English Language Teaching Educational Journal
Vol. 9, No. 2, August 2026, pp. 260-273

ISSN 2621-6485

Effects of talker and vowel context in Thai learners' discrimination of English dental fricatives

Patchanok Kitikanan [a,b,1], Ho Ting Chak [a,2] *

[a] Naresuan University, Tha Pho, Mueang Phitsanulok District, Phitsanulok 65000, Thailand
[b] The Chinese University of Hong Kong University, Sha Tin, Hong Kong
[1] nicenanina@gmail.com , [2] htcchak@gmail.com

*Corresponding author

ARTICLE INFO

Article history
Received 29 Marc 2026
Revised 07 August 2026
Accepted 26 August 2026
Available Online 02 September 2026

Keywords
Dental fricative
L2 speech perception
Talker variability
Vowel context
World Englishes

ABSTRACT

L2 speech perception is shaped by variability in both talker characteristics and phonetic context, yet how these sources of variability interact remains insufficiently understood. This study investigated how talker accent and vowel context interact to influence L2 speech perception among Thai learners. Thirty-eight Thai undergraduate learners completed an ABX discrimination task testing five English consonant contrasts (/f–θ/, /ð–d/, /θ–ð/, /f–t/, and /t–θ/) produced by five talkers (Thai, Philippine, Indian, British, and American English) across three vocalic environments (high, low, and back vowels). Accuracy was analysed using a binomial generalised linear mixed-effects model (GLMM). Results revealed a significant three-way interaction among consonant contrast, vowel context, and talker ($p < .001$). Talker effects were limited in high- and back-vowel contexts but pronounced in the low-vowel context. Similarly, vowel-context effects were non-uniform; for instance, low vowels degraded /ð–d/ discrimination for American, Indian, and Thai talkers, whereas back vowels impaired /t–θ/ discrimination specifically for the Thai talker. Overall, no single talker or vowel context provided a uniform perceptual advantage. These findings demonstrate that L2 phoneme discrimination is dynamically co-conditioned by phonetic context and talker variability, carrying important implications for models of L2 speech perception (PAM-L2, SLM-r, L2LP) and contextualised perception pedagogy.

How to Cite: Kitikanan, P., & Chak, H. T. (2026). Effects of talker and vowel context in Thai learners' discrimination of English dental fricatives. *English Language Teaching Educational Journal, 9*(2), 260-273. https://doi.org/10.12928/eltej.v9i2.16045

1. Introduction

1.1. L2 Speech Perception across Diverse Talkers and Phonetic Contexts

The global spread of English has led to the emergence and increasing use of diverse localised varieties, commonly conceptualised within the framework of *World Englishes* (Rose et al., 2021). In contemporary communication, English is frequently used among speakers from different linguistic backgrounds, exposing second language (L2) learners to substantial variation in pronunciation. Such variability is relevant to L2 speech perception because the phonetic realization of consonantal contrasts may differ across individual talkers and English varieties. English learners increasingly encounter phonetic variability across speakers and English-variety backgrounds (Boonsuk et al., 2021; Boonsuk & Fang, 2022; Huang & Hashim, 2024). Nevertheless, L2 speech perception research has traditionally relied heavily on native-speaker models (Pickering & Huang, 2022), providing relatively limited

evidence on how learners perceive difficult phonological contrasts when listening to talkers from diverse English-variety backgrounds.

Several models of L2 speech perception provide a theoretical basis for understanding these difficulties. The Perceptual Assimilation Model for L2 (PAM-L2; Best et al., 2007) proposes that non-native sounds are perceived in relation to existing L1 categories and that discrimination depends on how L2 sounds are assimilated to those categories. From this perspective, contrasts involving dental fricatives and their potential Thai substitutes, particularly /ð–d/, /t–θ/, and /f–θ/, may present perceptual difficulty when the members of a contrast are assimilated to similar L1 categories. The Second Language Linguistic Perception model (L2LP; Van Leussen & Escudero, 2015) further emphasizes learners' use and reweighting of acoustic cues in establishing L2 sound categories. The revised Speech Learning Model (SLM-r; Flege & Bohn, 2021) similarly highlights the roles of perceived L1–L2 similarity, linguistic experience, and the quantity and quality of input in phonetic category development. Collectively, these models suggest that L2 discrimination reflects not only learners' existing phonological categories but also the acoustic characteristics and variability of the speech input.

One source of such variability is the individual talker. Talkers differ in the acoustic and phonetic realization of speech, reflecting variation in individual pronunciation patterns and broader linguistic experience (Bradlow et al., 2023). In a World Englishes context, this variability is particularly relevant because learners encounter speakers from diverse English-variety backgrounds. However, familiarity with a particular variety should not be equated with acoustic similarity. Thai learners, for example, may have greater experience with Thai-accented English, but such familiarity does not necessarily make every contrast produced by a Thai talker easier to discriminate. A talker's particular realization may increase or decrease the acoustic distinctiveness of a contrast (Kapadia et al., 2023). Accordingly, the present study focuses on variation across individual talkers rather than assuming uniform perceptual effects associated with national accent categories.

Phonetic context can significantly alter talker-related perception patterns. Research demonstrates that neighboring vowels affect consonant identification through coarticulatory and spectral transitions (Kitikanan, 2017; Kalaiah & Bhat, 2017; Liu & Li, 2024; Pan et al., 2026), meaning consonant contrasts may be perceived differently depending on the surrounding vowel. Accordingly, this study examines vowel context as a primary phonetic factor shaping L2 English speech perception.

