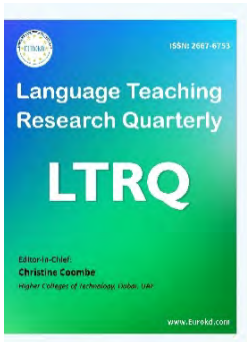


Language Teaching Research Quarterly

2024, Vol. 40, 29–45

Enhancing Pronunciation Learning through High Variability Phonetic Training: A Meta-analysis

Hassan Saleh Mahdi[1]*, Mohammed Ali Mohsen[2]

[1]Department of English, Faculty of Education, Hodeida University, Hodeida, Yemen
[2]College of Languages and Translation, Najran University, Najran, Saudi Arabia

Received *6 September 2023* ***Accepted*** *1 December 2023*

Abstract

High-Variability Phonetic Training (HVPT) has demonstrated effectiveness in second language (L2) acquisition. This study utilizes a meta-analysis to explore the influence of high variability on learning L2 pronunciation and identifies the factors that moderate this impact. The studies were collected using a keyword search in the SCOPUS database. In total, our meta-analysis incorporated 18 primary studies that presented results obtained from experimental and control group designs, encompassing a total of 22 effect sizes. The results of our meta-analysis revealed that the overall effect size of HVPT on L2 pronunciation was medium (g = 0.77). Specifically, the effect size was large for consonant sounds and lexical tones, particularly pronounced when utilized by advanced learners. For beginners, the effect size was medium. The context in which the training was conducted also played a role, with a medium effect size observed when sentences were used. Additionally, the number of talkers involved in the study influenced the effect size, with a medium effect size found when 5 to 8 talkers were included. In other cases, the reported effect sizes were either small or non-significant.

Keywords: *HVPT, Learners' Proficiency, Pronunciation, Meta-analysis*

Introduction

High-variability phonetic training (HVPT) has emerged as a notable and impactful approach in the realm of second language (L2) learning. This pedagogical method involves exposing learners to a diverse range of phonetic variations within the target language, thereby aiming to enhance their pronunciation skills and overall phonological competence. While HVPT has demonstrated its efficacy in several aspects of L2 acquisition, such as improved fluency and intelligibility, the existing body of research presents a more nuanced picture when it comes to its influence on consonants versus vowels. The core issue lies in the fact that results within the literature exhibit a degree of variability themselves, with researchers reporting differing

* Corresponding author.
E-mail address: almatari7@gmail.com
https://doi.org/10.32038/ltrq.2024.40.02

outcomes regarding the impact of HVPT on consonant and vowel sounds. In some studies, HVPT appears to yield significant improvements in learners' ability to accurately pronounce sounds (e.g., Bradlow et al., 2023; Mahdi et al., 2023; Zhang et al., 2023), contributing to clearer and more intelligible speech. Conversely, other investigations suggest that the effects of HVPT might not be as pronounced when it comes to vowel pronunciation (e.g., Giannakopoulou et al., 2017).

A few systematic reviews and meta-analyses were conducted to investigate the impact of High Variability Phonetic Training (HVPT) on the pronunciation of second language (L2) learners. For instance, Zhang et al. (2021) reviewed 18 studies. Their findings initially indicated a small effect size (Hedges' g = 0.46) for the immediate training results. However, this effect size decreased significantly (g = -0.04) after addressing outliers and accounting for publication bias. In contrast, when it came to immediate perceptual improvements, the effect size was medium (g = 0.56), but this was not the case for production gains, which were not statistically significant (g= −0.04). Additionally, the length of the training program and the format in which the talker presented the material were identified as potential factors that could influence immediate perceptual enhancements and the overall generalization of these gains. Another investigation, conducted by Thomson (2018) who synthesised the research on HVPT, scrutinizing 32 studies that employed HVPT as a method to enhance learners' ability to perceive and reproduce L2 sounds more accurately. The collective findings from these studies presented compelling evidence affirming the high efficacy of HVPT as a tool for pronunciation training, and moreover, indicated that the enhancements in pronunciation are enduring over time. While the previous synthesis and systematic reviews tackling push ahead to our understanding of the issue of HVPT, many unexplored moderators might affect the results of these studies. To address this gap, this study attempts to uncover many variables that remain not investigated in previous literature. It seeks to provide insights into how variables such as outcome measurement, target sounds, the number of stimuli, and context (i.e., words or sentences) may impact the incorporation of HVPT in L2 pronunciation learning.

Literature Review

HVPT

Individuals who acquire a second language (L2) during their adolescence or later often encounter challenges when it comes to perceiving and producing L2 sounds. These challenges can be mitigated through phonetic training, where L2 learners receive feedback regarding their perception and/or production of L2 sounds on a trial-by-trial basis. It has been observed that training incorporating highly diverse stimuli is more effective in facilitating the acquisition of L2 phonetic distinctions compared to training with limited variability, as evidenced by several studies (e.g., Lively et al., 1993; Logan et al., 1991; Wang et al., 1999).

The seminal work in this field was conducted by Logan et al. (1991). Their study aimed to assess whether training involving variability in speech sounds, stemming from diverse talkers with distinct voices and speech patterns, could assist learners in concentrating on pertinent acoustic differences when perceiving L2 sound distinctions. The outcomes indicated that participants who received training demonstrated enhanced proficiency in distinguishing between /r/ and /l/, particularly in specific phonetic settings. Also, Lively et al. (1994) replicated this procedure using monolingual Japanese speakers, affirming the improvements

and retention of knowledge three months later. Performance remained relatively stable even six months later, suggesting enduring alterations in how sound categories are stored in memory. HVPT gained recognition as an efficacious approach to speech perception training and evolved into a prominent research area within the field of speech sciences. Its implications for language instruction, particularly its potential to extend to novel listening contexts, are of considerable interest.

HVPT enhances learners' phonetic discrimination skills, enabling them to perceive subtle differences in sounds, leading to better pronunciation. It also boosts intelligibility by exposing learners to diverse speech patterns, making them effective communicators across various contexts. Early studies, such as the one conducted by Strange and Dittmann (1984) to teach Japanese speakers to distinguish between /r/ and /l/. These studies used synthetic stimuli but found limited generalization to naturally produced minimal pairs. Later research, exemplified by Logan et al. (1991), emphasized the importance of incorporating variability in training materials. They achieved success by including multiple natural exemplars spoken by various speakers and across different phonetic contexts. Further studies, like Lively et al. (1993) contrasted high variability (HV) input with low variability (LV) input, with HVPT becoming a standard in L2 phonetic training. This methodology has been applied to various language pairs, demonstrating its versatility in training learners on different phonetic contrasts. Additionally, some evidence suggests that HVPT can benefit production skills in addition to perception, as seen in studies by Bradlow et al. (1999) and Lambacher et al. (2005), although results are mixed in this regard. Despite the prevalence of HVPT in L2 phonetic training, few studies explicitly compared high and low variability training materials. Some studies, like Sadakata and McQueen (2013) found greater benefits for HV training in tasks of generalization, identification, and production. However, Giannakopoulou et al. (2017) did not find the same advantage for HV training in the English /i/-/ɪ/ contrast for adult participants.

