ESL learners' perceptions of using High-Variability Phonetic Training (HVPT) through YouGlish to improve pronunciation skills

RIZGAR QASIM MAHMOOD
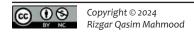
University of Wollongong, Australia

Abstract

This study explores ESL learners' perceptions of using high-variability phonetic training (HVPT) through YouGlish to improve pronunciation skills. Data were collected through an online questionnaire completed by 70 upper-intermediate level ESL learners and semi-structured interviews with 20 learners. The collected data were analyzed statistically and thematically. The results revealed that most learners had significantly positive attitudes towards using this online tool. They reported that YouGlish assisted them in accurately learning how to pronounce words and phrases from a variety of accents, including Australian, British, and American. Additionally, while learners had positive perceptions of using HVPT and YouGlish for pronunciation improvement, they emphasized the critical role of teachers in providing explicit corrective feedback (CF) and additional clarifications. Furthermore, the participants also reported some drawbacks of YouGlish. This study offers both theoretical and pedagogical implications for language learning and teaching.

Keywords: Accents; ESL learners; corrective feedback; High-Variability Phonetic Training (HVPT); pronunciation skills; YouGlish

Introduction

Pronunciation is a fundamental skill in spoken language, essential for effective communication. It significantly affects speakers' confidence levels. Unfortunately, it is often considered the most overlooked aspect of language learning. Levis and McCrocklin (2018) emphasized that pronunciation is essential to verbal communication, noting that every speaker has a unique accent. They argue that speaking a language without pronunciation is impossible, and understanding others requires adapting to their pronunciation. Similarly, Szpyra-Kozłowska (2014) asserted that "pronunciation is an important component of language without which no efficient oral communication is possible... people with poor pronunciation often lack the confidence to speak

up and try to say as little as possible" (pp. 2–4). Thus, teaching pronunciation is vital because "phonetic errors may lead to misunderstandings and even communication breakdowns" (p. 2). In a more recent work on pronunciation, Pennington and Rogerson-Revell (2019) asserted, "It is the crucial starting point for all spoken language since thoughts must be articulated in sound to be heard, and so to become a message that can be communicated to another person" (p. 1). Furthermore, they have seen pronunciation as "the common ground between speaker and listener that is needed for effective communication" (p. 1).

Previous studies have shown that teachers have used various approaches and techniques to teach pronunciation, including integrating technology into L2 pronunciation teaching. Given the rapid advances in Information and Communication Technology (ICT) worldwide, increasing interest is in integrating technologies into classrooms to equip learners for a progressively technology-dependent culture (Rahmati et al., 2021). While technology can be applied across various educational fields, it is particularly advantageous for enhancing pronunciation skills. It has been used in multiple ways to support pronunciation learning. An early example involved enabling students to practice recognizing and producing second language (L2) phonological properties, encompassing both segmental and suprasegmental elements. Blake (2011) noted that "Web-facilitated, hybrid, or fully virtual classes can provide learning environments conducive to successful second language development when properly integrated into the curriculum" (p. 19). Additionally, Goodwin (2020) stated that "effective pronunciation teaching involves (a) understanding the basic sound system of the language being taught, (b) setting realistic goals for achievement, (c) having an overall plan for instruction, and (d) incorporating a variety of techniques, tools, and technology" (p. 941). Previous research has reported mixed results regarding the effectiveness of technology in teaching pronunciation. In addition, inconclusive L2 learners' perceptions of using technology while teaching and learning pronunciation have been reported. The current study addresses the gap by examining ESL learners' perceptions of using High-Variability Phonetic [pronunciation] Training (HVPT) through YouGlish as the primary technological source for pronunciation improvement. HVPT has become a significant and effective second-language (L2) learning method. This teaching approach involves exposing learners to phonetic variations within the target language to improve their pronunciation and overall phonological competence (Mahdi & Mohsen, 2024). However, this study examines L2 learners' perceptions of HVPT and incorporates YouGlish as an authentic source of L2 input. To the researcher's knowledge, not much research has examined the topic under investigation in this approach. To fill this gap, the following research questions are addressed:

1. What are ESL learners' perceptions of using HVPT and YouGlish as online tools to improve their pronunciation?
2. To what extent is using HVPT and YouGlish adequate for L2 pronunciation improvement?

Literature Review

Conceptual Framework

The conceptual framework for the current study is drawn on Krashen's (1985) Input Hypothesis and Vygotsky's Sociocultural Theory (Lantolf, 2011; Vygotsky, 1978, 1986) to comprehensively understand how these methods facilitate pronunciation improvement. Krashen's Input Hypothesis suggests that language learners acquire new language structures most effectively when exposed to comprehensible input slightly beyond their current level of proficiency (i+1). HVPT aligns with this theory by providing diverse phonetic inputs, which are crucial for learners to grasp the phonological aspects of the target language (TL). Furthermore, the natural

order of language acquisition is supported as learners encounter varied phonetic contexts, promoting phonological development. The Affective Filter Hypothesis also plays a role, as engaging and authentic materials from YouGlish help lower learners' affective filters, making them more receptive to learning.

Vygotsky's Sociocultural Theory emphasises the importance of social interaction and the Zone of Proximal Development (ZPD) in cognitive development. YouGlish serves as a mediational tool, providing learners with immediate access to authentic pronunciation examples, which can scaffold their learning within the ZPD. Social interaction is also integral, as learners can discuss their experiences and receive feedback from peers or instructors, enhancing their phonetic skills through observation and imitation.