Taken together, previous research leaves several issues insufficiently explored. Dental fricatives remain perceptually problematic for Thai learners, yet their discrimination has received limited attention across talkers from diverse English-variety backgrounds and phonetic contexts. Moreover, although PAM-L2, L2LP, and SLM-r highlight perceptual assimilation, acoustic cue use, experience, and input variability, relatively little research has considered these principles in listening conditions involving individual talker variability within a World Englishes context. Examining talker-related patterns across different phonetic environments may therefore provide further insight into the variability characterizing L2 perception of consonant contrasts.

The present study investigates how Thai learners discriminate five English consonant contrasts (/f–θ/, /ð–d/, /θ–ð/, /f–t/, and /t–θ/) produced by individual speakers representing Thai, Philippine, Indian, British, and American English. National designations are used to outline the speakers' general linguistic backgrounds rather than to imply uniform accent categories. Given that each variety is instantiated by a single speaker, the analysis centers on exploratory, talker-specific patterns rather than broad population-level generalisations. Ultimately, this work offers exploratory insights into how L2 English consonant discrimination fluctuates across distinct talkers and phonetic environments, framing these observations within wider models of L2 speech perception and World Englishes.

1.2. Theoretical frameworks of L2 speech perception and learning

L2 speech perception is broadly understood through three major theoretical frameworks: the Perceptual Assimilation Model of Second Language Speech Learning (PAM-L2; Best et al., 2007), the Second Language Linguistic Perception model (L2LP; Van Leussen & Escudero, 2015), and the revised Speech Learning Model (SLM-r; Flege & Bohn, 2021). PAM-L2 posits that non-native speech sounds are perceptually mapped onto pre-existing L1 phonological categories based on perceived phonetic similarity (Best et al., 2007). Discrimination accuracy depends on the resulting mapping type: contrasts mapped onto two distinct native categories (Two-Category assimilation) yield high discrimination accuracy, whereas contrasts assimilated to a single native category (Single-Category assimilation) present

severe perceptual difficulty. This model is applicable to English dental fricatives (/θ, ð/) for Thai learners, as these target sounds lack direct phonemic equivalents in Thai and are often perceptually assimilated to local stop (/t, d/) or fricative (/f/) categories. However, PAM-L2 primarily treats category assimilation as a static, segment-level mapping, offering limited mechanisms for how surrounding phonetic context or talker-specific acoustic variation might alter this perceptual alignment.

Conversely, L2LP and SLM-r adopt more dynamic, cue-based perspectives on L2 learning. L2LP emphasises that non-native listeners must re-weight acoustic dimensions (e.g., duration, spectral properties, and formant transitions) to match the distributional input of the target language (Escudero & Boersma, 2004; Van Leussen & Escudero, 2015). When different talkers or adjacent phonetic contexts alter these acoustic cues, listener accuracy fluctuates based on how flexible their cue-weighting strategies are. Similarly, SLM-r frames L2 speech learning as a lifespan process where new phonetic categories emerge based on the quality and quantity of input (Flege & Bohn, 2021). Within SLM-r, exposure to varied phonetic exemplars is critical; listeners must extract stable, contrastive properties across diverse acoustic realisations to construct robust L2 category representations.

Together, these frameworks offer complementary perspectives on why L2 sound discrimination may vary across contrasts and phonetic contexts (Baese-Berk et al., 2022). PAM-L2 emphasises how perceived relationships between L1 and L2 sound categories shape discrimination, whereas L2LP and SLM-r additionally emphasise the role of experience and sensitivity to phonetic detail in the development of L2 sound categories. Talker variability represents an important source of acoustic variation in L2 speech learning (Uchihara et al., 2022; Zhang et al., 2021), while successful phonetic learning may also generalise across novel talkers and phonetic contexts (Zhang et al., 2021). Despite these theoretical advancements, most empirical work has evaluated L2 discrimination using isolated target segments produced by standard inner-circle native talkers. Few studies have examined how complex L2 contrasts behave when talker variability (spanning multiple World Englishes varieties) interacts directly with coarticulatory vocalic contexts. Examining these interacting variables extends current L2 perception models beyond traditional native-speaker-centered paradigms, providing a more ecologically valid account of L2 perception in multilingual settings (Jenkins, 2000).

1.3. L2 perception of English dental fricatives

English dental fricatives /θ/ and /ð/ are widely recognised as difficult sounds for L2 learners, particularly when these phonemes are absent from the L1 inventory (Kafryawan & Zulihi, 2023; Krug, 2024; Öztürk, 2025). For Thai learners, English dental fricatives have been associated with perceptual and production difficulties involving phonologically similar Thai consonants, particularly /t/, /d/, and /f/ (Kitikanan, 2017).

Previous research has identified several potentially difficult contrasts involving dental fricatives, including /ð–d/, /f–θ/, and /t–θ/ (Kitikanan, 2016, 2017). The difficulty of these contrasts may reflect similarities in phonetic properties between English dental fricatives and available L1 categories. For example, /θ/ may be perceived in relation to /t/ because of shared voicelessness and similarities in place of articulation (Sridhanyarat, 2017), or to /f/ because both are voiceless fricatives with similar acoustic characteristics (Kitikanan, 2017). Similarly, /ð/ may be perceived in relation to /d/ (Sridhanyarat, 2017). From a PAM-L2 perspective, such cross-linguistic similarities are expected to make discrimination more difficult when members of an L2 contrast are perceptually related to similar L1 categories. These proposed mappings, however, represent theoretically motivated interpretations rather than assimilation patterns directly measured in the present study.