Some studies examined the benefit of HVPT in improving lexical tone training. For example, Wang et al. (1999) employed a paradigm similar to Logan et al. (1991) and used four speakers for training. Participants were trained using real monosyllabic Mandarin words that varied in consonants, vowels, and syllable structure. The training involved identifying tones using diacritic representations. Native American English speakers showed significant improvement in tone identification, which generalized to new words and a new speaker. Similarly, Wang et al. (2003) explored whether this learning transferred to production, finding improvements primarily in pitch contour. In addition, Perrachione et al. (2011) directly compared high variability (HV) and low variability (LV) training materials in native American English speakers learning Mandarin tones. While LV outperformed HV during training, this trend did not continue during testing, suggesting that HV training was more effective for generalization. An interesting finding was the interaction between individual aptitude and training type, with high-aptitude learners benefiting more from HV training. In the same vein, Sadakata and McQueen (2014) also investigated input variability and individual aptitude in lexical tone training, using Dutch speakers learning Mandarin. They found a similar interaction between aptitude and variability, with high-aptitude learners benefiting more from HV training. Both studies suggest that HV training can facilitate learning tone contrasts, especially for individuals with higher aptitude. However, there was no consistent evidence that higher variability significantly enhanced generalization, which was somewhat surprising given the

expectation that exposure to multiple speakers should aid in dissociating tones from specific speakers. Several factors play a crucial role in determining the effectiveness of HVPT. They include learners' L2 proficiency, learning outcomes, target sounds, number of talkers, and number of stimuli.

Learners' L2 Proficiency

The proficiency level of second language learners plays a pivotal role in the effective utilization of HVPT. Some studies have suggested that variability can pose challenges for novice learners, as observed in studies by Chang and Bowles (2015) and Kingston (2003). For instance, Chang and Bowles (2015) discovered that individuals with no prior experience in learning Mandarin tones exhibited better performance in acquiring Mandarin tones when presented with monosyllabic words (which have less variation in tone contours) compared to disyllabic words (characterized by greater variation in tone contours). However, it is important to note that recent studies have presented a different perspective, suggesting that novice learners can still derive benefits from variability if they possess strong general perceptual abilities, such as the capacity to distinguish pitch contours. Antoniou and Wong (2015), Perrachione et al. (2011), and Sadakata and McQueen (2014) have demonstrated that novice learners with robust perceptual abilities can actually gain advantages from variability. The findings of these studies imply that variability might initially pose challenges for novice learners with limited perceptual abilities when learning L2 contrasts. Still, for those with heightened perceptual abilities, variability can be advantageous.

Learning Outcomes

Several studies have examined the strategies aimed at enhancing speech perception in a non-native language through perceptual phonetic training. This has been explored in various studies, including those conducted by Logan et al. (1991) and Lively et al. (1993). The predominant task commonly used in these training studies is the identification task. In this task, participants are presented with auditory stimuli and required to associate them with written words displayed on a screen. Feedback is provided after each response, as demonstrated in studies by Logan et al. (1991) and Huensch and Tremblay (2015). However, certain studies have incorporated both an identification task and a discrimination task. In the discrimination task, participants are called upon to determine whether the auditory stimuli they encounter are identical or not. This approach is exemplified in studies conducted by Cebrian and Carlet (2014) and Iverson et al. (2012). An advantage of incorporating the auditory discrimination task is that participants are not required to possess knowledge of how to read words in the language they are being trained in. Moreover, production is used in some studies (e.g., Davis, 2015; Dong et al., 2019; Zhang et al., 2021). The results of these studies indicate that perception yielded larger effect size than production tasks.

Target Sounds

Using HVPT to enhance L2 pronunciation can be influenced by the type of sounds taught. Generally, HVPT is used to improve L2 pronunciation of three categories: consonants, vowels, and lexical tones. In this context, lexical tones refer to the distinctive pitch patterns carried by the syllable of a word (Dong et al., 2019). Vowel sounds can be more challenging than

consonant sounds for several reasons. First, they are more sensitive to stress, requiring nuanced articulation and perception. Second, vowels carry more information in spoken language, making them crucial for comprehension. Third, the relationship between vowels and consonants can be complex due to phonotactic constraints, indicating that learning vowel sounds may involve more intricate rules. Fourth, vowel perception is tied to robust auditory processing, suggesting that it may be more demanding cognitively. Lastly, the perception of vowels is less stable across varying speaking rates, adding another layer of complexity. These challenges make vowels a particularly intricate component of language (Berg, 1990; Kachlicka et al., 2019; Kewley-Port et al., 2007; Magen, & Blumstein, 1993; Onishi et al., 2002). The findings of Zhang et al. (2021) found that all studies target consonant sounds yielded large effect sizes (e.g., Davis, 2015; Shehata, 2013). On the other hand, studies target vowel sounds yielded either non-significant or small or medium effect sizes (e.g., Brosseau-Lapré, 2013; Giannakopoulou, 2017; MacDonald, 2012; Zhang et al., 2021). Regarding lexical tones the results were conflicting ranging from negative effect size (e.g., Wiener, et al., 2020) to large effect size (e.g., Silpachai, 2020).

Number of Talkers

The impact of HVPT on improving L2 pronunciation can be significantly affected by the inclusion of a diverse number of speakers. Integrating multiple speakers into HVPT may be useful because exposing L2 learners to a wide range of speakers helps them become accustomed to the inherent diversity they may encounter when communicating with native speakers. This assumption warrants further investigation because two studies, Davis (2015) involving 10 speakers and Shehata (2013) involving 3 speakers, produced identical results with a large effect size. However, it's worth noting that quantifying talker variability solely based on the number of talkers may not provide a precise measure. Some talkers are likely to exhibit greater variability in their speech productions than others, and this within-talker variability may also impact perception (Newman et al., 2001). For example, it has been proposed that increased talker variability may be more prominent when talker identity carries more informative cues (Kleinschmidt, 2019). Consequently, further research in this area would be valuable to better understand the intricacies of talker variability and its effects on language perception and pronunciation learning.