The study investigates ESL learners' perceptions of HVPT through YouGlish by examining how they perceive the effectiveness of varied phonetic input and the role of YouGlish as a mediational tool. It also assesses the impact of social interactions on their learning outcomes. By integrating HVPT with YouGlish, the study aims to determine how these methods can improve pronunciation skills and increase learners' confidence in their language abilities. Combining these theoretical perspectives offers a nuanced understanding of the potential benefits and challenges of using HVPT through YouGlish in ESL pronunciation training.

Technology and L2 Pronunciation Teaching

Digital technologies like smartphones, computers, and tablets have become integral to our daily lives. Many teachers and learners also frequently use these technologies in language classrooms and for self-study, though not universally. The field of computer-assisted language learning (CALL) has seen significant growth in recent decades, and interest in computer-assisted pronunciation training (CAPT) has similarly increased, leading to a surge in web-based and mobile apps and resources (Rogerson-Revell, 2021). Digital technologies, generally, offer a significant advantage in their ability to motivate and engage learners, especially those who are young or digitally proficient. However, these tools and resources must be user-friendly and reliable to avoid discouraging less experienced users. Furthermore, technology has profoundly impacted teaching and learning environments by diversifying learning contexts. In the past, educators often categorized teaching as either face-to-face or online. Today, however, both face-to-face and online settings offer a multitude of options (Kessler, 2018).

Technology is crucial in enhancing pronunciation skills in second language (L2) learning contexts. It facilitates various approaches, such as practising L2 phonological elements and integrating virtual and hybrid classes effectively into curricula (Blake, 2011). Goodwin (2020) underscores that effective pronunciation teaching involves understanding language sound systems, setting realistic goals, planning instruction comprehensively, and utilizing diverse techniques and technologies.

Researchers like Fouz-González (2020) and O'Brien et al. (2018) highlight technology's potential to improve pronunciation through personalized feedback and autonomous learning significantly, reducing teacher workload (Hincks, 2015; McCrocklin, 2016). The COVID-19 pandemic underscored technology's importance in enabling remote learning and preparing educators for unexpected disruptions (Herold, 2020; Pastor, 2020; UNESCO, 2023). Wallace and Choi (2022) argue that technology equips educators and students with essential tools for effective pronunciation instruction. Therefore, it can be inferred that technology enhances traditional teaching methods and addresses contemporary challenges in L2 pronunciation instruction, making learning more accessible, personalized, and resilient to disruptions like the COVID-19 pandemic.

Despite the increasing availability of resources for second language (L2) pronunciation, relatively limited attention has been given to the role of technology in this area. Wang and

Munro (2004) addressed this gap by examining the effectiveness of computer-based training in enhancing L2 learners' perception and production of segmental contrasts. Their study involved a trainee group of 16 participants (13 Mandarin, three Cantonese) and a control group of five participants (four Mandarin, one Cantonese). Over two months, the trainee group received training using a computer sound synthesizer focusing on three pairs of English vowels (i, ɪ, u, ʊ, ɛ, æ)—the training aimed to shift learners' attention from vowel duration to vowel quality. Pre-tests, post-tests immediately after training, and delayed post-tests three months later demonstrated significant improvements in perceptual performance and the ability to apply learned knowledge to new contexts among the trainee group. In contrast, the control group showed no noticeable improvement. Furthermore, Wang and Young (2015) developed ASR-based CALL to enhance English pronunciation among both adult and young learners. This system integrates formative and summative feedback approaches, utilizing implicit and explicit strategies tailored to learners' needs. The study involved two learner groups, highlighting that formative feedback supported improvements in speaking articulation, while summative feedback enabled learners to reflect on their progress. Adult learners benefitted from implicit feedback methods, such as model pronunciation and audio recasts within complete sentences, whereas young learners preferred explicit feedback, such as textual information, for self-correction. Additionally, the research identified that learners prefer different media modalities; audio feedback was found more effective for adults, whereas young learners preferred a combination of textual and audio feedback. Despite these positive findings, the study noted that there was no statistically significant difference between the pre-test and post-test scores of the young learners, indicating potential areas for further investigation or refinement in future research. It can be concluded that previous studies have employed various technological tools to enhance L2 learners' pronunciation skills. However, it is essential to investigate those tools' nature and specific types.

HVPT

HVPT involves exposure to various speech sounds through computer-assisted pronunciation training (CAPT). Learners are exposed to diverse speech samples from different speakers and phonetic contexts, aiming to enhance the perception of challenging non-native sounds (Melnik & Peperkamp, 2021). Logan et al. (1991) assert that HVPT is the most widely utilised method for improving the perception of L2 speech sounds, noting significant improvements in tasks like word recognition and perception of English word-initial /h/ even four months post-training.

HVPT utilises multiple natural examples of target sounds across varied phonetic contexts, often employing minimal pair examples (Logan et al., 1991; Shinohara & Iverson, 2018). Alternative terms for HVPT include high variability identification procedure (Pisoni et al., 1994), high variability perceptual training (Bradlow et al., 1997), and high-variability auditory training (Iverson et al., 2012). However, HVPT is widely accepted among scholars and researchers (Thomson, 2018), and recognised as a simple, cost-effective technique for aiding adults in learning new language sounds (Liberman, 2008). Thomson (2018) suggests that HVPT emerged to offer concentrated, high-quality training, resulting in perceptual improvements among advanced L2 learners previously considered resistant to change after initial exposure. O'Brien et al. (2018) emphasize the importance of incorporating diverse speakers and engaging input types, clear real-time feedback with visual aids, guidance on pronunciation improvement, and the ability to establish measurable goals for practical CAPT resources.

Thomson (2018) synthesized 32 studies and concluded that HVPT is highly effective in pronunciation training, noting that improvements persist over time. He defined HVPT as HVPT refers to perceptual training (most often focusing on segmental features), in which the auditory

training stimuli include numerous samples, produced by multiple talkers, in varied phonetic contexts" (p. 208). Similarly, Levis (2016) emphasized the importance of perceptual training in pronunciation teaching, highlighting that improved perception skills lead to better production skills and expand the options available for developing these skills.