Identification and discrimination tasks have commonly been used to investigate learners' perception of non-native consonants (Cebrian et al., 2024). Perceptual difficulties may persist when the acoustic cues distinguishing unfamiliar consonants are subtle or variably realised. However, many previous studies have employed stimuli from a limited range of talkers, particularly speakers of standard native English varieties, potentially underrepresenting the talker variability encountered in global English communication (Jeong & Lindemann, 2025).

Research specifically examining Thai learners' perception of English dental fricatives remains relatively limited. Previous studies have frequently addressed production (e.g., Kitikanan, 2016; Sridhanyarat, 2017), with fewer studies investigating perceptual discrimination across different talkers and phonetic environments. Moreover, individual talkers may differ in their acoustic realization of the same phonological contrast. Examining discrimination across talkers from different English-variety

backgrounds may therefore provide insight into how robustly Thai learners perceive difficult consonant contrasts under variable speech input.

1.4. Effects of phonetic and sociophonetic variability on L2 perception

L2 speech perception is sensitive to variability in the speech input, including variability associated with talkers and phonetic contexts (Shejaeya et al., 2024). One important source of phonetic variability is vowel context, which can influence the perceptual salience of adjacent consonants through coarticulation (Mohan & Maruthy, 2021). Coarticulatory effects may modify spectral cues and formant transitions, thereby affecting the perception of consonantal features (Kalaiah & Bhat, 2017; Liu & Li, 2024). Previous studies (Mingxing, 2022; Liu & Li, 2024; Pan et al., 2026) have shown that vowel context can influence speech perception, with low-vowel contexts facilitating the perception of some consonant contrasts. For example, Mingxing (2022) used a speeded AX discrimination task with 24 listeners and found that discrimination accuracy was lower and response times were longer for [si–ɕi] than for [sa–ɕa], suggesting that the contrast was easier to discriminate in the low-vowel [a] context than in the high-vowel [i] context. Similarly, Liu and Li (2024) conducted two speeded AX discrimination experiments comparing [_a] and [_i] contexts. Their findings showed that the [n–l] contrast was perceptually more distinct in the [_a] context, particularly when it occurred with a high-initial tone. Together, these findings suggest that vowel context can affect consonant discrimination and that a low-vowel context may enhance perceptual distinctiveness.

Accent variability represents another important source of variation in L2 speech input (Alshangiti et al., 2023; Luthra, 2024; Zhang et al., 2021, 2025). Within the framework of World Englishes, learners encounter speakers from diverse linguistic backgrounds whose productions may differ phonetically (Schneider, 2011; Georgiou, 2024). Contemporary English language teaching in Thailand increasingly recognises diverse English accents and Global Englishes perspectives (Boonsuk et al., 2021; Ulla et al., 2025). Research indicates that non-native or accented speech negatively influences L2 speech perception (Miao, 2024; Pérez-Ramón, 2024). For instance, Miao (2024) tested 687 listeners from diverse native-language backgrounds who evaluated speakers with various accents, including native North American English as well as moderate and heavy Chinese accents. While accented speech generally had a negative impact on proficiency ratings, the study identified a specific threshold: speakers with moderate accents were evaluated just as favorably as native English speakers. Another study is from Pérez-Ramón (2024) with Japanese and Spanish listeners evaluating Spanish-accented English. It found that foreign accent variations are easier to process when words closely resemble native speech patterns. These studies lead to the hypothesis that because British and American English speakers are native, Thai learners will find their speech easier to differentiate than speech from Philippine, Indian, or Thai speakers. This study addresses the following research questions:

1. RQ1: How does discrimination accuracy vary across talkers and vowel contexts?
2. RQ2: Do the effects of talker and vowel context on discrimination accuracy vary across consonant contrasts?

2. Method

2.1. Stimuli

A total of 15 English lexical items were used as stimuli in this study. For consistency with the experimental design, the three stimulus sets are referred to as *high-vowel*, *low-vowel*, and *back-vowel conditions*. These labels describe the broad phonetic characteristics used in stimulus selection rather than mutually exclusive phonological vowel classes, as the individual lexical items contain different vowel phonemes. The study focused on five consonant contrasts: /f–θ/, /ð–d/, /θ–ð/, /f–t/, and /t–θ/. Each consonant contrast was embedded in three different words, with each word corresponding to one of the three vowel contexts. For the high vowel environment, the selected words were *"feel"*, "*thing"*, "*this"*, "*deep"*, and "*tea"*. The low vowel set included "*fat"*, "*thank"*, "*that"*, "*dad"*, and "*tag"*. The back vowel context consisted of "*food"*, *"thought"*, "*those"*, "*doom"*, and "*two"*. Because each consonant-by-vowel condition was represented by a limited set of lexical items, vowel-context effects cannot be fully disentangled from item-specific lexical and acoustic. The stimuli were produced by five male talkers with Thai, Philippine, Indian, British, and American English backgrounds, all affiliated with academic institutions. Because each English variety was represented by one talker, talker identity and variety

background were not independently manipulated. The variable is therefore referred to as talker condition, and comparisons represent differences among the five individual talkers rather than population-level differences among English varieties. Each talker was recorded separately in a sound-treated booth at Naresuan University using a Focusrite audio interface and TASCAM TM-280 condenser microphone at 48 kHz/24-bit resolution. The words were presented via PowerPoint, and talkers produced each item in the carrier sentence *"Say ____ again"* under the first researcher's supervision.

2.2. Participants

Thirty-eight Thai undergraduate students (six males), aged 18–20 years and majoring in fields other than English, participated in the study. None had received formal instruction in English phonetics. They had studied English for approximately 15 years through formal education in Thailand, and many had received grades of D or F in English courses. No independent standardised proficiency assessment was administered; therefore, course grades are reported only as background information and are not treated as a standardised measure of English proficiency. No formal audiometric screening was conducted.