Number of Stimuli

Stimuli in HVPT pertain to the speech samples or auditory inputs that are made available to learners of a second language during their training. These stimuli encompass various mediums such as recorded materials, audio clips, or live speech delivered by either native speakers or language models representing the target language. When learners are exposed to only a limited number of stimuli, it may not adequately prepare them for the wide array of pronunciation patterns they may encounter in real-life situations. Consequently, having a substantial pool of stimuli at their disposal enables learners to engage with a diverse range of phonetic expressions and accents, thereby enhancing their adaptability in communication.

Research in the field of second language learning support the idea that a greater number of stimuli in HVPT can result in more effective improvement of pronunciation. Some studies, including those by Lively et al. (1993), Logan et al. (1991), and Pisoni (1993), suggested that

because listeners are believed to store specific instances or exemplars, exposing them to a wide variety of natural stimuli proves superior to limited exposure. This is because the distributions of natural stimuli enable learners to discern which cues are reliable, thus contributing to the formation of a multidimensional categorization framework for these stimuli.

To find out the effect of HVPT on learning L2 pronunciation, a comprehensive analysis is needed. Consequently, the present study aims to address the following research questions:

RQ_1: What is the overall impact of incorporating HVPT into the process of learning second/foreign language pronunciation?

RQ_2: What are the effects of participant-related factors (e.g., L2 proficiency level) and treatment-related factors (e.g., target sounds, learning outcomes, number of talkers, number of stimuli, and context) on the utilization of HVPT as a tool for pronunciation learning?

Methods

Research Design and Research Methodology

In order to evaluate the impact of HVPT on second language pronunciation, a meta-analysis was carried out. Meta-analysis is a statistical technique used to combine findings from various experimental studies. The results of these studies are converted into effect sizes, which quantify the difference between the post-test scores of the experimental and control groups. This process involved several sequential stages, including: (1) conducting a thorough literature review; (2) defining the criteria for study selection; (3) encoding and organizing the data; and (4) calculating the effect sizes.

Searching the Literature

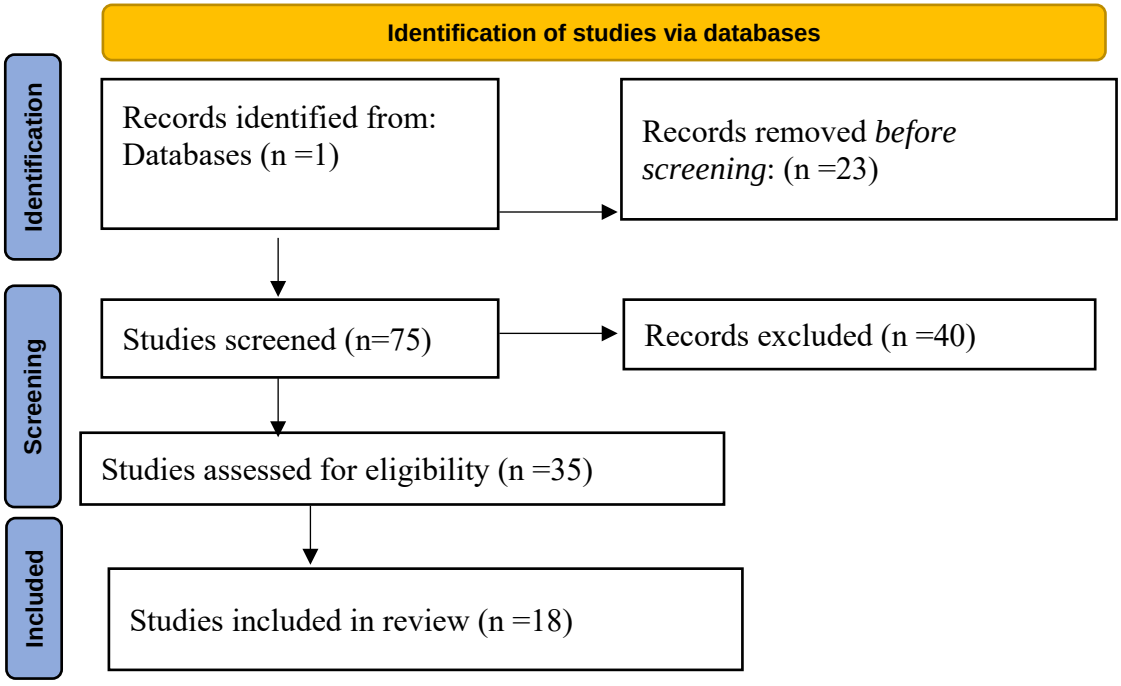
In our effort to identify appropriate studies for our analysis, we initiated a search using the SCOPUS database. We located a total of 75 articles matching the following keywords: 'High-variability phonetic training,' 'High-variability pronunciation training,' and 'HVPT.' Subsequently, we scrutinized the abstracts of these studies to ascertain their suitability for inclusion in our research. We did not impose any specific time constraints on this search and considered studies published up to September 9, 2023. Some studies were excluded from our analysis because they did not align with the predefined criteria for this meta-analysis. Consequently, we identified 18 articles that met our criteria and are marked with an asterisk (*) in the reference list.

Study Inclusion Criteria

In order to identify the appropriate studies for this analysis, specific inclusion criteria were established. Each study had to satisfy the following conditions: (1) It had to be an experimental or quasi-experimental study. (2) The central focus of the study had to revolve around the utilization of HVPT as a primary tool. (3) In the experimental group, participants were required to employ HVPT as a training tool. (4) The study had to report relevant statistical information such as means, standard deviations, and participant number for each group involved. (5) The primary dependent variable under investigation in the study had to be pronunciation achievement. The inclusion of studies meeting these criteria is depicted in Figure 1.

Figure 1

PRISMA Flowchart of the Present Meta-analysis



Coding of Study Characteristics

Several factors influence L2 pronunciation when using HVPT. In this meta-analysis, we examined eight factors, which were coded and used as moderators in the study. These moderators encompass learners' L2 proficiency, learning outcomes (i.e., identification or production), target sounds (consonants or vowels), treatment duration, the number of talkers, the number of stimuli, and the context (i.e., words and sentences).

