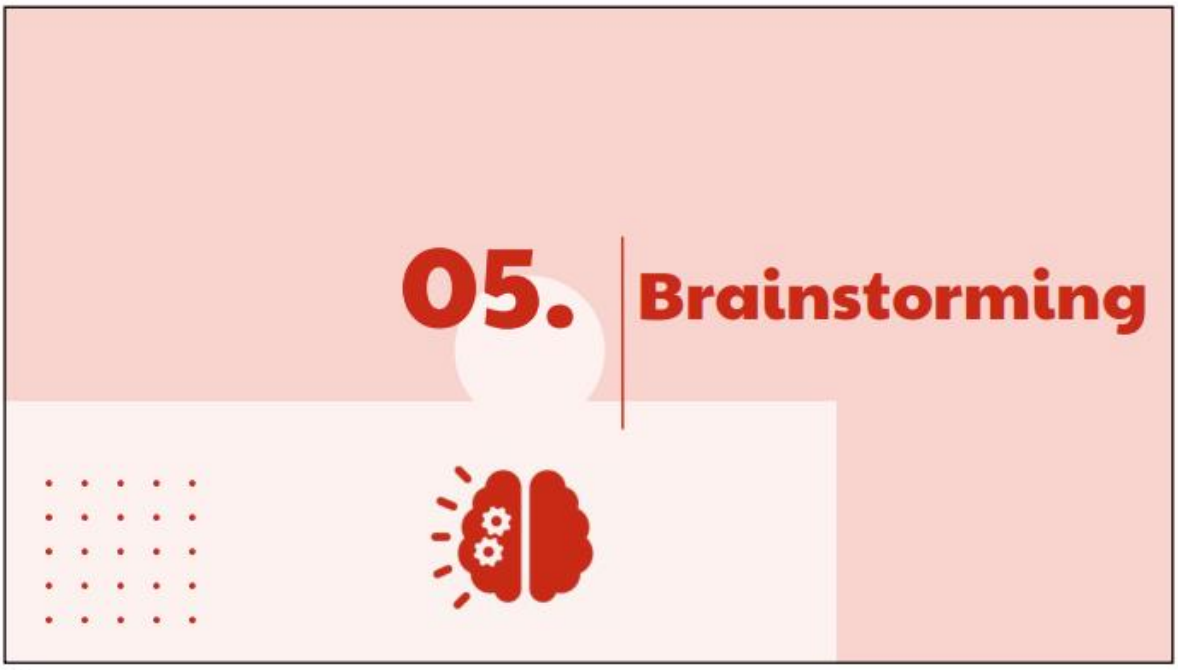
Summary

Activities:
3 - 6 - 7.

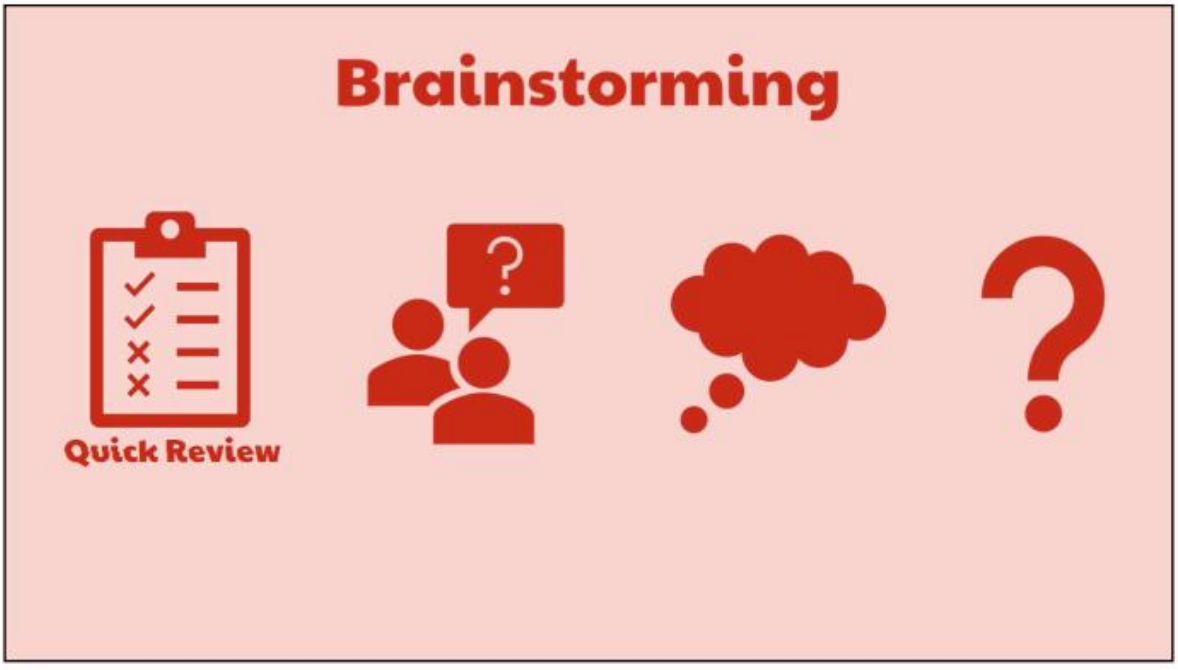
25

Coffee Break!