2.3. Data collection

The experiment was implemented in Praat using the ExperimentMFC interface (Boersma & Weenink, 2023) with a tailored script. Prior to the main task, participants viewed a three-minute instructional video in Thai that explained the procedure. The video also included three sample trials to help participants become familiar with the task format. After the instruction phase, participants proceeded to the main experiment. In each trial, they listened to a sequence of three words and were required to decide whether the final word was more similar to the first or the second word by clicking the corresponding option on the screen. All auditory stimuli consisted of English words containing the target consonant, followed by the first 100 milliseconds of the subsequent vowel. For each consonant contrast, four presentation patterns were used. For instance, the /f–θ/ contrast (e.g., *"feel–thing"*) was presented in the following sequences: (1) "feel–thing–feel" (ABA), (2) "feel–thing–thing" (ABB), (3) "thing–feel–feel" (ABB), and (4) "thing–feel–thing" (ABA). To minimise predictability and repetition effects, the order of trials was randomised using the <PermuteBalancedNoDoublets> function in Praat. Each participant completed a total of 300 trials, derived from five consonant contrasts, three vowel contexts, four presentation patterns, and five speakers. The experiment took approximately 30 to 45 minutes to complete, and participants received 150 Baht as compensation. Ethical approval for the study was obtained from the Survey and Behavioural Research Ethics Faculty Sub-Committee of the Chinese University of Hong Kong, Faculty of Education (EDU2026-027).

2.4. Data analysis

Data were transferred from Microsoft Excel. Descriptive statistics were presented with frequency. Regarding inferential statistics, the analyses were conducted in *R* using *RStudio* (RStudio Team, 2026). Discrimination accuracy was coded as a binary outcome (1 = correct, 0 = incorrect) and analysed using a generalised linear mixed-effects model (GLMM) with a binomial distribution and logit link function. The model was fitted using the *glmmTMB* package (McGillycuddy et al., 2025). Consonant contrast, vowel context, talker condition, and all two- and three-way interactions among these variables were entered as fixed effects. Participant was included as a random intercept to account for repeated observations within participants. Thus, the fixed-effects structure tested the effects of consonant contrast × vowel context × talker condition simultaneously.

A participant random-slope structure for vowel context was initially examined. However, this model resulted in singular convergence and a non-positive-definite Hessian matrix, indicating that the more complex random-effects structure was not adequately supported by the data. The random-effects structure was therefore simplified to a participant random intercept while retaining the full fixed-effects structure.

The statistical significance of the fixed effects and their interactions was evaluated using Type III Wald chi-square tests. Because the model included a three-way interaction, significant interaction effects were followed up using estimated marginal means with the *emmeans* package (Lenth, 2024). To examine talker-related differences, pairwise comparisons among talker conditions were conducted separately within each consonant contrast × vowel-context combination. To directly evaluate vowel-context effects, pairwise comparisons among vowel contexts were conducted within each consonant contrast × talker combination. Pairwise comparisons were adjusted for multiple testing using the Bonferroni procedure.

For pairwise comparisons, effects were expressed as odds ratios (ORs) with 95% confidence intervals (CIs). An odds ratio greater than 1 indicated higher odds of a correct discrimination response for the first

condition relative to the second, whereas an odds ratio below 1 indicated lower odds. Statistical significance was evaluated at α = .05. Model convergence, residual behaviour, and overdispersion were examined as part of model diagnostics.

3. Findings and Discussion

3.1. Descriptive Result

Descriptive analyses examined discrimination accuracy across the five talkers within each vowel-context condition and consonant contrast. For /ð–d/, the highest accuracy was observed for the American talker in the back-vowel condition, the Indian talker in the high-vowel condition, and the British talker in the low-vowel condition. The Thai talker showed the lowest accuracy in both the high- and low-vowel conditions.

For /θ–ð/, the Thai talker showed the highest accuracy in the back-vowel condition, whereas the British and Indian talkers showed the highest accuracy in the high-vowel condition and the British talker in the low-vowel condition. For /f–t/, the Thai talker showed the highest accuracy in the back-vowel condition but the lowest accuracy in the high- and low-vowel conditions; the highest accuracy in these latter conditions was observed for the American and Philippine talkers, respectively.

For /f–θ/, the highest accuracy was observed for the American and Thai talkers in the back-vowel condition, the Thai talker in the high-vowel condition, and the Indian talker in the low-vowel condition. For /t–θ/, the American talker consistently showed the highest accuracy across the three vowel-context conditions, whereas the Thai talker showed the lowest. Overall, no single talker consistently yielded the highest discrimination accuracy across contrasts and vowel-context conditions, although a more consistent descriptive pattern was observed for /t–θ/. Table 1 presents the frequency of correct responses across consonant contrasts, vowel-context conditions, and talkers.

Table 1. Frequency of correct responses across contrasts, vowel contexts, and talkers (total = 152)

	American	British	Indian	Philippine	Thai	American
/ð–d/	Back	141	139	130	140	139
	High	134	135	138	133	126
	Low	127	135	121	131	103
/θ–ð/	Back	138	140	138	135	143
	High	136	137	137	135	132
	Low	126	143	133	142	135
/f-t/	Back	139	139	139	137	142
	High	143	141	141	142	140
	Low	138	140	139	143	135
/f–θ/	Back	137	134	131	135	137
	High	133	124	139	138	143
	Low	129	123	141	116	140
/t–θ/	Back	141	139	134	137	117
	High	142	141	135	141	135
	Low	143	138	137	138	134

Note. The orange boxes indicate the highest correct discrimination scores, whereas the yellow boxes indicate the lowest correct discrimination scores.