Effect Size Calculation

In this meta-analysis, we employed Hedge's g as the measure to evaluate the effectiveness of HVPT on pronunciation. Hedge's g calculates effect sizes by dividing the observed mean difference in a study by the combined standard deviation of that study. The formula for Hedge's g is as follows:

$$\text{Hedge's g: } g = (M1 - M2) / SD_{pooled}$$

Here, M1 represents the mean of group 1, M2 represents the mean of group 2, and SD_{pooled} is the combined estimate of the population standard deviation (Borenstein et al., 2009).

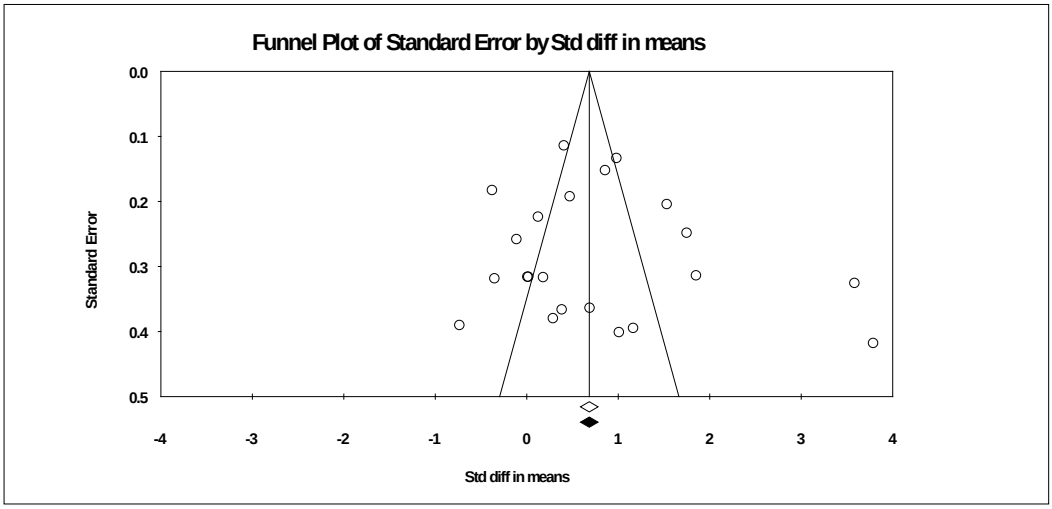
Traditionally, effect sizes are classified as large (0.80 or above), medium (0.50–0.80), small (0.20–0.50), or negligible (less than 0.20) based on Cohen's benchmark (1988). However, our study follows the scale proposed by Plonsky and Oswald (2014), where effect sizes are interpreted using the following benchmarks: 0.40 (small), 0.70 (medium), and 1.0 (large). Additionally, our study utilizes a random-effects model. This approach is common in language learning studies because effect sizes tend to vary among different studies (Borenstein et al., 2009). The data analysis was conducted using Comprehensive Meta-Analysis (Version 3).

Publication Bias

To investigate the possibility of publication bias, a funnel plot was generated. As illustrated in Figure 2, the funnel plot exhibited two key characteristics. First, the data points gathered closely around the 95% confidence interval line, and most of the circles tended to scatter symmetrically on both sides of the mean effect size. This distribution provides evidence indicating the absence of publication bias within the collection of primary studies. When the number of effect sizes on both sides of the average is roughly equal, it signals the inclusion of studies reporting both negative and positive effects. Publication bias would be a concern if the meta-analysis exclusively incorporated studies reporting positive effect sizes.

Figure 2

Funnel Plot of Publication Bias



Results

The analysis of these studies covered several categories. First, we delve into the overall impact of HVPT on pronunciation. Following that, we scrutinize factors or variables that could potentially influence the process of pronunciation learning when HVPT is used.

The Overall Effect of Integration of Computers on Pronunciation

The overall effect sizes of 18 studies with 22 effect sizes are shown in Table 1.

Table 1

Overall Effect of Integration Computer on Pronunciation

k*	Confidence intervals		P-value	Q-value	*Df*	I-squared	Hedge's g
	Lower limit	Upper limit					
22	0.405	1.134	.000	287.012	21	92.68	.77

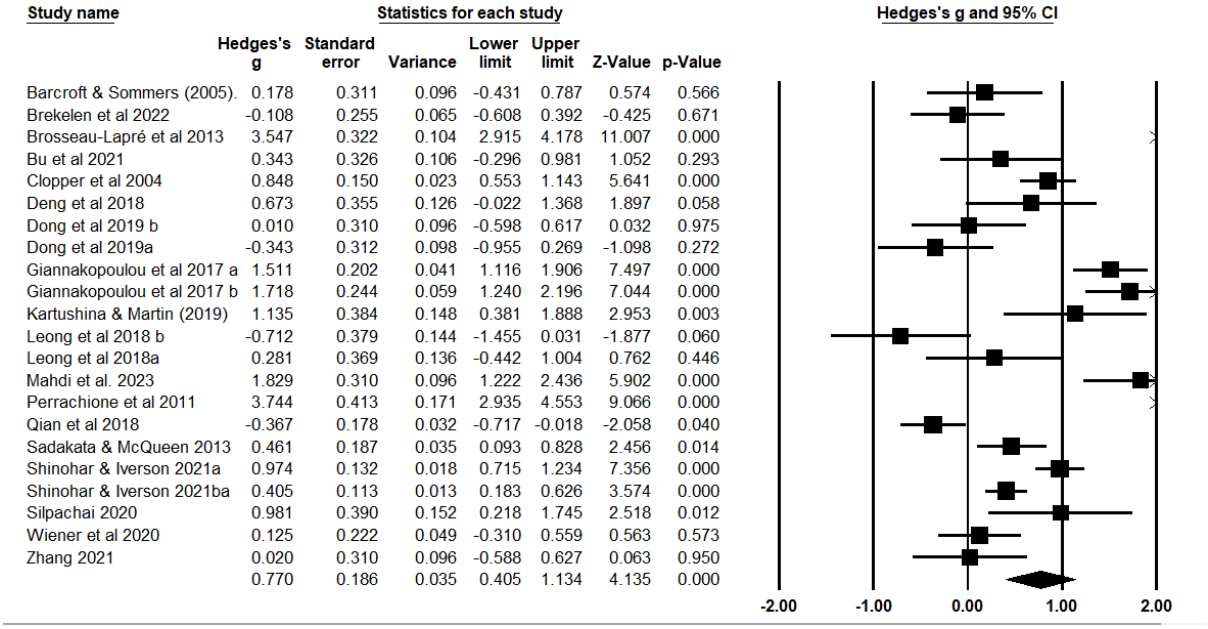
Note: k= number of effect sizes calculated. The level of significance is .05.