26



27



28

06. Final Remarks

29

Discussion: ASR for pronunciation teaching can…

- Promote learner **autonomy**.
- Be used in various **learning contexts**.
- Offer **individualized instruction** based on learners' specific needs.
- Facilitate integrating the **pronunciation** component in language courses.
- Serve as a **complement** to classroom instruction and suitable for **extra-class activities**.

30

Discussion: ASR can offer to the **learner** more…

- Oral **production** opportunities.
- Orthographic **feedback**.
- Contact with the target language.
- **Motivation** and **interest** to study the target language.

31

Any questions or comments?

32

Further Information

PDF
p. 14

https://forms.gle/D9EBFMNVcPB2XpZA9

Survey's Deadline:
November 14th

33

Further Reading

PDF
p. 14

34

Final Remarks

'Educational technology is only as good as the **humans** behind it.'

(REVELL-ROGERSON, 2021)

35

References

ASHWELL, T.; ELAM, J. R. How accurately can the google web speech API recognize and transcribe Japanese L2 english learners' oral production? **JALT CALL Journal**, v. 13, n. 1, p. 59–76, 2017.

BASHORI, M. et al. Web-based language learning and speaking anxiety. **Computer Assisted Language Learning**, v. 0, n. 0, p. 1–32, 2020.

BASHORI, M. et al. Effects of ASR-based websites on EFL learners' vocabulary, speaking anxiety, and language enjoyment. **System**, v. 99, n. April, p. 102496, 2021.

BURNS, A.; SEIDLHOFER, B. Speaking and pronunciation. Em: **An introduction to applied linguistics**. [s.l.] Routledge, 2020. v. 3rd Editionp. 240–258.

CELCE-MURCIA, M.; BRINTON, D. M.; GOODWIN, J. M. **Teaching pronunciation hardback with audio CDs (2): A course book and reference guide**. 2. ed. Cambridge: Cambridge University Press, 2010.

DERWING, T. M. Utopian Goals for Pronunciation Teaching. **Proceedings from the 1st Conference of Pronunciation in Second Language Learning and Teaching**, n. October 2010, p. 24–37, 2010.

DIZON, G.; TANG, D. Intelligent personal assistants for autonomous second language learning: An investigation of Alexa. **JALT CALL Journal**, v. 16, n. 2, p. 107–120, 2020.

GONÇALVES, A. R.; SILVEIRA, R. Intelligibility research in Brazil: empirical findings and methodological issues. **Revista Horizontes de Linguistica Aplicada**, v. 14, n. 1, p. 51–81, 2015.

GOTTARDI, W.; ALMEIDA, J. F. DE; TUMOLO, C. H. S. Automatic speech recognition and text-to-speech technologies for L2 pronunciation improvement: reflections on their affordances. **Texto livre**, v. 15, 2022.

HARMER, J. **The practice of English language teaching**. [s.l.] Pearson longman, 2015. v. 5th edition

INCEOGLU, S.; LIM, H.; CHEN, W. H. Asr for EFL pronunciation practice: Segmental development and learners' beliefs. **Journal of Asia TEFL**, v. 17, n. 3, p. 824–840, 2020.

JURAFSKY, D.; MARTIN, J. H. **Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition**: Unpublished manuscript. USA: [s.n.].

36

References

KIM, I. S. Automatic speech recognition: Reliability and pedagogical implications for teaching pronunciation. **Educational Technology and Society**, v. 9, n. 1, p. 322–334, 2006.

LEVIS, J.; SUVOROV, R. Automatic Speech Recognition. Em: CHAPELLE, C. A. (Ed.). **The encyclopedia of applied linguistics**. New York: Wiley-Blackwell, 2013. p. 316–323.

LIAKIN, D.; CARDOSO, W.; LIAKINA, N. Mobilizing Instruction in a Second-Language Context: Learners' Perceptions of Two Speech Technologies. **Languages**, v. 2, n. 3, p. 11, 2017.

MCCROCKLIN, S. Automatic Speech Recognition : Making It Work for Your. n. January 2015, 2018.

MCCROCKLIN, S.; EDALATISHAMS, I. Revisiting Popular Speech Recognition Software for ESL Speech. **TESOL Quarterly**, v. 54, n. 4, p. 1086–1097, 2020.

MCCROCKLIN, S.; HUMAIDAN, A.; EDALATISHAMS, I. ASR dictation program accuracy: Have current programs improved? **Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference**, n. June, p. 191–200, 2019.

MCCROCKLIN, S. M. Dictation programs for pronunciation learner empowerment. **Proceedings of the 5th pronunciation in second language learning and teaching conference.**, n. September, p. 30–39, 2014.

MOUSSALLI, S.; CARDOSO, W. Intelligent personal assistants: can they understand and be understood by accented L2 learners? **Computer Assisted Language Learning**, v. 33, n. 8, p. 865–890, 2020.

MROZ, A. Seeing how people hear you: French learners experiencing intelligibility through automatic speech recognition. **Foreign Language Annals**, v. 51, n. 3, p. 617–637, 2018.