Overall, the descriptive results indicate that discrimination accuracy varied across talkers, vowel-context conditions, and consonant contrasts. No single talker consistently facilitated or hindered discrimination across all conditions.

3.2. Results from GLMM

A binomial generalised linear mixed-effects model was fitted to discrimination accuracy, with consonant contrast, vowel context, talker condition, and their interactions as fixed effects and participant as a random intercept. Type III Wald chi-square tests showed a significant main effect of vowel context, $\chi^2(2) = 8.20$, $p = .017$, whereas the main effects of consonant contrast, $\chi^2(4) = 1.39$, $p = .846$, and talker, $\chi^2(4) = 7.80$, $p = .099$, were not significant. Significant two-way interactions were found between

consonant contrast and talker, $\chi^2(16) = 36.96$, $p = .002$, and between vowel context and talker, $\chi^2(8) = 23.30$, $p = .003$. The contrast × vowel interaction was not significant, $\chi^2(8) = 8.57$, $p = .380$.

Importantly, there was a significant three-way interaction between consonant contrast, vowel context, and talker, $\chi^2(32) = 69.79$, $p < .001$. This interaction indicates that talker-related differences in discrimination accuracy varied according to both consonant contrast and vowel context. Given this interaction, follow-up comparisons were conducted to examine talker differences within each contrast × vowel combination and vowel-context differences within each contrast × talker combination. Pairwise comparisons were Bonferroni-adjusted.

a) Talker effects across vowel contexts and consonant contrasts

Talker differences were relatively limited in the back-vowel context, except for the /θ–t/ contrast. For this contrast, the odds of a correct response were significantly higher for the American (OR = 6.56, 95% CI [2.00, 21.55], $p < .001$), British (OR = 5.22, 95% CI [1.66, 16.35], $p < .001$), Indian (OR = 3.23, 95% CI [1.11, 9.36], $p = .020$), and Philippine talkers (OR = 4.25, 95% CI [1.41, 12.84], $p = .002$) than for the Thai talker. In other words, Thai listeners discriminated this contrast more accurately when it was produced by the American, British, Indian, and Philippine talkers than when it was produced by the Thai talker.

Talker differences were also limited in the high-vowel context. Significant differences occurred only for /θ–f/. The British talker yielded significantly lower odds of correct discrimination than the Indian talker (OR = 0.29, 95% CI [0.09, 0.95], $p = .033$) and the Thai talker (OR = 0.18, 95% CI [0.05, 0.65], $p = .002$). This suggests that for /θ–f/ in the high-vowel context, Thai learners discriminated the contrast less accurately when it was produced by the British talker than when it was produced by either the Indian or Thai talker.

More extensive talker differences occurred in the low-vowel context. For /ð–d/, the Thai talker yielded significantly lower odds of correct discrimination than the American (OR = 3.86, 95% CI [1.47, 10.12], $p < .001$), British (OR = 7.25, 95% CI [2.56, 20.58], $p < .001$), Indian (OR = 2.61, 95% CI [1.03, 6.58], $p = .037$), and Philippine talkers (OR = 5.19, 95% CI [1.92, 14.07], $p < .001$). The results imply that for /ð–d/ in the low-vowel context, Thai learners discriminated the contrast significantly less accurately when it was produced by the Thai talker than when it was produced by the American, British, Indian, or Philippine talker. For /ð–θ/, the American talker yielded significantly lower odds than the British (OR = 0.21, 95% CI [0.06, 0.75], $p = .006$) and Philippine talkers (OR = 0.24, 95% CI [0.07, 0.83], $p = .012$). In other words, for /ð–θ/ in the low-vowel context, Thai learners showed significantly lower discrimination accuracy for the American talker than for the British and Philippine talkers.

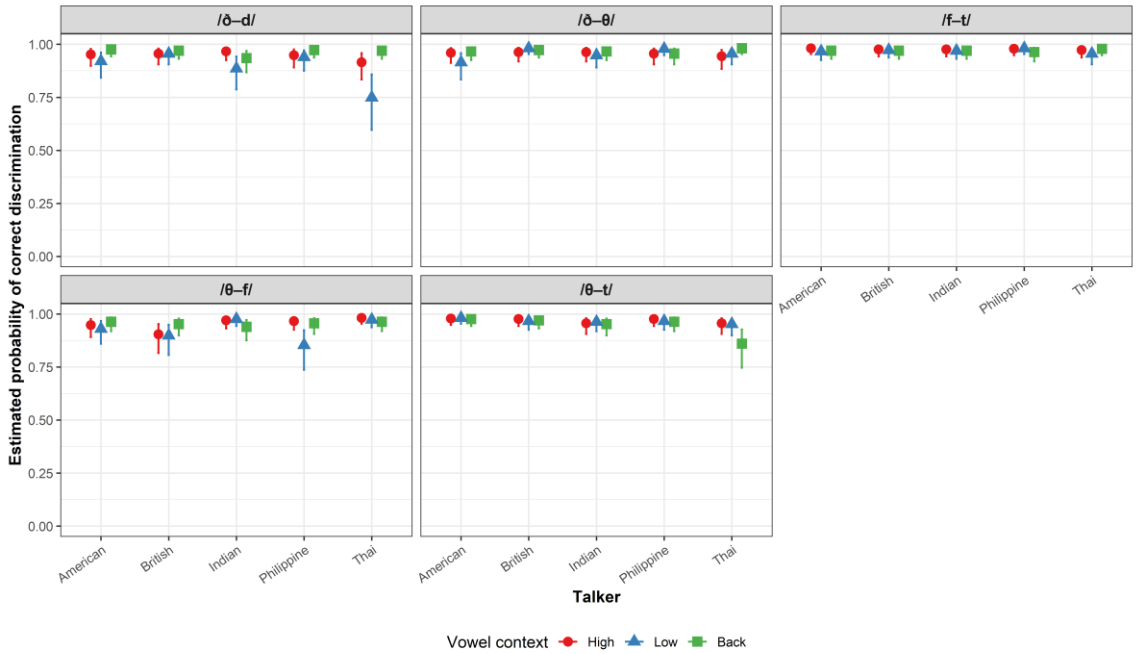

Fig 1. Model-estimated probabilities of correct discrimination across talkers, by vowel context and consonant contrast. Error bars represent 95% confidence intervals.