Research across languages, including studies by Carlet and Cebrian (2015), Lee and Lyster (2016), and Shinohara and Iverson (2018), supports the effectiveness of HVPT and computer-based training in enhancing learners' ability to differentiate between vowels and consonants. Furthermore, Qian et al. (2018) highlighted the challenges learners face in perceiving L2 phonological contrasts and the impact on their speech intelligibility. Their study developed a computational approach for segmental perception training, successfully improving learners' ability to distinguish phonemic contrasts. Wiener et al. (2020) explored explicit phonetic instruction combined with high variability input in Mandarin tone production, showing that exposure to varied speech contexts enhances learners' ability to produce accurate tone contours when combined with explicit instruction. Based on the reviewed literature, it can be concluded that previous results emphasized HVPT's benefits and effectiveness. However, none of those studies have reported how the learners were exposed to target sounds and pronunciation across various contexts. Therefore, I am utilising YouGlish as the primary L2 pronunciation input provision source in the current study.

YouGlish and L2 Pronunciation Training

YouGlish functions as a video-assisted dictionary where learners can hear words pronounced in authentic contexts via YouTube videos, facilitating meaningful learning experiences (Saed et al., 2021). This tool allows students not only to hear word pronunciations from video characters but also to select different English accents (American, British, Australian), providing access to native or fluent speakers for accurate pronunciation, intonation, stress, and word usage (Fu & Yang, 2019; Sukmawati et al., 2024; Topal, 2023). Furthermore, YouGlish also offers phonetic transcription in both Modern IPA and Traditional IPA formats, enhancing learners' ability to understand pronunciation details (Mahmood, 2024). Additionally, the tool provides nearby words, further aiding learners in contextualizing and reinforcing their understanding of word usage and pronunciation nuances (Mahmood, 2024). In addition, learners using YouGlish can also control the speed of pronunciation (normal, slow, fast) and switch between different videos to ensure they capture all aspects of word pronunciation effectively (Fu & Yang, 2019). This tool allows for independent use, eliminating the need for teacher supervision, and encourages active engagement by enabling learners to explore English language usage in authentic contexts rather than passively consuming content (Saed et al., 2021). The effectiveness of YouGlish in enhancing speaking skills has garnered attention in the literature, prompting the need for further research to evaluate its benefits comprehensively.

Previous studies have investigated the effectiveness of using YouGlish and online videos to enhance L2 learners' pronunciation skills. For example, Fu and Yang (2019) investigated the impact of the YouGlish video-online pronunciation dictionary and found that it enhances oral communicative skills in pronunciation, intonation, and word usage. Similarly, Syafiq et al. (2021) observed in their study that using YouTube videos, including tools like YouGlish, significantly improves communicative speaking skills. They noted that besides pronunciation and intonation, other aspects of English learning, such as grammatical structures, vocabulary (lexis), fluency, and content proficiency, were also strengthened. Similarly, Quispe-Vargas et al. (2024) conducted a study to assess the impact of YouGlish on various subskills of English-speaking proficiency, including fluency, grammatical appropriateness, pronunciation,

communicative interaction, and vocabulary. The study found a significant improvement, with an average increase of 45.26 points between pre-test and post-test scores. This indicates that using YouGlish positively enhances English-speaking proficiency among university students. Based on the previous results, it can be concluded that YouGlish is effective in L2 pronunciation development. However, these studies have not examined how L2 learners perceive technological tools such as HVPT and YouGlish as the primary source of L2 pronunciation improvements. Therefore, it is crucial to examine how learners perceive these tools.

L2 Learners' Perceptions of HVPT and YouGlish

Learners' perceptions play a crucial role in shaping the effectiveness and outcomes of teaching and learning methods. Understanding how learners perceive instructional approaches, resources, and environments can influence their engagement, motivation, and, ultimately, their learning success (Black & Wiliam, 1998; Eccles & Wigfield, 2002; Hattie, 2009; Nguyen & Hung, 2021; Nguyen et al., 2021). Therefore, it is vital to understand how L2 learners perceive the technological tools used in L2 classrooms, especially to teach and learn L2 pronunciation.

Based on the existing literature, little has been written on L2 learners' perceptions of HVPT and YouGlish. However, several previous studies investigated learners' perceptions of YouGlish alone. For example, Sukmawati et al. (2024) found a mix of positive and negative feedback. Both types of feedback were used to gauge learners' perspectives on incorporating YouGlish to improve pronunciation skills. Similarly, Prastyo et al. (2022) examined university students' perceptions of using YouGlish to improve their pronunciation. The results indicated that students generally view the implementation of YouGlish positively, perceiving several benefits from its use in learning English pronunciation. The research findings highlight students' perceptions of YouGlish's impact on their pronunciation acquisition. In the same vein, Topal (2023) explored learners' perceptions of YouGlish as an L2 supporting source. The results revealed that participants identified authentic input, content diversity, and extra-curricular learning opportunities as strengths of the website. Conversely, weaknesses cited included the absence of progress checks and assessments and the perception of a competitive learning environment. Participants also reported incremental improvement in pronunciation at the segmental level.

In summary, the effectiveness of HVPT and YouGlish has been investigated based on the reviewed literature. However, these two techniques have been examined separately. Previous research has also explored L2 learners' perceptions of these techniques separately. Therefore, the current study aims to explore ESL learners' perceptions of HVPT and YouGlish as the primary authentic input sources for L2 pronunciation improvement.