MUNRO, M. J.; DERWING, T. M. Intelligibility in research and practice: Teaching priorities. Em: REED, M.; LEVIS, J. M. (Eds.). **The Handbook of English Pronunciation**. [s.l.] Wiley Online Library, 2015. p. 375–396.

NERI, A.; CUCCHIARINI, C.; STRIK, H. ASR-based corrective feedback on pronunciation: Does it really work? **INTERSPEECH 2006 and 9th International Conference on Spoken Language Processing, INTERSPEECH 2006 - ICSLP**, v. 4, n. May 2014, p. 1982–1985, 2006.

37

References

NERI, A.; CUCCHIARINI, C.; STRIK, H. The effectiveness of computer-based speech corrective feedback for improving segmental quality in L2 Dutch. **ReCALL**, v. 20, n. 2, p. 225–243, 2008.

PENNINGTON, M. C.; ROGERSON-REVELL, P. English pronunciation teaching and research. **Londres: Palgrave Macmillan**, v. 10, p. 978–988, 2019.

REVELL-ROGERSON, P. M. Computer-Assisted Pronunciation Training (CAPT): Current Issues and Future Directions. **RELC Journal**, v. 52, n. 1, p. 189–205, 2021.

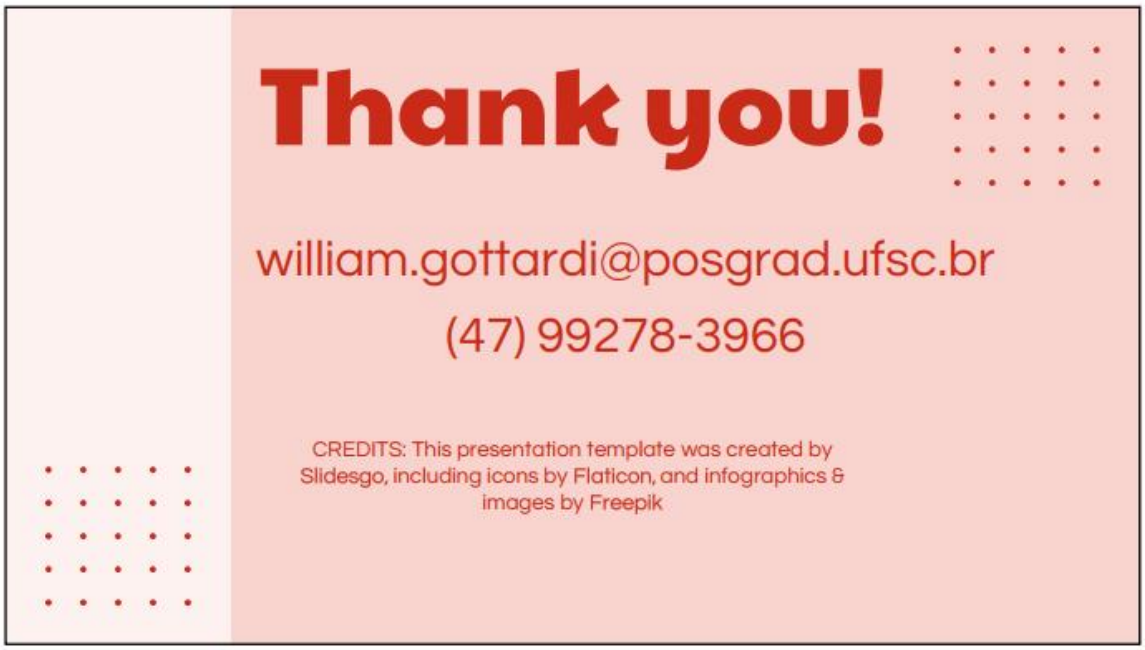
Souza, H. K., & Gottardi, W. (2022). HOW WELL CAN ASR TECHNOLOGY UNDERSTAND FOREIGN-ACCENTED SPEECH?. In SciELO Preprints. https://doi.org/10.1590/010318138668782v61n32022

SILVEIRA, R. et al. Percepção, produção e inteligibilidade do inglês falado por usuários brasileiros. **Perspectivas atuais de aprendizagem e ensino de línguas. Florianópolis: LLE/CCE/UFSC**, p. 237–283, 2017.

YU, D.; DENG, L. **Automatic Speech Recognition A Deep Learning Approach**. London: Springer, 2015.

ZIMMER, M.; SILVEIRA, R.; ALVES, U. K. **Pronunciation Instruction for Brazilians: Student's Book**. [s.l.] Cambridge Scholars Publishing, 2009.

38


Thank you!
william.gottardi@posgrad.ufsc.br
(47) 99278-3966
CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik
39

APPENDIX H - Participants' answers to the background questionnaire

Section I - Personal Information (A) - Demographic Information[11]

P*	Name	Age	Gender	City where you live	State where you live	Country where you live
P1	Bruno de Azevedo	33	Male	Pato Branco	Paraná	Brazil
P2	Micaella de Lima	25	Female	Blumenau	Santa Catarina	Brazil
P3	Debora Cristofolini	34	Female	Timbó	Santa Catarina	Brazil
P4	Janaina Fernanda de Almeida	24	Female	Brusque	Santa Catarina	Brazil
P5	Mauricio de Bortolli Lattmann	31	Male	Camboriú	Santa Catarina	Brazil
P6	Luana Tiburi Dani Gauer	36	Female	Caxias do Sul	Rio Grande do Sul	Brazil
P7	Luana Garbin Baldissera	32	Female	Varese	Varese	Italy
P8	Andreik Edvan Rocha	26	Male	Curitiba	Paraná	Brazil
P9	Jason Memmott	42	Male	Florianópolis	Santa Catarina	Brazil
P10	Antônio Carlos Neto	29	Male	Caxias	Maranhão	Brazil
P11	Beata Walesiak	40	Female	Warsaw	Warsaw	Poland
P12	Ángeles Roxana Cañete	34	Female	Clorinda	Formosa	Argentina

Source: author
*P = Participant code.