For /θ–f/ in the low-vowel context, several additional differences were observed. The British talker yielded lower odds than both the Indian (OR = 0.22, 95% CI [0.07, 0.74], p = .004) and Thai talkers (OR = 0.25, 95% CI [0.08, 0.81], p = .009). The Indian talker yielded higher odds than the Philippine talker (OR = 6.94, 95% CI [2.12, 22.74], p < .001), whereas the Philippine talker yielded lower odds than the Thai talker (OR = 0.16, 95% CI [0.05, 0.52], p < .001). This suggests that for /θ–f/ in the low-vowel context, Thai learners showed significantly better discrimination for the Indian and Thai talkers than for the British and Philippine talkers. No significant talker differences were found for /f–t/ in any of the three vowel contexts. Figure 1 shows model-estimated probabilities of correct discrimination across talkers, by vowel context and consonant contrast.

b) Vowel-context effects across talkers and consonant contrasts

Follow-up comparisons also showed that vowel-context effects were not uniform across contrasts and talkers. For the American talker, a significant vowel effect occurred for /ð–d/, with higher odds of correct discrimination in the back than in the low-vowel context (OR = 3.49, 95% CI [1.22, 9.98], p = .013). For the Indian talker, /ð–d/ discrimination was significantly better in the high than in the low-vowel context (OR = 3.71, 95% CI [1.41, 9.78], p = .004). These results suggest that for /ð–d/, discrimination was poorer in the low-vowel context for both talkers, whereas better performance occurred in the back-vowel context for the American talker and the high-vowel context for the Indian talker.

For the Philippine talker, significant vowel-context differences occurred for /θ–f/. The odds of correct discrimination were higher in both the back (OR = 3.73, 95% CI [1.49, 9.32], p = .002) and high-vowel contexts (OR = 4.97, 95% CI [1.91, 12.92], p < .001) than in the low-vowel context. These imply that when /θ–f/ was produced by the Philippine talker, Thai learners discriminated the contrast significantly more accurately in both the back- and high-vowel contexts than in the low-vowel context.

The Thai talker showed the most extensive vowel-context differences. For /ð–d/, performance was higher in the back than in the high-vowel context (OR = 2.98, 95% CI [1.09, 8.15], p = .028) and substantially higher in the back than in the low-vowel context (OR = 10.74, 95% CI [4.17, 27.68], p < .001). Performance was also higher in the high than in the low-vowel context (OR = 3.61, 95% CI [1.60, 8.15], p < .001). Overall, for /ð–d/ in Thai speech, discrimination accuracy was highest in the back-vowel context, intermediate in the high-vowel context, and lowest in the low-vowel context. For /ð–θ/, performance was higher in the back than in the high-vowel context (OR = 3.12, 95% CI [1.01, 9.66], p = .048). This means that Thai learners showed significantly higher discrimination accuracy in Thai speech in the back-vowel context than in the high-vowel context.

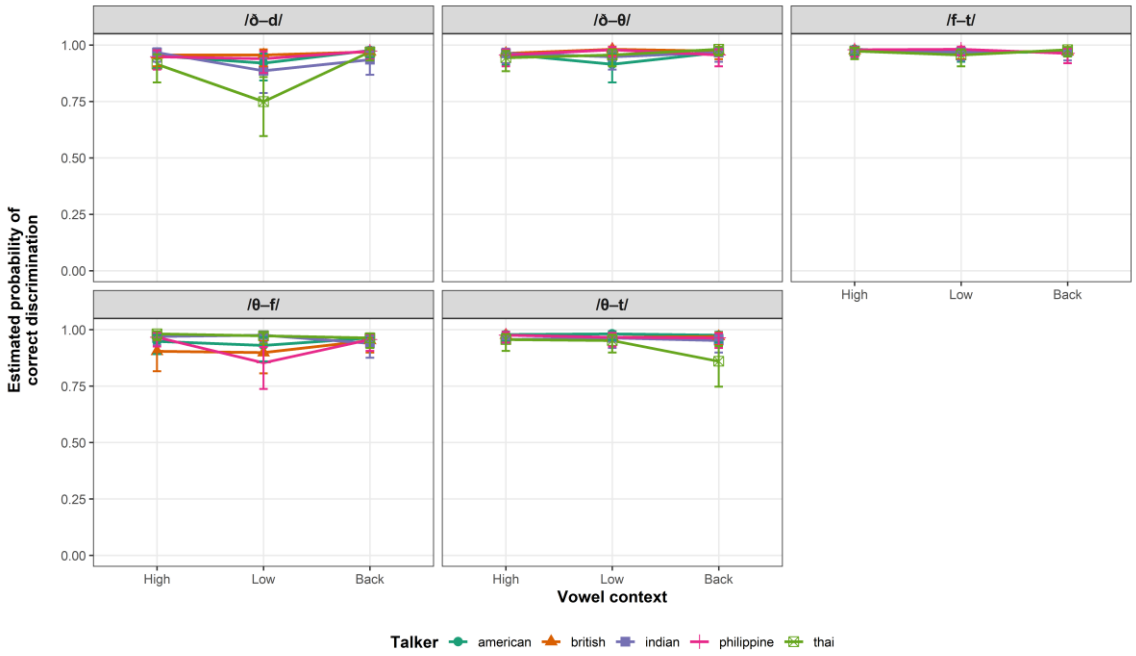

Fig 1. Model-estimated probabilities of correct discrimination across vowel contexts, by talker and consonant contrast. Error bars represent 95% confidence intervals.