Methodology

Participant and Context

This study was conducted in an ESL context where the learner participants had enrolled in English language courses in Australia. The study sample comprised 70 upper-intermediate proficiency level learners (25 males and 45 females). Their ages ranged between 19 and 35, with a mean age of 23. They were from various first language (L1) backgrounds. Most were from Latin American countries such as Brazil (*n* = 14), Colombia (*n* = 16), Ecuador (*n* = 5), Chile (*n* = 4), and Peru (*n* = 2), while some were from Asian countries, mainly Japan (*n* = 15), Thailand (*n* = 5), Vietnam (*n* = 3), Korea (*n* = 4), and Taiwan (*n* = 2). These learners took a placement test, and the results identified their levels as upper-intermediate. They were enrolled in the Cambridge First Certificate B1 course, focusing on the four primary language skills: speaking, listening, reading, and writing.

Data Collection Tool

The data were collected through a mixed-method approach: a well-structured questionnaire and a semi-structured interview. The survey questionnaire was designed using multiple-choice and Likert scale items. These items focused on exploring the participants' perceptions of using HVPT and YouGlish to improve their pronunciation skills. The data were collected at the end of a longitudinal course lasting 12 weeks. At the end of the course, the learners were asked to fill out the questionnaire anonymously. Subsequently, interviews were conducted with 20 volunteer participants better to understand their perceptions of the topic under exploration.

HVPT and YouGlish Implementation

In the first week of the course, learners were introduced to YouGlish. The teacher (i.e., the researcher) thoroughly demonstrated how to use YouGlish, explaining key features such as selecting various accents, adjusting the playback speed, and interpreting the phonetic transcription of the words. Learners were given practical examples and time to familiarize themselves with the tool through hands-on practice. In addition, the concept of HVPT was explained comprehensively, illustrating how it, combined with YouGlish, could help learners improve their pronunciation by exposing them to high-variability input from native speakers in authentic contexts.

Throughout the 12-week course, whenever any pronunciation issue or specific sound was taught or explained, YouGlish was incorporated into the instruction. Learners were consistently encouraged to use YouGlish for practice inside and outside the classroom but without explicit CF from the teacher. The teacher recommended that learners use YouGlish whenever practicing the learned words, encountering new words or phrases, and fostering continuous exposure to authentic pronunciation examples.

Importantly, learners also had access to YouGlish independently outside of class hours. This self-directed use of YouGlish provided valuable insights into learners' actual learning experiences. By tracking how frequently and in what contexts learners engaged with the tool beyond the classroom, it became possible to gather data on the software's impact on their pronunciation skills. This ongoing access and use allowed for a more comprehensive understanding of the role YouGlish played in their pronunciation development, as it extended the learning process beyond structured classroom activities and supported individual practice, reinforcing the principles of HVPT.

Data Analysis

Quantitative and qualitative data were collected, and two data analysis methods were executed. First, the quantitative data from the survey questionnaire were coded and then analysed descriptively using Jamovi (version 2.4.8). In contrast, the qualitative data from the interviews were analysed thematically, meaning that common themes and patterns were extracted and interpreted from the responses.

Results

The advancement of technology and the availability of several tools for pronunciation improvement have left L2 learners with many options for use during L2 learning and pronunciation development. Therefore, this study specifically focused on HVPT and YouGlish.

Descriptive Results

Seven constructs were generated from the survey questionnaire: Frequency of Using YouGlish, Improved Pronunciation, Understanding Accent Differences, Helpfulness for

Learning Pronunciation, Need for Explicit Feedback, Importance of Teacher Role, and Reading Phonetic Symbols. Participants were required to choose their responses on a Likert scale. The survey constructs reveal important insights into how ESL learners perceive the use of YouGlish and HVPT in improving their pronunciation skills. Table 1 shows the descriptive statistics for each construct.

As illustrated in Table 1, *the frequency of YouGlish use* is the first construct analysed. With a mean score of ($M = 4.21$, $SD = 0.740$), it is evident that the learners frequently used YouGlish. This indicates that learners believe HVPT and YouGlish to be effective methods to improve their pronunciation, and this is evident in the low value of the ($SD = 0.740$). Therefore, it can be inferred that the learners' active engagement with YouGlish suggests that it was a regularly utilized resource in their language learning process.

Moving to *improved pronunciation*, the results revealed that learners' perceptions of HVPT and YouGlish were positive ($M = 4.40$, $SD = 0.493$), indicating that their pronunciation has improved. Furthermore, the low SD of (0.493) indicates a high level of consensus and consistency among the learners regarding this benefit. This consistent perception is crucial as it highlights YouGlish and HVPT's perceived effectiveness in aiding pronunciation improvement, a core component of language proficiency.

In addition to the above two constructs, the questionnaire explored how HVPT and YouGlish assisted the participants' *understanding of different accents* and how helpful they were for Learning Pronunciation. The results revealed positive perceptions of both constructs equally ($M = 4.40$, $SD = 0.493$). It can be inferred that learners found YouGlish and HVPT helpful in improving their pronunciation and understanding different accents. The high scores suggest that learners appreciated the exposure to various accents through YouGlish, which is vital for their overall listening and speaking skills in English.