Section I - Personal Information (B) - Work and Study Background Information)

P*	Q1*	Q2*	Q3*	Q4*	Q5*
P1	Letras - UNIARP - 2013	No	11	Public	Technical/ Vocational
P2	Letras/ Furb (2018) e Pedagogia/ Uniasselvi (2022)	Metodologias de ensino para educação Bilingue/ UniOpet (2020) e Coordenação pedagógica/ Unina (2022)	8	Private	Bilingual Education, Elementary School I

[11] The e-mail address was omitted to protect Participants' personal data.

P3	Pedagogia 2009 Uniasselvi / Letras - Português/Inglês 2014 FURB	Mestrado em Educação FURB 2023	15	Public	Elementary School I, Elementary School II
P4	Letras - FURB - 2019	Mestrado - Estudos linguísticos e literários de língua inglesa - UFSC - 2022	7	Private	Young Learners, Bilingual Education, Elementary School II
P5	Business Administration - UNIVALI - 2015	Linguistics in English - UFSC - March 2023	12	Private	Private teacher
P6	Licenciatura em Letras (Português-Inglês) - Universidade de Caxias do Sul/RS - 2009	Mestrado em Letras, Cultura e Regionalidade - Universidade de Caxias do Sul/RS - 2013 Doutorado em Letras - UFRGS - 2023	15	Public	High School, Technical/ Vocational, Higher Education
P7	Letras Port Ingles / Universidade Estadual do Oeste do Paraná	MA English - Linguistics and Literature (UFSC 2020) and Phd in English - Linguistics and Literature (UFSC 2020)	14	Both	Elementary School I, High School, Higher Education, Language School, Private teacher
P8	Languages Portuguese and English and its respective literature - FURB, 2019	English Teaching Methodology - UNINTER, 2020	6	Private	Language School, Private teacher
P9	Bachelor of Arts (Japanese) University of the Sunshine Coast (AUS)	No	4	Private	Private teacher
P10	Letras Português/Inglês e Respectivas Literaturas - Universidade Estadual do Maranhão - UEMA	Mestrado em Literatura - Universidade Federal do Piauí - UFPI	10	Private	Elementary School I, Elementary School II, High School, Higher Education, Language School, Private teacher

P11	English Studies, Institute of English Studies, University of Warsaw	PhD program, English Studies, Institute of English Studies, University of Warsaw	16	Both	ELSA
P12	Teacher of English	Diploma in Phonetics and Phonology	10	Public	High School, Higher Education

Source: author

*P = Participant code. *Q1 = Which undergraduate course(s) have you taken/are you taking (write the name of the course(s), the institution(s), and the (expected) year of completion)?. *Q2 = Have you taken/are you taking any postgraduate course (write the name of the course the institution and the (expected) year of completion)?. *Q3 = Years of experience teaching English. *Q4 = In which sector do you work?. *Q5 = Mark the context(s) of education you are currently working in: (you can check more than one answer)

Section II - Information about participants' teaching practices

P*	**Q1***	**Q2***	**Q3***	**Q4***	**Q5***	**Q6***
P1	7	5	8	8	0	0
P2	5	10	4	10	0	0
P3	8	9	9	9	7	8
P4	8	6	6	9	2	3
P5	8	6	8	5	3	0
P6	7	7	9	9	2	0
P7	9	7	10	10	0	0
P8	9	10	10	10	1	1
P9	7	6	5	8	0	0
P10	9	10	10	10	10	0
P11	9	10	10	9	3	7
P12	9	10	10	10	5	5

Source: author

*P = Participant code. *Q1 = How confident are you in teaching pronunciation?. *Q2 = How often do you teach pronunciation in your classes?. *Q3 = How confident are you in using digital resources (computer, cell phone, projector, websites, apps) in your pedagogical practices?. *Q4 = How often do you use digital resources in your classes?. *Q5 = How often do you use Google Translate in your classes?. *Q6 = How often do you use any Automatic Speech Recognition (ASR) tool as a resource for teaching pronunciation?

APPENDIX I – Participants' answers to the survey

Section I - Teacher Development Needs

P*	Q1*	Q2*	Q3*
P1	Teaching techniques, Technological resources	7	8
P2	Theory, Teaching techniques, Technological resources	9	9
P3	Teaching techniques, Technological resources	9	10
P4	Teaching techniques, Practical ideas on how to use the resources	7	9
P5	Teaching techniques, Technological resources	7	8
P6	Theory, Teaching techniques, Digital resources	7	8
P7	Theory	10	10
P8	Theory, Digital resources	10	10
P9	Teaching techniques	7	5
P10	Theory, Teaching techniques, Digital resources	10	10
P11	Teaching techniques	10	10
P12	Teaching techniques, Digital resources	10	7

Source: author

*P = Participant code. *Q1 = In which area do you need more courses/workshops to make you feel more comfortable teaching pronunciation?. *Q2 = I feel comfortable to use ASR to teach pronunciation after attending the workshop. *Q3 = I would use ASR to improve my own pronunciation.