A different pattern emerged for the Thai talker on /θ–t/. Here, the odds of correct discrimination were lower in the back-vowel context than in both the high-vowel context (OR = 0.28, 95% CI [0.11, 0.71], p = .003) and the low-vowel context (OR = 0.31, 95% CI [0.13, 0.77], p = .006). Thus, the back-vowel context was associated with poorer discrimination of /θ–t/ relative to the other two vowel contexts. Figure 2 presents model-estimated probabilities of correct discrimination across vowel contexts, by talker and consonant contrast.

Overall, discrimination accuracy was jointly shaped by talker, vowel context, and consonant contrast rather than by any one factor independently. Neither talker nor vowel context exerted a uniform effect; instead, their effects were contrast-specific and mutually dependent.

3.3. Discussion

The present study examined talker-related and phonetic context variation in Thai learners' discrimination of five English consonant contrasts (/f–θ/, /ð–d/, /θ–ð/, /f–t/, and /θ–t/), focusing particularly on dental fricatives across three vowel environments. The statistical results (GLMM) revealed that discrimination accuracy is driven by a significant three-way interaction between consonant contrast, vowel context, and talker condition. No single talker or phonetic context yielded a uniform perceptual advantage or disadvantage across all conditions.

The hypothesis that a low-vowel context would enhance perceptual distinctiveness received only partial support. For /t–θ/, Thai listeners showed higher odds of correct discrimination in both the low- and high-vowel contexts than in the back-vowel context, indicating that the low vowel facilitated discrimination relative to the back vowel. For the remaining contrasts, however, the low-vowel context either did not differ from the other vowel contexts or was associated with lower discrimination accuracy. One explanation is that the perceptual effect of vowel context is not determined by vowel height alone, but the following vowel modifies the acoustic relationship between the consonants being compared. Vowel context can influence fricative spectra and consonant-to-vowel formant transitions, but the usefulness of these cues varies across consonantal contrasts (Yeni-Komshian & Soli, 1981). Moreover, previous research with Thai learners has shown that the perceptual assimilation of English /θ/ itself varies across vowel contexts: /θ/ was more frequently assimilated to Thai /f/ in high- and low-vowel contexts but to /s/ in a back-vowel context (Kitikanan, 2017). Thus, a low vowel may increase the perceptual distance between some consonants, as appears to have occurred for /t–θ/, while reducing or leaving unchanged the perceptual distance between others. The present findings therefore suggest that vowel-context effects arise from an interaction between coarticulatory acoustic information and contrast-specific L1–L2 perceptual mappings rather than from a general advantage associated with low vowels.

The hypothesis that Thai learners would find speech produced by native talkers easier to discriminate than speech produced by non-native talkers also received only limited support. Higher odds of correct discrimination for native speakers than non-native talkers were observed in only a small number of conditions. For /t–θ/ in the back-vowel context, discrimination was more accurate for the American and British talkers than for the Thai talker, while for /ð–d/ in the low-vowel context, the American and British talkers also yielded higher discrimination accuracy than the Thai talker. In most other conditions, however, no consistent advantage for the native talkers over non-native talkers was observed. This suggests that talker nativeness alone did not determine perceptual distinctiveness for the Thai listeners. Exposure to native-speaker productions does not necessarily guarantee accurate perception among L2 learners (Borodkin, 2022). One possible explanation is that L2 listeners do not necessarily benefit from the phonetic characteristics of native speech because their perception is shaped by L1-based cue weighting and their experience with L2 speech (Zhang et al., 2024). More importantly, the present results indicate that talker effects depend on the particular consonant contrast and vowel context. Certain talkers may have realised contrastive acoustic cues more distinctly in particular phonetic environments, thereby facilitating discrimination.

The complex three-way interaction observed among consonant contrast, vowel context, and talker accent provides crucial empirical nuance to current models of L2 speech perception. Under the frameworks of PAM-L2 (Best & Tyler, 2007) and SLM-r (Flege & Bohn, 2021), foreign contrast discrimination depends heavily on how L2 phones map onto L1 phonetic categories, as well as the precision of acoustic cue detection. However, our findings demonstrate that category assimilation and phonetic distance are not fixed properties of isolated target phonemes. Instead, as predicted by dynamic exemplar-based approaches embedded in L2LP (Van Leussen & Escudero, 2015), perceptual mapping is highly contextualised: adjacent vocalic environments dynamically alter the available acoustic cues

(e.g., formant transitions, duration, and spectral burst properties). The significant degradation or facilitation of specific contrasts in particular vowel contexts—such as the severe reduction in /ð–d/ discrimination in low-vowel contexts—highlights that L2 phonetic category formation operates on larger contextual units (such as CV syllables) rather than context-free segmental representations.