One of the controversial topics in L2 acquisition is the provision of CF. Therefore, we aimed to explore whether HVPT and YouGlish are sufficient without the teacher providing feedback. The results showed that most learners desired explicit feedback ($M = 4.11$, $SD = 0.692$). However, according to the results, some learners might feel confident in their ability to self-correct using YouGlish and HVPT. In contrast, others might still prefer more traditional feedback from an instructor. These findings lead us to the following construct: the crucial role of teachers. The results highlight teachers' critical role in language learning and pronunciation ($M = 4.51$, $SD = 0.503$). Despite the availability and perceived usefulness of technological tools, learners still value the guidance and support provided by their teachers. The high median score of

Table 1 *Descriptive results of the learners' perceptions of using HVPT and YouGlish*

	N	Mean	SD	Minimum	Maximum
Frequency of YouGlish Use	70	4.21	0.740	3	5
Improved Pronunciation	70	4.40	0.493	4	5
Understand Accent Differences	70	4.40	0.493	4	5
Helpful for Learning Pronunciation	70	4.40	0.493	4	5
Explicit Feedback Needed	70	4.11	0.692	3	5
Teacher Role Importance	70	4.51	0.503	4	5
Reads Phonetic Symbols	70	2.39	0.490	2	3

5 indicates that many learners rated the importance of the teacher's role at the maximum level, reaffirming the teacher's integral role even in technology-enhanced learning environments.

One of the significant components of YouGlish is the provision of the phonetic transcription of any given word in Modern IPA and Traditional IPA. Therefore, we decided to explore if the participants used this feature to improve their pronunciation. According to the results in Table 1, reading phonetic symbols had a significantly lower mean score ($M = 2.39, SD = 0.490$). This result suggests learners were less engaged or found less value in reading phonetic symbols. The lower scores might indicate that learners either find phonetic symbols challenging to understand or less relevant to their learning preferences compared to more interactive tools like YouGlish.

In conclusion, the descriptive statistics reveal that learners frequently use YouGlish and find it highly beneficial for improving pronunciation, understanding accent differences, and overall pronunciation learning. They appreciate the role of teachers in their learning journey and have varying opinions on the necessity of explicit feedback. The lower engagement with reading phonetic symbols suggests a preference for more practical and interactive learning tools. These insights can inform educators and curriculum designers in enhancing ESL learning experiences by integrating practical technological tools while maintaining essential teacher support.

Interview Results

The qualitative data gathered through semi-structured interviews was used to better understand how learners perceive HVPT (High Variability Phonetic [Pronunciation] Training) and YouGlish after using them for 12 weeks to improve their pronunciation. This qualitative approach was necessary to complement the quantitative data, which could not provide detailed insights into the learners' experiences and preferences regarding these tools. To fulfill the aim, these questions were asked: *1) What is your perception of HVPT and YouGlish after using them for 12 weeks to improve your pronunciation? 2) What are the advantages of using them? 3) What are the drawbacks of HVPT and YouGlish?* Then, based on the collected data, the following themes were constructed.

Perception of HVPT and YouGlish

Participants generally expressed positive sentiments when asked about their perception of HVPT and YouGlish after 12 weeks of use, particularly highlighting YouGlish. One participant emphasized its practicality, stating, "YouGlish is a very beneficial online and technology tool that I can use whenever I am unsure how to pronounce a word." This sentiment was echoed by others who appreciated the tool's ability to access authentic pronunciation examples. For instance, one participant noted, "It is a great online source to listen to authentic and accurate pronunciation by native speakers." The tool's capability to offer multiple examples for practising pronunciation and accent was also praised, with a participant commenting, "It helps you to learn the accurate pronunciation and helps you to practice your desired accent with as many examples as possible."

Advantages of Using YouGlish

Participants identified several advantages of using YouGlish. One significant benefit was its role in helping learners notice and correct their pronunciation errors. One participant said, "YouGlish helps me to notice my pronunciation and the errors I make with native speaker's pronunciation, and I can repeat and learn accurate pronunciation." The tool's ability to expose learners to various accents and improve their comprehension in real-life conversations was also highlighted. For example, one participant shared, "YouGlish is great for learning how to listen

to various accents, and it helped me to understand different speakers better when I was talking to them at work." Another participant felt that YouGlish boosted their confidence in pronunciation and accent, stating, "YouGlish helped me to notice that I make different sounds than native speakers, and it helped me practice and repeat them more. Now, I am more confident with my pronunciation and accent."

Furthermore, the interview results revealed that YouGlish was not only seen as effective and beneficial for sound production, such as vowels and consonants (i.e., segmental features), but also perceived as an effective tool for improving stress and intonation. For example, a participant stated, "Before using YouGlish, I knew *'record'* can be used as a *verb* and *noun*, but I never knew the differences and always pronounced it the same. However, now, I know how stress changes its form from a noun to a verb."

Drawbacks of HVPT and YouGlish

Despite the positive feedback, participants also pointed out some drawbacks of HVPT and YouGlish. A common issue was the difficulty in understanding phonetic symbols, as one participant expressed, "I believe YouGlish is a great resource, but I cannot read the phonetic symbols, and I really need a teacher to give me explicit feedback on the sounds that I cannot make even after listening and watching the examples on YouGlish." This need for explicit feedback from a teacher was a recurring theme, with another participant stating, "I used YouGlish, and it was great, but I need to learn pronunciation from a teacher as they give us very accurate and explicit feedback." Furthermore, while YouGlish was appreciated for its utility, some participants felt that teachers could offer more tailored feedback and better address individual pronunciation challenges. One participant summarized this sentiment by saying, "I believe this tool is helpful, but teachers can be more effective as they give you more feedback and they understand what problems we have during speaking and pronunciation, and as a Spanish speaker, I know we usually face challenges when it comes to sounds like /ð/, /θ/, /ʃ/, and /tʃ/, therefore, teacher's explicit feedback would be more effective and beneficial. Similarly, the participants from the mentioned Asian countries also reported similar challenges with specific sounds, such as /r/ and /l/, especially among the Japanese learners. They emphasized that in-class training and explicit CF would be more effective than listening to various speakers or watching videos.

Discussion

This study examined ESL learners' perceptions of using HVPT and YouGlish to enhance their pronunciation. This section analyses the current study's findings concerning the research questions and in the broader context of existing research.