Section II - ASR Accessibility

P*	Q1*	Q2*	Q3*	Q4*
P1	9	10	online, hybrid (online and face-to-face)	External noise, Helping students interpret the orthographic feedback provided by the ASR
P2	9	10	online, face-to-face, hybrid (online and face-to-face)	Access to microphones/ headphones with microphone, External noise, Helping students interpret the orthographic feedback provided by the ASR
P3	10	10	hybrid (online and face-to-face)	Access to microphones/ headphones with microphone, External noise

P4	10	10	online, hybrid (online and face-to-face)	Access to a personal computer (PC) or mobile devices (cell phones, tablets, and laptops), External noise, Assist a big group of students (around 35) at the same time.
P5	10	10	hybrid (online and face-to-face)	Helping students interpret the orthographic feedback provided by the ASR
P6	8	10	online, hybrid (online and face-to-face)	Access to microphones/ headphones with microphone, External noise, Helping students interpret the orthographic feedback provided by the ASR
P7	10	10	online, face-to-face, hybrid (online and face-to-face)	Access to microphones/ headphones with microphone, External noise, Helping students interpret the orthographic feedback provided by the ASR
P8	10	8	hybrid (online and face-to-face)	Access to stable internet connection, Access to a personal computer (PC) or mobile devices (cell phones, tablets, and laptops), Access to microphones/ headphones with microphone
P9	8	10	hybrid (online and face-to-face)	External noise, Helping students interpret the orthographic feedback provided by the ASR
P10	10	10	online, face-to-face, hybrid (online and face-to-face)	Access to stable internet connection, Access to a personal computer (PC) or mobile devices (cell phones, tablets, and laptops), External noise, Helping students interpret the orthographic feedback provided by the ASR
P11	10	9	online, face-to-face, hybrid (online and face-to-face)	Access to stable internet connection, External noise
P12	7	5	hybrid (online and face-to-face)	Access to stable internet connection

Source: author

*P = Participant code. *Q1 = It is easy to use Google Translate's voice recognition feature. *Q2 = My students have access to Google Translate to practice pronunciation. *Q3 = In which context(s) would you use ASR to teach pronunciation?. *Q4 = Which of the ASR possible limitations is/are difficult to overcome considering your teaching context?.

Section III - ASR Affordances

P*	Q1*	Q2*	Q3*	Q4*	Q5*	Q6*
P1	8	9	10	10	8	an auxiliary resource for the regular classes, an out of class supplement/ homework, a learning tool suggestion for self-studying, a pronunciation self-assessment tool suggestion
P2	8	5	10	5	8	an auxiliary resource for the regular classes, an out of class supplement/ homework, a learning tool suggestion for self-studying
P3	10	10	10	10	10	an auxiliary resource for the regular classes, a learning tool suggestion for self-studying, a pronunciation self-assessment tool suggestion
P4	10	5	10	8	8	an auxiliary resource for the regular classes, an out of class supplement/ homework, a learning tool suggestion for self-studying, a pronunciation self-assessment tool suggestion
P5	8	7	8	9	9	an auxiliary resource for the regular classes, an out of class supplement/ homework, a learning tool suggestion for self-studying, a pronunciation self-assessment tool suggestion
P6	9	8	8	8	9	an out of class supplement/ homework, a learning tool suggestion for self-studying, a pronunciation self-assessment tool suggestion
P7	8	8	10	10	10	an auxiliary resource for the regular classes, an out of class supplement/ homework, a learning tool suggestion for self-studying
P8	10	9	9	10	9	an auxiliary resource for the regular classes, an out of class supplement/ homework, a learning tool suggestion for self-studying
P9	8	5	7	8	9	an out of class supplement/ homework, a learning tool suggestion for self-studying, a pronunciation self-assessment tool suggestion
P10	10	10	10	10	10	an auxiliary resource for the regular classes, an out of class supplement/ homework, a learning tool suggestion for self-studying, a pronunciation self-assessment tool suggestion
P11	10	9	10	9	9	an out of class supplement/ homework, a learning tool suggestion for self-studying, a pronunciation self-assessment tool suggestion
P12	9	7	7	7	6	an auxiliary resource for the regular classes, a learning tool suggestion for self-studying

Source: author

*P = Participant code. *Q1 = ASR facilitates the teaching of pronunciation. *Q2 = My students would be interested in using ASR to improve their own pronunciation. *Q3 = I would use ASR as a complementary tool for teaching pronunciation in my classes. *Q4 = ASR can encourage/motivate learners to produce more output outside the classroom. *Q5 = ASR transcription (orthographic feedback) can be beneficial to the development of learner's pronunciation. *Q6 = I would use ASR for pronunciation teaching as:.