The examination employed the standardized mean difference as the measure of outcomes. A random-effects model was applied to the dataset. Table 1 presents the overall effect size of employing HVPT in L2 pronunciation, indicating a medium effect size (g = 0.77). Additionally, the Q-test for heterogeneity and the I^2 statistic are reported. The confidence intervals ranged from 0.405 to 1.134. Consequently, the average outcome significantly deviated from zero.

According to the Q-test, it appears that the true outcomes exhibit heterogeneity (Q= 287.012, p= 0.000). This implies that HVPT is an effective method for learning L2 pronunciation. The forest plot, depicted in Figure 3, offers a visual representation of the analysis for context.

Figure 3

Forest Plot of Overall Effect Size of HVPT on L2 Pronunciation



Additional information regarding the moderators that influence the application of HVPT in L2 pronunciation was explored, and the analysis of these studies was conducted based on various categories.

Table 2

Variables Analysis

Variable	Categories	k*	Hedge g	Confidence intervals		P-value	Q-value	I-squared
				Lower limit	Upper limit			
L 2 proficiency	Beginner	8	0.90	0.75	1.04	0.000	236.36	87.91
	Intermediate	5	0.28	0.03	0.53			
	Advanced	4	1.13	0.88	1.39			
	NA	5	0.28	0.11	0.46			
Outcome	Identification	13	0.65	0.52	0.78	0.000	286.95	92.68
	Production	9	0.68	0.54	0.82			
Target sounds	Vowels	5	0.53	0.38	0.68	0.000	189.54	93.74
	Consonant	8	1.06	0.89	1.23			
	Both	6	0.076	-0.11	0.27			
	Lexical	3	1.91	1.51	2.31			
Number of talkers	2-4	13	0.55	0.41	0.69	0.000	281.66	93.01
	5-8	9	0.77	0.64	0.91			
Number of stimuli	1-50	1	0.82	0.68	0.95	0.00	283.30	82.46
	51-100	2	-0.20	-0.72	0.31			
	101-200	4	0.27	-0.00	0.55			
	Above 200	3	0.66	0.49	0.82			
Context	Words	20	0.64	0.53	0.74	0.00	284.28	92.68
	Sentences	2	0.88	0.61	1.16			

**K=number of effect sizes*

The first variable initially examined was the L2 learners' proficiency. The learners' L2 proficiency was categorized into four groups: beginners, intermediate, advanced, and mixed. In instances where the primary study did not specify the learners' L2 proficiency, "not available" (NA) was used as a label. The results shown in Table 2 revealed that beginner learners who utilized HVPT demonstrated a medium effect size (g=0.90), while intermediate learners exhibited small effect sizes (g=0.28). Advanced learners had a large effect size (g=1.13). It is worth noting that confidence intervals were positive for all cases. Additionally, a statistical analysis using the Q-test revealed significant differences in effect sizes among these proficiency categories (Q=236.36, p = 0.00, I-squared=87.91).

The second variable examined was the outcome of the language treatments, which were categorized into two groups: identification and production. The results showed that studies measuring sound identification and sound production both had relatively small effect sizes (g=0.65 and g=0.68, respectively). Confidence intervals were positive in both cases, indicating consistent findings. Similar to proficiency levels, the Q-test demonstrated significant differences in effect sizes among these outcome categories (Q = 286.95, p = 0.00, I-squared= 92.68).

The third variable considered was the specific target sounds in the studies. The target sounds were divided into four categories: vowels, consonants, both vowels and consonants, and lexical sounds. The results revealed that studies targeting consonant sounds and lexical sounds had large effect sizes (g=1.06 and g=1.91, respectively). On the other hand, studies focusing on vowel sounds reported a small effect size (g=0.53), and those targeting both vowels and consonants showed non-significant results (g=0.07). Confidence intervals were positive for consonants, vowels, and lexical categories but included zero for the combined category (both consonants and vowels). Similar to previous analyses, the Q-test indicated significant

differences in effect sizes among these target sound categories (Q = 189.54, p = 0.00, I-squared= 93.74).

The fourth variable was the number of talkers involved in the studies. It was classified into two categories: from two to four talkers; and five to eight talkers. Results showed that studies involving two to four talkers had a small effect size (g=0.55), while those with more talkers (five to eight) had a medium effect size (g=0.77). Confidence intervals were positive in both cases. Once again, the Q-test demonstrated significant differences in effect sizes between these two categories (Q = 281.66, p = 0.00, I-squared= 93.01).

The fifth factor considered was the number of stimuli used in the studies. It was categorized into four groups: 1-50 stimuli, 51-100 stimuli, 101-200 stimuli, and above 200 stimuli. Studies employing fewer stimuli (1-100) reported a medium effect size (g=0.82). Those using stimuli from 100 to 200 had non-significant effect sizes (g=0.27), and studies using over 200 stimuli reported a small effect size (g=0.66). In contrast, studies employing stimuli in the 51-100 range reported a negative non-significant effect size (g=-0.20). With the exception of the 51-100 stimuli category, confidence intervals were positive. The Q-test again indicated significant differences in effect sizes among these stimulus categories (Q = 283.30, p = 0.00, I-squared= 82.46).

The sixth factor examined was the context in which the sounds were used. It was categorized into two groups: words and sentences. Studies used words as a context reported a small effect size (g=0.64), while those employing sentences had a medium effect size (g=0.88). Confidence intervals were positive in both cases, and the Q-test showed significant differences in effect sizes between these context categories (Q = 284.28, p = 0.00, I-squared= 92.68).

Discussion

The study aims at synthesizing the previous studies to explore the overall effect size of implementing HVPT in learning L2 pronunciation. Regarding the first research question, the results provided valuable insights into the impact of using HVPT on the pronunciation of learners in their second language (L2). It was found that, when considering the collective effect across all the examined factors and categories, the overall effect size of employing HVPT for improving L2 pronunciation was categorized as having a medium effect size. This means that the use of HVPT demonstrated a notable and discernible influence on the improvement of L2 pronunciation skills among the learners. The medium effect size suggests that, on average, there was a substantial and meaningful improvement in pronunciation as a result of employing this particular training method. It is indicative of the effectiveness of HVPT in enhancing the learners' ability to articulate and reproduce sounds accurately in their second language. The finding of this meta-analysis contradicts the finding of Zhang et al. (2021) who found that small effect size was reported.