The first research question addressed learners' perceptions: "What are ESL learners' perceptions of using HVPT and YouGlish as online tools to improve their pronunciation?" The quantitative and qualitative data analyses revealed that learners perceived these tools as beneficial for improving pronunciation. This perception aligns with previous research highlighting the positive impact of technology-assisted pronunciation tools on L2 learning (O'Brien et al., 2018; Rogerson-Revell, 2021; Wang & Munro, 2004). In particular, learners noted that YouGlish's feature of offering multiple accents (American, British, and Australian) helped boost their confidence when communicating with native speakers. This finding echoes prior work by Darcy (2018), who pointed out that exposure to multiple accents improves learners' ability to perceive and reproduce pronunciation variations accurately. YouGlish's role in enabling repeated exposure and practice also helped learners identify and correct mispronunciations, reinforcing the findings of Kartal and Korucu-Kis (2020).
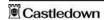
However, regarding HVPT, the study findings require a more detailed examination of its fundamental principles and how they were reflected in the learners' experiences. HVPT, or high-variability phonetic training, is rooted in two fundamental principles: high-variability input and minimal pair contrasts (Logan et al., 1991). High-variability input involves exposing learners to diverse samples of the target sounds produced by different speakers across varied phonetic contexts to help learners generalize their learning to new situations (Bradlow et al., 1997). Extensive research supports this principle, showing that repeated exposure to varied pronunciations enables learners to internalise a second language's phonetic features (Thomson, 2018). The minimal pair contrast strategy ensures learners focus on specific phonemic distinctions that are particularly challenging in the L2 context (Iverson et al., 2012). For example, learners practice contrasting sounds like /r/ and /l/ to enhance their ability to differentiate sounds that are not distinct in their L1.

In this study, learners reported that HVPT, when integrated with YouGlish, helped them better understand and practice complex phonological contrasts, especially when listening to native speakers in varied contexts. This reflects the core principle of high-variability input, as learners were exposed to authentic linguistic samples from various speakers through YouGlish. However, participants also needed more explicit feedback on specific sounds that continued to challenge them, such as /r/ and /l/ for Japanese learners and /ʃ/ and /tʃ/ for Spanish-speaking learners. This finding suggests that while HVPT offers substantial benefits, it may be insufficient without the complement of targeted CF from instructors—an issue also highlighted by Prastyo et al. (2022). The participants' reliance on teacher feedback suggests that the minimal pair aspect of HVPT alone may not fully address all learners' needs in cases where phonological distinctions are complicated.

The second research question asked: "To what extent is using HVPT and YouGlish adequate for L2 pronunciation improvement?" While both tools were perceived as helpful, the study identified gaps in the efficacy of these tools when used independently. For instance, learners noted that HVPT and YouGlish provided substantial input but lacked detailed corrective mechanisms for challenging sound distinctions. These findings align with previous research that advocates combining technological tools with explicit instruction from teachers to optimize pronunciation learning (Mahmood, 2024). Participants in this study recognized the teacher's essential role in offering explicit pronunciation guidance, echoing findings from other studies that emphasized the importance of CF in L2 pronunciation training (O'Brien et al., 2018; Thomson, 2018).

The findings of this study can be contextualised within the theoretical frameworks of Krashen's (1985) Input Hypothesis and Vygotsky's Sociocultural Theory (1986). Krashen's Input Hypothesis, which posits that learners benefit most from comprehensible input slightly beyond their current proficiency level (i+1), is supported by YouGlish's and HVPT's provision of authentic, varied linguistic input. Both tools help learners develop phonological awareness and improve pronunciation skills by exposing learners to diverse pronunciations in real-world contexts. Similarly, Vygotsky's Sociocultural Theory, particularly the Zone of Proximal Development (ZPD) concept, highlights the importance of social interaction and scaffolding in learning. YouGlish, as a mediational tool, supports learning through repeated exposure to authentic speech in varied linguistic contexts, helping learners move towards more accurate pronunciation within their ZPD.

In conclusion, HVPT and YouGlish substantially benefit L2 pronunciation training by providing high-variability input and authentic linguistic contexts. However, they are most effective when complemented by explicit teacher CF and CF strategies. This balanced approach is essential to fully support learners' pronunciation development at all levels, particularly for phonetic contrasts that remain difficult despite extensive exposure.
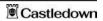
Conclusion

This study underscores the potential of integrating technological tools like HVPT and YouGlish to enhance learners' pronunciation. Participants perceived these tools positively, benefiting from access to diverse pronunciation models and opportunities for self-directed learning. However, effectively utilizing these tools requires proper training and instructional guidance to maximize their impact. Despite their perceived benefits, the study highlights the indispensable role of teachers in pronunciation instruction. Participants needed explicit instructor feedback and guidance, particularly in addressing pronunciation challenges. This underscores that while technological tools can facilitate independent learning and exposure to authentic language input, they complement rather than replace the crucial role of teachers in providing tailored instruction and CF. Effective language instruction should integrate technological resources and traditional teaching methods to comprehensively support ESL learners' pronunciation development.

The findings of this study offer theoretical insights into the effective integration of technological tools like HVPT and YouGlish in ESL pronunciation instruction. They underscore the validity of Krashen's Input Hypothesis, demonstrating that providing learners with authentic and varied language input can enhance pronunciation skills by fostering phonological awareness and accurate pronunciation patterns. Moreover, Vygotsky's Sociocultural Theory is supported by YouGlish as a mediational tool within the ZPD, facilitating interactive learning experiences and enabling learners to receive timely feedback crucial for pronunciation improvement. Pedagogically, these insights suggest that while technological tools offer valuable resources for independent learning and exposure to diverse language models, their practical implementation requires structured training and instructional support. Integrating YouGlish and HVPT into ESL curriculum design can enrich learning environments by supplementing traditional teaching methods with engaging, interactive tools. However, to optimize learning outcomes, instructors must play a central role in providing explicit pronunciation instruction and personalized feedback tailored to learners' needs.