Section IV - Activities Appraisal

Questions 1 to 13:

P*	**Q1***	**Q2***	**Q3***	**Q4***	**Q5***	**Q6***	**Q7***	**Q8***	**Q9***	**Q10***	**Q11***	**Q12***	**Q13***
P1	10	10	9	10	5	3	3	10	8	6	6	10	10
P2	10	10	10	10	10	10	9	8	7	5	7	8	8
P3	10	10	10	10	10	10	10	10	10	10	9	10	9
P4	10	10	9	7	7	8	5	10	8	10	6	10	8
P5	9	7	9	8	8	7	6	8	9	9	8	9	9
P6	8	9	9	9	9	9	9	9	9	9	10	9	9
P7	10	10	10	10	10	8	7	10	10	10	10	10	10
P8	10	10	10	10	10	10	9	9	10	10	10	9	8
P9	10	10	10	10	10	7	8	10	10	9	8	10	10
P10	10	10	5	10	10	10	8	5	10	10	10	10	10
P11	10	9	7	10	7	3	6	5	10	10	8	9	9
P12	8	9	8	9	8	8	9	8	8	8	8	9	6

Source: author

*P = Participant code. *Q1 = The activities have a clear and reasonable learning focus. *Q2 = The activities have addressed linguistic features related to pronunciation (ex. minimal pairs, vowel quality, segments, prosody, accuracy, fluency). *Q3 = The activities have an appropriate degree of difficulty for the learners considering the target level indicated by the researcher during the workshop. *Q4 = The activities can be adapted according to learners' individual characteristics (e.g., age, computer experience, learning style). *Q5 = The activities allow learners to use the language for interpreting and constructing meaning. *Q6 = The activities direct learner's attention primarily toward the meaning of the language. *Q7 = The activities show a strong correspondence between the tasks and what learners would expect to see outside the classroom. *Q8 = The activities address real pronunciation difficulties faced by Brazilian learners of English. *Q9 = The activities can help learners develop autonomous learning strategies. *Q10 = The activities can offer a positive teaching/learning experience with the digital technology used (Google Translate). *Q11 = The activities can be implemented considering the particular constraints of my class and language program. *Q12 = The activities are adaptable to my teaching context in relation to the equipment used (e.g., microphone, earphone, computer, laptop, cellphone). *Q13 = The activities are adaptable to my teaching context in relation to the infrastructure required (e.g., internet, wi-fi, computer lab).

Questions 14 and 15:

P*	Q14*	Q15*
P1	Activity 1 – Pronunciation Self-assessment, Activity 2 – Vowel Contrast, Activity 4 – Paragraph-reading Task	Activity 7 – Role-play Activity
P2	Activity 1 – Pronunciation Self-assessment, Activity 2 – Vowel Contrast, Activity 5 – Tongue Twisters	Activity 1 – Pronunciation Self-assessment
P3	Activity 1 – Pronunciation Self-assessment, Activity 2 – Vowel Contrast	Activity 1 – Pronunciation Self-assessment
P4	Activity 1 – Pronunciation Self-assessment, Activity 2 – Vowel Contrast, Activity 4 – Paragraph-reading Task, Activity 5 – Tongue Twisters, Activity 7 – Role-play Activity	Activity 5 – Tongue Twisters
P5	Activity 1 – Pronunciation Self-assessment, Activity 2 – Vowel Contrast, Activity 5 – Tongue Twisters	Activity 5 – Tongue Twisters
P6	Activity 1 – Pronunciation Self-assessment, Activity 2 – Vowel Contrast, Activity 5 – Tongue Twisters, Activity 7 – Role-play Activity	Activity 2 – Vowel Contrast
P7	Activity 1 – Pronunciation Self-assessment, Activity 2 – Vowel Contrast, Activity 6 – Monitoring Worksheet, Activity 7 – Role-play Activity	Activity 1 – Pronunciation Self-assessment
P8	Activity 1 – Pronunciation Self-assessment, Activity 2 – Vowel Contrast, Activity 4 – Paragraph-reading Task	Activity 2 – Vowel Contrast
P9	Activity 1 – Pronunciation Self-assessment, Activity 2 – Vowel Contrast, Activity 7 – Role-play Activity	Activity 1 – Pronunciation Self-assessment
P10	Activity 1 – Pronunciation Self-assessment, Activity 2 – Vowel Contrast, Activity 3 – Shadow Reading, Activity 4 – Paragraph-reading Task, Activity 5 – Tongue Twisters, Activity 6 – Monitoring Worksheet, Activity 7 – Role-play Activity	Activity 5 – Tongue Twisters
P11	Activity 1 – Pronunciation Self-assessment, Activity 2 – Vowel Contrast, Activity 3 – Shadow Reading, Activity 4 – Paragraph-reading Task, Activity 5 – Tongue Twisters	Activity 4 – Paragraph-reading Task

P12	Activity 2 – Vowel Contrast, Activity 5 – Tongue Twisters, Activity 7 – Role-play Activity	Activity 7 – Role-play Activity

Source: author

*P = Participant code. *Q1 = Which activity(ies) presented during the workshop would you use with your students?. *Q2 = Among all the proposed activities, which one was your favorite?.