Regarding the second research question, the results of this study present a multifaceted analysis of the variables influencing the effectiveness of HVPT on L2 pronunciation. Each variable examined provides valuable insights into the nuanced nature of L2 pronunciation improvement through HVPT. The findings related to learners' L2 proficiency levels reveal that the impact of HVPT varies based on learners' proficiency. HVPT can be beneficial for advanced and beginner L2 learners. Advanced learners demonstrated a large effect size. This finding indicates that HVPT can be effective for learners who are already proficient in their

L2. Advanced learners might be aiming for more nuanced and subtle improvements in their pronunciation, and HVPT appears to provide the necessary refinement to achieve this. The substantial effect size suggests that advanced learners can benefit significantly from HVPT, potentially closing the gap between near-native and native-like pronunciation. Similarly, the medium effect size observed for beginner learners indicates that HVPT has a substantial impact on improving pronunciation skills among those who are just starting to learn a new language. This finding aligns with the notion that beginners can benefit from the early exposure to diverse phonetic patterns and articulatory variations provided by HVPT, which can help establish a solid foundation for accurate pronunciation. On the other hand, intermediate learners exhibited a smaller effect size. This suggests that HVPT can be valuable for learners who have already developed some level of proficiency but may still struggle with certain phonetic aspects. For intermediate learners, HVPT could serve as a supplementary tool to refine their pronunciation skills further. This finding contradicts the conclusions of Chang and Bowles (2015) and Kingston (2003), who argued that variability hampers the learning process in novice learners. Conversely, this finding aligns with the findings of some studies (e.g., Antoniou & Wong, 2015; Perrachione et al., 2011; Sadakata & McQueen, 2014), which suggest that novice learners can still derive advantages from variability if they possess strong general perceptual skills, such as the ability to discern pitch contours. These results indicate that variability can impede the initial learning of L2 contrasts for novice learners with limited perceptual abilities, whereas L2 learners with heightened perceptual abilities stand to benefit from variability.

The study also examined the effectiveness of language treatments in two outcome categories: identification and production. Both categories had relatively small effect sizes, indicating that the treatments had a small impact on improving L2 pronunciation. This finding is in line with Zhang et al. (2021) who found that the effect size was small for identification, and it was much smaller for production. The study also analyzed the impact of language treatments on specific target sounds, including consonants, vowels, both vowels and consonants, and lexical tones. Consonant sounds showed a large positive effect, suggesting effective treatment. Lexical tones also had large effect size, indicating highly effective treatment. Vowel sounds had a small effect size while treatments targeting both vowels and consonants had a non-significant impact. This finding is in line with Zhang et al. (2021) and Mahdi et al. (2023) who found that the effect size was large for consonant sounds, and it was small for vowel sounds.

The finding that studies with more talkers (five to eight) yielded a medium effect size, while those with fewer talkers (two to four) resulted in a smaller effect size, highlights the significance of auditory diversity. When learners are exposed to a greater variety of speakers, they are likely to encounter a wider range of pronunciation patterns, accents, and speech styles. This exposure can be particularly beneficial for learners aiming to achieve a more adaptable and natural-sounding pronunciation. Educators and curriculum designers can draw valuable insights from this result. It suggests that incorporating a variety of audio resources, featuring multiple talkers, into pronunciation training materials can be advantageous. This approach can help learners develop the ability to understand and produce sounds in diverse contexts, contributing to more authentic and flexible pronunciation skills. This finding supports the idea that there is no difference in the results based on the number of talkers. However, the finding

of this study contradicts the findings of Davis (2015) and Shehata (2013), who reported large effect sizes for both categories.

The results indicate a clear trend: studies that employed a moderate quantity of stimuli (ranging from 1 to 100) reported a medium effect size in pronunciation improvement. This suggests that a sufficient quantity of practice material is crucial for achieving a meaningful and noticeable impact on L2 pronunciation. The medium effect size suggests that learners benefit from a balanced amount of exposure to different phonetic elements and contexts. Studies using an extensive number of stimuli (above 200) reported a smaller effect size. This result indicates a potential diminishing return on the effectiveness of HVPT when learners are immersed with an excess of practice material. It suggests that, beyond a certain point, additional stimuli may not significantly contribute to further pronunciation improvement. This finding could be related to cognitive load considerations, where an overwhelming number of stimuli may lead to reduced focus and retention. On the other hand, the non-significant effect sizes observed for studies with 100 to 200 stimuli and the negative non-significant effect size for studies with 51 to 100 stimuli raise questions about the challenges of using a limited set of stimuli. Limited stimuli may not provide learners with sufficient exposure to various phonetic variations, potentially hindering their ability to generalize and adapt their pronunciation skills to different contexts and speech patterns. These findings emphasize the importance of striking a balance between the quantity and quality of stimuli in HVPT programs. Providing an adequate number of stimuli is essential for exposure and practice, but it must be complemented with thoughtful selection and design of stimuli that capture relevant phonetic nuances and challenges. Moreover, the negative non-significant effect size observed for studies with 51 to 100 stimuli suggests that there may be a threshold below which the quantity of stimuli becomes insufficient for meaningful improvement.

The distinction between using words and sentences as contexts for HVPT interventions highlights the significance of contextual learning. The medium effect size for sentence-based training indicates that incorporating sounds into meaningful linguistic contexts can enhance pronunciation learning. However, the small effect size for word-based training suggests that further investigation is needed to determine the most effective methods for teaching individual sounds in isolation.

Pedagogical Implications

From a practical standpoint, the finding of the study implies that language educators should consider incorporating HVPT into their instruction. Language instructors can design lessons that integrate HVPT techniques, helping learners achieve more accurate and native-like pronunciation. Also, the positive correlation between a higher number of talkers the effect of HVPT emphasizes the importance of auditory diversity. Incorporating varied audio resources with multiple talkers in pronunciation training materials can enhance learners' adaptability and authenticity in pronunciation. Curriculum designers should consider including diverse audio stimuli to enrich learners' exposure. The study underscores the need for a balanced approach to the quantity and quality of stimuli in HVPT programs. While a moderate quantity (1 to 100 stimuli) yields a medium effect size, an excess of stimuli (above 200) may lead to diminishing returns. Educators should carefully select stimuli to ensure exposure to various phonetic elements while avoiding overwhelming learners. In addition, the distinction between word-

based and sentence-based training highlights the significance of contextual learning. While sentence-based training shows a medium effect size, further research is needed to determine the most effective methods for teaching individual sounds in isolation. Educators can explore the integration of meaningful linguistic contexts to enhance pronunciation learning.