Finally, several limitations should be acknowledged when interpreting the findings of this study. The study's reliance on self-reported perceptions of pronunciation improvement through HVPT and YouGlish may introduce response bias, potentially overestimating the tools' effectiveness. Objective measures of pronunciation proficiency and longitudinal studies are needed to validate the perceived benefits of these technological tools over time. Additionally, while participants valued YouGlish and HVPT for their accessibility and diverse content, challenges such as the need for more targeted feedback from teachers highlight ongoing areas for improvement in pronunciation instruction. Future research should address these limitations by employing larger, more diverse samples and incorporating objective assessments to understand better the nuanced impact of technological tools on ESL pronunciation development.
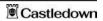
References

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Blake, R. J. (2011). Current trends in online language learning. *Annual Review of Applied Linguistics*, *31*, 19–35. https://doi.org/10.1017/S026719051100002X

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *J Acoust Soc Am*, *101*(4), 2299–2310. https://doi.org/10.1121/1.418276

Carlet, A., & Cebrian, J. (2015). Identification vs. discrimination training: Learning effects for trained and untrained sounds. International Congress of Phonetic Sciences.

Darcy, I. (2018). Powerful and effective pronunciation instruction: How can we achieve it? *CATESOL Journal, 30*(1), 13–45. https://eric.ed.gov/?id=EJ1174218

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(2002), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Fouz-González, J. (2020). Using apps for pronunciation training: An empirical evaluation of the English File Pronunciation app. *Language Learning & Technology*, *24*(1), 62–85. https://doi.org/10125/44709

Fu, J. S., & Yang, S.-H. (2019). Exploring how YouGlish facilitates EFL learners' speaking competence. *Journal of Educational Technology & Society*, *22*(4), 47–58. https://www.jstor.org/stable/26910184

Goodwin, J. (2020). Pronunciation teaching methods and techniques. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 941–950). Wiley-Blackwell.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Herold, B. (2020). The scramble to move America's schools online. https://www.edweek.org/technology/the-scramble-to-move-americas-schools-online/2020/03

Hincks, R. (2015). Technology and learning pronunciation. In *The handbook of English pronunciation* (pp. 505–519). https://doi.org/10.1002/9781118346952.ch28

Iverson, P., Pinet, M., & Evans, B. G. (2012). Auditory training for experienced and inexperienced second-language learners: Native French speakers learning English vowels. *Applied Psycholinguistics*, *33*(1), 145–160. https://doi.org/10.1017/S0142716411000300

Kartal, G., & Korucu-Kis, S. (2020). The use of Twitter and YouGlish for the learning and retention of commonly mispronounced English words. *Education and Information Technologies*, *25*(1), 193–221. https://doi.org/10.1007/s10639-019-09970-8

Kessler, G. (2018). Technology and the future of language teaching. *Foreign Language Annals*, *51*(1), 205–218. https://doi.org/10.1111/flan.12318

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.

Lantolf, J. P. (2011). The sociocultural approach to second language acquisition: Sociocultural theory, second language acquisition, and artificial L2 development. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (pp. 24–47). Routledge.

Lee, A. H., & Lyster, R. (2016). Effects of different types of corrective feedback on receptive skills in a second language: A speech perception training study. *Language Learning*, *66*(4), 809–833. https://doi.org/10.1111/lang.12167

Levis, J. M. (2016). Research into practice: How research appears in pronunciation teaching materials. *Language Teaching*, *49*(3), 423–437. https://doi.org/10.1017/S0261444816000045

Levis, J. M., & McCrocklin, S. (2018). Reflective and effective teaching of pronunciation In M. Zeraatpishe, A. Faravani, H. R. Kargozari, & M. Azarnoosh (Eds.), *Issues in applying SLA theories toward reflective and effective teaching* (pp. 77–89). Brill https://doi.org/10.1163/9789004380882_007

Liberman, M. (2008). HVPT. *Language Log*. https://languagelog.ldc.upenn.edu/nll/?p=328

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: a first report. *J Acoust Soc Am*, *89*(2), 874–886. https://doi.org/10.1121/1.1894649

Mahdi, H. S., & Mohsen, M. A. (2024). Enhancing pronunciation learning through High Variability Phonetic Training: A meta-analysis. *Language Teaching Research Quarterly*, 29–45. https://doi.org/10.32038/ltrq.2024.40.0240

Mahmood, R. Q. (2024). The impact of visual corrective feedback on pronunciation accuracy in L2 sound production: Empirical evidence. In T. C. Bang, C. H. Nguyen, & H. P. Bui (Eds.), *Exploring contemporary English language Education practices* (pp. 158–189). IGI Global. https://doi.org/10.4018/979-8-3693-3294-8.ch008

McCrocklin, S. M. (2016). Pronunciation learner autonomy: The potential of Automatic Speech Recognition. *System*, *57*, 25–42. https://doi.org/10.1016/j.system.2015.12.013

Melnik, G. A., & Peperkamp, S. (2021). High-Variability Phonetic Training enhances second language lexical processing: evidence from online training of French learners of English. *Bilingualism: Language and Cognition*, *24*(3), 497–506. https://doi.org/10.1017/S1366728920000644

Nguyen, L. T., & Hung, B. P. (2021). Communicative pronunciation teaching: Insights from the Vietnamese tertiary EFL classroom. *System*, *101*, 102573. https://doi.org/10.1016/j.system.2021.102573