APPENDIX J – Qualitative data grouped and coded

I Category - Teacher Development Needs

P*	Comment
P1	I think the hands-on part is the richer part of your workshop {orally}
P2	I loved your workshop. I think that you could add 1h to have more time to discuss the topics and provide feedback to people that are participating. {survey}
P3	-
P4	-
P5	I really enjoyed the workshop, it was clear, with a nice flow and an interesting topic as well. I just wish we were given more time to express ourselves. {survey} [Final Remarks] The review of the literature was well summarized, was quite short, was, uh, great. [...] I just think we should have more time mainly in the beginning … for us to have some time to talk about our experience with Google Translate. {orally}
P6	-
P7	Even though a 3h online workshop may seem a long time initially, the time really went by fast, and I don't see how it could have been shorter! The workshop was always interactive, with moments of explanation and hands on.... all the activities proposed were well explained and I believe this made all the difference in showing how we can take advantage of the tool. {survey}
P8	I really enjoyed the workshop. {survey}
P9	It [the workshop] was very clearly and logically presented. {survey}
P10	[Final Remarks] Uh, first of all, I'd like to thank you because it was a good opportunity to know how we can use this technology… like in a webinar. {orally}
P11	A very useful workshop and a great handout - seriously! {survey}
P12	The workshop was amazing and highly useful. {survey}

Source: author
Annotated text is presented between square brackets.
The source of speech is presented between curly brackets (chat, orally, or survey).
*Participant code.

II Category - ASR Accessibility

P*	Comment
P1	-
P2	-
P3	-
P4	-
P5	-
P6	-

P7	[Answer to the question "Do you think your students would enjoy the activities?] Sorry, I said maybe because I don't know. It's so personal. Maybe I need to learn where my students take the class, because if there is a noise, maybe you know because when they say something on the microphone, I listen to the TV or other people speaking, so maybe this could be one reason for them not to, and so that's why I don't say yes, definitely. I need to do a research first. [...] Yeah, because maybe they're taking the class in the library too, and that they can't speak, you know, so maybe could be frustrating. Depending on these issues. {orally}
P8	So basically, what we were discussing [regarding the discussion on the breakout room] like my cons. The problem, you know, about this ASR, uhm, is accessibility, you know, 'cause I remember working with these types of things, especially I try to use, uhm, my phone, because… like… there were some lessons, almost no one had phones, you know, but they had but with no Internet, so I had to share my Internet with them. So, in public schools, I think this is the issue, but as a teacher we can try to find ways to work and bring different things like this. Once you showed us so. Yep, just really nice this workshop. {orally}
P9	[Regarding the discussion on the breakout room] OK, yeah, obviously different group of students [this participant teaches private students only] so the accessibility is not usually an issue, it's just time because everyone is so busy doing, they say or whatever else it is. But yeah, but I think definitely could be useful for the demographic (population) I teach usually. {orally}
P10	[Final Remarks] When I was talking in the in the other room [breakout room discussion] I was always saying that there are a lot of different realities, especially in Brazil according to the region. For example, I'm from Maranhão, so I'm a teacher here in Maranhão and it's really difficult sometimes to teach pronunciation for the students, especially when you have different realities. We have the public school realities. Uh, the English courses reality - I teach in English courses too, the university, and it's really difficult according to this reality, especially when these students they don't have like this background from the language. Sometimes they are really raw (not experienced) in the classrooms like starting, uh, to study English for their first time, and for example, it's something that it demands time. It demands to do things in advance. We need to set an environment for them in order to make them understand the concept and understand the tool and how they can use it, but it's really, really interesting. {orally}
P11	[Final Remarks] I just wanted to say that you know, for my own research also, I can share that even if we are planning to use technology in the classroom, we need to take into consideration the fact that we need to devote some time and it's not going to be 5 minutes that maybe half an hour or even an hour or two. It was to kind of create the habit of using some technology in the classroom or just demonstrate to our students how to access this technology. Ensure that everyone has access because I've learned that out of 20 people, there's usually two or three who will have problems with downloading a certain app or with access to the Internet, and you know that they won't be able to use a particular technology and you will have to stay after class depending on how you teach the course and help them out. And so, it's really good to have like a pre started session which you actually check the technicalities and see whether people can use the tool or the activity right? And so, I would very much advise everyone to have that before they introduce any technology in the classroom. {orally}

P12	-

Source: author
Annotated text is presented between square brackets.
The source of speech is presented between curly brackets (chat, orally, or survey).
*Participant code.