Conclusion

This study aimed to provide a comprehensive assessment of the impact of HVPT on learning L2 pronunciation, considering various factors and categories. The results offer valuable insights into the overall effectiveness of HVPT in enhancing L2 pronunciation skills. Our findings indicate that, when considering all the factors and categories collectively, HVPT has a medium effect size on improving L2 pronunciation. This medium effect size indicates a significant and meaningful improvement in pronunciation as a result of employing HVPT.

Furthermore, the impact of HVPT varies based on learners' proficiency levels, with substantial effects observed for both advanced and beginner learners. Advanced learners benefit significantly from HVPT. For beginners, HVPT helps establish a solid foundation for accurate pronunciation. Intermediate learners also benefit, albeit to a lesser extent, suggesting HVPT can be a valuable supplementary tool for them. The importance of auditory diversity is evident, as studies with more talkers yield a medium effect size. This highlights the benefit of exposing learners to a wider range of pronunciation patterns and speech styles. Educators can draw insights from this, emphasizing the value of incorporating diverse audio resources into pronunciation training materials. Moreover, the quantity of stimuli plays a crucial role, with a balanced quantity resulting in a medium effect size. Excessive stimuli may not significantly contribute to further improvement, potentially due to cognitive load considerations, while limited stimuli may hinder generalization and adaptation of pronunciation skills. Contextual learning is also significant, with sentence-based training showing a medium effect size. However, further research is needed to determine the most effective methods for teaching individual sounds in isolation.

Limitations and Suggestions for Future Studies

While this study considered various factors, there may be additional variables that could influence the effectiveness of HVPT, such as individual learner characteristics, motivation, and the duration of training. Future studies should explore these variables to provide a more comprehensive understanding. Secondly, the study's findings may primarily apply to specific languages or language pairs. The effectiveness of HVPT could differ when applied to different language combinations. Future research should investigate a broader range of languages to assess its generalizability. Thirdly, the study may have focused on short-term effects. It's essential to investigate the long-term retention of improved pronunciation skills to assess the lasting impact of HVPT on L2 learners. Future studies may investigate the role of individual learner characteristics, such as age, aptitude, and language learning strategies, in moderating the impact of HVPT on pronunciation improvement. Understanding how these factors interact with HVPT can provide tailored recommendations for specific learner profiles. Additionally, future studies can employ longitudinal research methods to track pronunciation improvement over an extended period. This would help determine the sustainability of the gains achieved through HVPT and provide insights into the optimal frequency and duration of training.

Moreover, the effectiveness of HVPT can be compared across various language pairs to identify potential language-specific factors that influence its impact. This approach can assist educators in tailoring HVPT programs to address the unique challenges posed by different languages. The study also suggests that future research may examine how the effectiveness of HVPT may vary in different learning contexts, such as classroom settings, online courses, or self-study environments. Understanding the adaptability of HVPT to various educational settings is essential for its practical application.

ORCID

https://orcid.org/0000-0003-4023-8219

https://orcid.org/0000-0003-3169-102X

Acknowledgements

Not applicable.

Funding

The authors are thankful to the Deanship of Scientific Research at Najran University for funding this work under the Research Groups Funding program grant code (NU/RG/SEHRC/12/7).

Ethics Declarations

Competing Interests

No, there are no conflicting interests.

References

Antoniou, M., & Wong, P. C. (2015). Poor phonetic perceivers are affected by cognitive load when resolving talker variability. *The Journal of the Acoustical Society of America*, *138*(2), 571–574. https://doi.org/10.1121/1.4923362

*Barcroft, J., & Sommers, M. S. (2005). Effects of acoustic variability on second language vocabulary learning. *Studies in Second Language Acquisition*, *27*, 387–414. https://psycnet.apa.org/doi/10.1017/S0272263105050175

Berg, T. (1990). The differential sensitivity of consonants and vowels to stress. *Language Sciences*, *12*(1), 65-84. https://doi.org/10.1016/0388-0001(90)90024-B

Borenstein M., Hedges L. V., Higgins J. P. T. & Rothstein H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons.

Bradlow, A.R., Akahane-Yamada, R., Pisoni, D.B. & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics, 61*(5), 977–985. https://doi.org/10.3758/BF03206911

Bradlow, A. R., Bassard, A. M., & Paller, K. A. (2023). Generalized perceptual adaptation to second-language speech: Variability, similarity, and intelligibility. *The Journal of the Acoustical Society of America*, *154*(3), 1601-1613. https://doi.org/10.1121/10.0020914

*Brekelmans, G., Lavan, N., Saito, H., Clayards, M., & Wonnacott, E. (2022). Does high variability training improve the learning of non-native phoneme contrasts over low variability training? A replication. *Journal of Memory and Language*, *126*, 104352. https://doi.org/10.1016/j.jml.2022.104352

*Brosseau-Lapré, F., Rvachew, S., Clayards, M., & Dickson, D. (2013). Stimulus variability and perceptual learning of nonnative vowel categories. *Applied Psycholinguistics*, *34*(3), 419–441. https://doi.org/10.1017/S0142716411000750

*Bu, L., Nagano, M., Harel, D., & McAllister, T. (2021). Effects of practice variability on second-language speech production training. *Folia Phoniatrica et Logopaedica*, *73*(5), 384-400. https://doi.org/10.1159/000510621

Cebrian, J., & Carlet, A. (2014). Second-language learners' identification of target-language phonemes: A short-term phonetic training study. *Canadian Modern Language Review*, *70*(4), 474-499. https://doi.org/10.3138/cmlr.2318

Chang, C. B., & Bowles, A. R. (2015). Context effects on second-language learning of tonal contrasts. *The Journal of the Acoustical Society of America*, *138*(6), 3703-3716. https://doi.org/10.1121/1.4937612

*Clopper, C. G., & Pisoni, D. B. (2004). Effects of talker variability on perceptual learning of dialects. *Language and Speech*, *47*(3), 207-238. https://doi.org/10.1177%2F00238309040470030101

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Davis, A. K. (2015). *The interaction of language proficiency and talker variability in learning* (Unpublished Doctoral dissertation), The University of Arizona.