Nguyen, L. T., Hung, B. P., Duong, U. T. T., & Le, T. T. (2021). Teachers' and learners' beliefs about pronunciation instruction in tertiary English as a foreign language education. *Front Psychol*, *12*, 739842. https://doi.org/10.3389/fpsyg.2021.739842

O'Brien, M. G., Derwing, T. M., Cucchiarini, C., Hardison, D. M., Mixdorff, H., Thomson, R. I., Strik, H., Levis, J. M., Munro, M. J., Foote, J. A., & Levis, G. M. (2018). Directions for the future of technology in pronunciation research and teaching. *Journal of Second Language Pronunciation*, *4*(2), 182–207. https://doi.org/10.1075/jslp.17001.obr

Pastor, C. K. L. (2020). Sentiment analysis on synchronous online delivery of instruction due to extreme community quarantine in the Philippines caused by COVID-19 pandemic. *Asian Journal of Multidisciplinary Studies*, *3*(1), 1–6. https://asianjournals.org/online/index.php/ajms/article/view/207

Pennington, M. C., & Rogerson-Revell, P. (2019). *English pronunciation teaching and research: Contemporary perspectives*. Palgrave Macmillan.

Pisoni, D. B., Lively, S. E., & Logan, J. S. (1994). Perceptual learning of nonnative speech contrasts: Implications for theories of speech perception. In *The development of speech perception: The transition from speech sounds to spoken words.* (pp. 121–166). The MIT Press.

Prastyo, Y. D., Dharmawan, Y. Y., & Amelia, S. F. (2022). Student's perceptions on the implementation of YouGlish in learning English pronunciation at English department. *Lectura: Jurnal Pendidikan*, *13*(1), 42–54. https://doi.org/10.31849/lectura.v13i1.9225

Qian, M., Chukharev-Hudilainen, E., & Levis, J. (2018). A system for adaptive high-variability segmental perceptual training: Implementation, effectiveness, transfer. *Language Learning & Technology*, *22*(1), 69–96. https://doi.org/10125/44582

Quispe-Vargas, M., Laura-De La Cruz, K. M., Talavera-Mendoza, F., Manzur-Vera, G., Pérez-Postigo, G., Turpo-Gebera, O., & Diaz-Zavala, R. (2024). The impact of YouGlish on English speaking competency in higher education [English language, speaking competency, YouGlish, video]. *2024*, *14*(2), 14. https://doi.org/10.3926/jotse.1710

Rahmati, J., Izadpanah, S., & Shahnavaz, A. (2021). A meta-analysis on educational technology in English language teaching. *Language Testing in Asia*, *11*(1), 7. https://doi.org/10.1186/s40468-021-00121-w

Rogerson-Revell, P. M. (2021). Computer-assisted pronunciation training (CAPT): Current issues and future directions. *RELC Journal*, *52*(1), 189–205. https://doi.org/10.1177/0033688220977406

Saed, H. A., Haider, A. S., Al-Salman, S., & Hussein, R. F. (2021). The use of YouTube in developing the speaking skills of Jordanian EFL university students. *Heliyon*, *7*(7), e07543. https://doi.org/10.1016/j.heliyon.2021.e07543

Shinohara, Y., & Iverson, P. (2018). High variability identification and discrimination training for Japanese speakers learning English /r/−/l. *Journal of Phonetics*, *66*, 242–251. https://doi.org/10.1016/j.wocn.2017.11.002

Sukmawati, Romayanti, V., Abeng, A. T., Ibrahim, M., Hasanuddin, & Rasmiyati, A. (2024). Student's perceptions of YouGlish as instructional technology and media for learning English pronunciation. *IJOLEH : International Journal of Education and Humanities*, *3*(1), 86–95. https://doi.org/10.56314/ijoleh.v3i1.225

Syafiq, A. N., Rahmawati, A., Anwari, A., & Oktaviana, T. (2021). Increasing speaking skill through YouTube video as English Learning material during online learning in pandemic Covid-19. *Elsya: Journal of English Language Studies*, *3*(1), 50–55. https://doi.org/10.31849/elsya.v3i1.6206

Szpyra-Kozłowska, J. (2014). *Pronunciation in EFL instruction: A research-based approach*. Multilingual Matters.

Thomson, R. I. (2018). High Variability [Pronunciation] Training (HVPT): A proven technique about which every language teacher and learner ought to know. *Journal of Second Language Pronunciation*, *4*(2), 208–231. https://doi.org/10.1075/jslp.17038.tho

Topal, I. H. (2023). YouGlish as a pronunciation resource: Voices from Turkish EFL learners. *i-manager's Journal of Educational Technology*, *20*(3), 7.

UNESCO. (2023). *UNESCO's education response to COVID-19*. https://www.unesco.org/en/covid-19/education-response/initiatives

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Vygotsky, L. S. (1986). *Thought and language*. MIT Press.

Wallace, L., & Choi, J. (2022). Preparing second language educators to teach pronunciation with technology. In S. McCrocklin (Ed.), *Technological resources for second language pronunciation learning and teaching: Research-based approaches* (pp. 31–53). Rowman & Littlefield.

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System*, *32*(4), 539–552. https://doi.org/10.1016/j.system.2004.09.011

Wang, Y.-H., & Young, S. S.-C. (2015). Effectiveness of feedback for enhancing English pronunciation in an ASR-based CALL system. *Journal of Computer Assisted Learning*, *31*(6), 493–504. https://doi.org/10.1111/jcal.12079

Wiener, S., Chan, M. K. M., & Ito, K. (2020). Do explicit instruction and high variability phonetic training improve nonnative speakers' Mandarin tone productions? *Modern Language Journal*, *104*(1), 152–168. https://doi.org/10.1111/modl.12619