III Category - ASR Affordances

P*	Comment
P1	With short questions the intonation is ok-ish, maybe my previous question was too long... It's certainly a good resource, though. {survey}
P2	-
P3	It was a good moment to rethink about google translate. {survey}
P4	-
P5	[After the hands-on part] You should take Google translator with a grain of salt for sure. So, it [Google Translate] is important. It is a nice app, but you should take it with a grain of salt so students should use other resources as well. Well, we should use other resources as well as teachers. We shouldn't just, uh, trust what Google Translate says we should use other sources as well. {orally} [Final Remarks] When it comes to Google Translate, I'm going to start doing some things with my students. Which is basically teaching them that there's more to Google Translate than just typing the word, 'cause, I mean, let's be clear, everybody uses Google Translate and it came a long way since, uh, it started. I would say… Google Translate, It's now in elementary school, whereas Alexa is in kindergarten. 'cause when you talk to Alexa it doesn't get what you're saying. Even though you're speaking in Portuguese or in English, whereas when you work with Google Translate it came a long way from when it, uh, started, so I think it's improved a lot and I'm going to start teaching my students not to just type the word, but let's imagine they heard… they listen to a song, and they got a word, but they don't know how to spell that word. So, OK, they can pronounce it. So, I think it's important for us, teachers, not only, uh, provide them these apps or, I don't know, things that they can use, but also explain the other features they could use. {orally}
P6	-
P7	[While practicing tongue-twisters with Google Translate's ASR feature] For me it worked well, but i think i spoke very slowly […] now i spoke faster and it worked too! {chat}
P8	I have been using Siri too in some of my lessons, however I have never thought I could use Google Translate like that. I know for sure it helped me. {survey}
P9	Pronunciation is often completely neglected in language instruction, and this is a tool that can really help. {survey}

P10	[While practicing tongue-twisters with Google Translate's ASR feature] it worked pretty well! {chat} [Final Remarks] I'm an English teacher and I usually use the Google Translate in my classrooms, but not in this perspective, yes? so this perspective is really interesting in your research. It is really interesting too. Congratulations and I like it so much. [...] It's another vision, another way [of using Google Translate], and probably I will use it in the future with my students. {orally}
P11	[While exploring Google Translate] Just a shame GT can't save translations longer than 300 characters... {chat}
P12	[While exploring Google Translate's ASR feature] The app recognized my speech very fast! {chat}

Source: author
Annotated text is presented between square brackets.
The source of speech is presented between curly brackets (chat, orally, or survey).
*Participant code.

IV Category - Activities Appraisal

P*	**Comment**
P1	The activities were really good! (*Translated speech from Portuguese*) {survey}
P2	-
P3	-
P4	[Answer to the question "Do you think your students would enjoy the activities?] I think it will depend on the profile of your students. So, for example, I'm working with teenagers, and they don't like anything… like my group of students at the moment. They don't like anything, so I know that this type of activities wouldn't work with them, but on the other hand I have other groups... but yeah, it would work for sure. So, I think it's part of your job to know your students well and then you can decide which activities would work according to the profile of your students. {orally} I think this [activity 1] could be also like a nice way to use the technology if you are trying to raise awareness to the intonation. {orally}

P5	[After the hands-on part] I just wanna comment on the on the second activity. I really liked the second activity and even though it is intended for beginners. I wouldn't say it's just beginners 'cause most of our learners, they have questions about “oh is it cheap? Is it cheap? Is it …?” Yep, so I really liked the activity even though I work with private students, I think for people that work with other classroom settings it would be nice as a variation for them to be like working in pairs and they would come up with some sentences and then the partner would have to guess what is the word that the person is trying to say. But it's also very important for the teacher to, yeah, monitor and check if the word is being is not being mispronounced, But I'm totally going to use the second activity and the last one you gave. It's more like for B1, B2 [learners], I don't know… a bit more advance, but I also really liked it. […] But yeah, just that I'm totally going to use the second activity. I think it's really important for them to know the difference between these words and the example of “fill” and “feel”… I, uh, I gave two options: [speaking to Google Translate] I said “can you fill?” Not “[fill] it” I just said “can you fill” without the “it” in the end, so “can you fill” and then the Google Translate transcribed sentiu [feel] and then I said “can you fill my glass”, and then it [Google Translate] put like “fill” encher, so sometimes it works by the context as everybody was saying before so. {orally} [After returning from the breakout room] So, as we were discussing [on the breakout room], we thought the activities 1 and activity 2 would be the ones that we would use the most. Most mainly because of the fact that the third one, the shadow reading, would be a bit tricky, not just for the teacher but also for the students to be able to perform it well even though it is a nice activity, it would demand a lot of time to be able to do it right? 'cause you don't want to do a pronunciation activity where it's not effective. So you don't want to do it like “oh, let me do it in 5 minutes”, just 'cause I think it's nice... it [the pronunciation activity] needs to be, uh, effective, right? So, uh, the tongue-twister [activity 5], the second one about the words that have similar pronunciation, and the first one - to complete with their own information - would be the ones we [participants] would use the most. {orally} [Final Remarks] I'm gonna for sure use some of the activities you provided into my practice. {orally}
P6	I loved the tips! (*Translated speech from Portuguese*) {survey}
P7	[the workshop provided] a wide range of activities that suit all groups of students (and may also be adapted). {survey} I loved the activities you created. I wanna use it with my Portuguese students as well {chat} [Answer to the question "Do you think your students would enjoy the activities?] By the way, the activities, uh, that you created? Yes, I would definitely say yes, but I need to learn about the context that they're taking the classes, that's why, uh, maybe, sometimes like homework is a better idea… that's why, like, they all [the activities] offer alternatives… like, homework is always an alternative. Uh, if you teach private students is one thing… a class is a completely different thing, right? You have to say if it's possible or not. {orally}
P8	-
P9	-

P10	I really liked the research and the activities! {survey}
P11	-
P12	-

Source: author
Annotated text is presented between square brackets.
The source of speech is presented between curly brackets (chat, orally, or survey).
*Participant code.