*Deng, Z., Chandrasekaran, B., Wang, S., & Wong, P. C. M. (2018). Training-induced brain activation and functional connectivity differentiate multi-talker and single-talker speech training. *Neurobiology of Learning & Memory*, *151*, 1–9. https://doi.org/10.1016/j.nlm.2018.03.009

*Dong H, Clayards M, Brown H, Wonnacott E. (2019). The effects of high versus low talker variability and individual aptitude on phonetic training of Mandarin lexical tones. *PeerJ*, *7*, e7191. https://doi.org/10.7717/peerj.7191

*Giannakopoulou, A., Brown, H., Clayards, M., & Wonnacott, E. (2017). High or low? Comparing high and low-variability phonetic training in adult and child second language learners. *PeerJ*, *5*, e3209. https://doi.org/10.7717%2Fpeerj.3209

Huensch, A., & Tremblay, A. (2015). Effects of perceptual phonetic training on the perception and production of second language syllable structure. *Journal of Phonetics*, *52*, 105-120. https://doi.org/10.1016/j.wocn.2015.06.007

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

Kachlicka, M., Saito, K., & Tierney, A. (2019). Successful second language learning is tied to robust domain-general auditory processing and stable neural representation of sound. *Brain and Language*, *192*, 15-24. https://doi.org/10.1016/j.bandl.2019.02.004

*Kartushina, N., & Martin, C. D. (2019). Talker and acoustic variability in learning to produce nonnative sounds: Evidence from articulatory training. *Language Learning*, *69*(1), 71-105. https://doi.org/10.1111/lang.12315

Kewley-Port, D., Burkle, T. Z., & Lee, J. H. (2007). Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *122*(4), 2365-2375. https://doi.org/10.1121/1.2773986

Kingston, J. (2003). Learning foreign vowels. *Language and Speech*, *46*(2-3), 295-348. https://doi.org/10.1177/00238309030460020201

Kleinschmidt, D. (2019). Structure in talker variability: How much is there and how much can it help? *Cognition and Neuroscience*, *34*(1), 43–68. https://doi.org/10.1080/23273798.2018.1500698

Lambacher, S. G., Martens, W. L., Kakehi, K., Marasinghe, C. A., & Molholt, G. (2005). The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics*, *26*(2), 227-247. https://doi.org/10.1017/S0142716405050150

*Leong, C. X. R., Price, J. M., Pitchford, N. J., & van Heuven, W. J. (2018). High variability phonetic training in adaptive adverse conditions is rapid, effective, and sustained. *PLOS One*, *13*(10), e0204888. https://doi.org/10.1371/journal.pone.0204888

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America*, *94*(3), 1242–1255. https://doi.org/10.1121/1.408177

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories. *The Journal of the Acoustical Society of America*, *96*(4), 2076–2087. https://doi.org/10.1121/1.410149

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English/r/and/l: A first report. *The Journal of the Acoustical Society of America*, *89*(2), 874-886. https://doi.org/10.1121/1.1894649

MacDonald, R. M. M. (2012). *Counteracting age related effects in L2 acquisition: Training to distinguish between French vowels* (Unpublished Doctoral dissertation), The University of Edinburgh.

Magen, H. S., & Blumstein, S. E. (1993). Effects of speaking rate on the vowel length distinction in Korean. *Journal of Phonetics*, *21*(4), 387-409. https://doi.org/10.1016/S0095-4470(19)30227-X

Mahdi, H. S., Alkhammash, R., & Al-Athwary, A. A. (2023). Using high variability phonetic training as a contextualized tool in the development of English consonant clusters pronunciation among Saudi EFL learners. *Education and Information Technologies*, 1-20. https://doi.org/10.1007/s10639-023-12113-9

Newman, R., Clouse, S., & Burnham, J. (2001). The perceptual consequences of within-talker variability in fricative production. *The Journal of the Acoustical Society of America*, *109*(3), 1181–1196. https://doi.org/10.1121/1.1348009

Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, *83*(1), 13-23. https://doi.org/10.1016/s0010-0277(01)00165-2

*Perrachione, T. K., Lee, J., Ha, L. Y., & Wong, P. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *The Journal of the Acoustical Society of America*, *130*(1), 461-472. https://doi.org/10.1121/1.3593366

Pisoni, D. B. (1993). Long-term memory in speech perception: Some new findings on talker variability, speaking rate and perceptual learning. *Speech Communication*, *13*(1–2), 109–125. https://doi.org/10.1016/01676393(93)90063-Q

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, *64*, 878–912. https://doi.org/10.1111/lang.12079

*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://dx.doi.org/10125/44582

*Sadakata, M., & McQueen, J. M. (2013). High stimulus variability in nonnative speech learning supports formation of abstract categories: Evidence from Japanese geminates. *The Journal of the Acoustical Society of America*, *134*(2), 1324–1335. https://doi.org/10.1121/1.4812767

Sadakata, M., & McQueen, J. M. (2014). Individual aptitude in Mandarin lexical tone perception predicts effectiveness of high-variability training. *Frontiers in Psychology*, *5*, 1318. https://doi.org/10.3389/fpsyg.2014.01318

Shehata, A. K. (2013). *When variability matters in second language word learning: Talker variability and task type effects*. (Unpublished doctoral dissertation), University of Utah.

*Shinohara, Y., & Iverson, P. (2021). The effect of age on English/r/-/l/perceptual training outcomes for Japanese speakers. *Journal of Phonetics*, *89*, 101108. https://doi.org/10.1016/j.wocn.2021.101108

*Silpachai, A. (2020). The role of talker variability in the perceptual learning of Mandarin tones by American English listeners. *Journal of Second Language Pronunciation*, *6*(2), 209–235. https://doi.org/10.1075/jslp.19010.sil

Strange, W., & Dittmann, S. (1984). Effects of discrimination training on the perception of/rl/by Japanese adults learning English. *Perception & psychophysics*, *36*(2), 131-145. https://doi.org/10.3758/BF03202673

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

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America*, *113*, 1033– 1043. https://doi.org/10.1121/1.1531176

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *The Journal of the Acoustical Society of America*, *106*(6), 3649-3658. https://doi.org/10.1121/1.428217

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

*Zhang, X., Cheng, B., & Zhang, Y. (2021). The role of talker variability in nonnative phonetic learning: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research*, *64*(12), 4802-4825. https://doi.org/10.1044/2021_jslhr-21-00181

Zhang, X., Cheng, B., Zou, Y., Li, X., & Zhang, Y. (2023). Cognitive factors in nonnative phonetic learning: Impacts of inhibitory control and working memory on the benefits and costs of talker variability. *Journal of Phonetics*, *100*, 101266. https://doi.org/10.1016/j.wocn.2023.101266