Furthermore, the finding that talker-specific accent variability interacts non-uniformly with vowel context challenges static implementations of L2 perceptual models. While PAM-L2 accounts for L1-to-L2 category mapping and SLM-r focuses on how experience updates phonetic representations over time, neither model explicitly formalises how non-native listeners process talker variability across varying vocalic environments. In line with L2LP's emphasis on acoustic input distributions and optimal perception, our results show that listeners do not apply a single uniform cue-weighting strategy across talkers. Rather, talker accent alters the acoustic realisation of coarticulated sequences, shifting whether a given vowel context aids or hinders contrast discrimination (e.g., Thai speech yielding the highest /θ–ð/ discrimination in the back vowel context, but the lowest in the high vowel context). Integrating context-dependent talker variability into SLM-r and PAM-L2 is therefore essential for modeling how L2 learners construct robust, generalised phonetic categories in real-world communicative settings.

From a pedagogical perspective, these findings highlight the necessity of incorporating context-dependent variability into L2 pronunciation and perception training as suggested by the study of Uchihara, Karas and Thomson (2025). Traditional high-variability phonetic training (HVPT) often prioritises exposure to multiple talkers, but our results demonstrate that talker accent alone is insufficient (Brekelmans et al., 2022); training must also systematically account for coarticulatory vowel environments. Because performance on specific target contrasts—such as /ð–d/ or /θ–f/—varied dramatically depending on whether they were flanked by high, low, or back vowels, instructors and material developers should move away from presenting L2 target sounds in isolated or static contexts. Instead, perceptual training regimes should adopt a multi-dimensional framework that pairs target consonant contrasts across a diverse range of vocalic environments and talker accents (both native and non-native L2 talkers), such as in videogame (Saito et al., 2022). By exposing learners to these coarticulatory acoustic variations, pedagogical interventions can promote more adaptive cue-weighting strategies among L2 learners (Zhang et al., 2023), thereby facilitating the development of more flexible and robust phonetic categories that better reflect the variability of real-world communicative settings.

4. Conclusion

In conclusion, this study demonstrates that L2 speech perception is dynamically shaped by the interaction of speaker background, phonetic context, and specific consonant contrasts, rather than by any single factor in isolation. The significant three-way interaction reveals that talker and vowel effects are highly interdependent and contrast-specific. While native and non-native speaker differences varied considerably depending on the surrounding vowel, low-vowel contexts generally introduced greater perceptual difficulty. Ultimately, these findings highlight that L2 speech discrimination is non-uniform, underscoring the critical need to account for both talker variability and immediate phonetic environment when evaluating non-native phonological processing.

Some limitations should be acknowledged. First, each English-variety background was represented by only one talker; therefore, talker-specific characteristics cannot be separated from broader variety-related variation, and the findings should not be generalised to entire English varieties. Second, the limited lexical sampling means that vowel-context patterns cannot be fully disentangled from item-specific lexical and acoustic characteristics. Future research should include multiple talkers and lexical items per condition, direct acoustic measures, and more diverse learner groups (Kostromitina et al., 2025). Third, participants' prior exposure to different varieties of English and their English proficiency were not systematically assessed, which may have contributed to individual differences in discrimination performance. Although participants reported no known hearing difficulties, formal audiometric screening was not conducted.

Acknowledgment

The authors express their profound gratitude to all individuals who played a role in the research undertaking, including those involved in the collection and examination of data, as well as those who provided constructive feedback.

Declarations

Author contribution	:	The authors are responsible for the entire research project, from initiating the ideas to revising the manuscript.
Funding statement	:	This research did not receive any funding.
Conflict of interest	:	There is no conflict of interest in this work.
Ethical declaration	:	We, as authors acknowledge that this work has been written based on ethical research that conforms with the regulations of my university and that we have obtained permission from the relevant institute when collecting data. We support ELTEJ in maintaining high standards of personal conduct, practicing honesty in all our professional practices and endeavors.
Additional information	:	No additional information is available for this paper.

REFERENCES

Alshangiti, W., Evans, B. G., & Wibrow, M. (2023). Investigating the effects of speaker variability on Arabic children's acquisition of English vowels. *Arab World English Journal, 14*(1), 3–27.

Baese-Berk, M. M., Chandrasekaran, B., & Roark, C. L. (2022). The nature of non-native speech sound representations. *The Journal of the Acoustical Society of America, 152*(5), 3025–3034. https://doi.org/10.1121/10.0015230

Best, C. T., Tyler, M., Bohn, O., & Munro, M. (2007). Nonnative and second-language speech perception. In M. J. Munro & O. S. Bohn (Eds.), *Language experience in second language speech learning* (Vol. 17, pp. 13-34). John Benjamins.

Boersma, P., & Weenink, D. (2023). *Praat: Doing phonetics by computer (Version 6.4.60)*. In *Glot International* http://www.praat.org/

Boonsuk, Y., Ambele, E. A., & McKinley, J. (2021). Developing awareness of Global Englishes: Moving away from 'native standards' for Thai university ELT. *System*, *99*, 102511. https://doi.org/10.1016/j.system.2021.102511 02511

Boonsuk, Y., & Fang, F. (2022). Perennial language learners or competent language users: An investigation of international students' attitudes towards their own and native English accents. *RELC Journal*, *53*(1), 40–55. https://doi.org/10.1177/0033688220926715 6715.

Borodkin, K., Gassner, T., Ershaid, H., & Amir, N. (2022). tDCS modulates speech perception and production in second language learners. *Scientific Reports*, *12*(1), 16212. https://doi.org/10.1038/s41598-022-20512-0

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

Cebrian, J., Gavaldà, N., Gorba, C., & Carlet, A. (2024). Differential effects of identification and discrimination training tasks on L2 vowel identification and discrimination. *Studies in second language acquisition*, *46*(4), 1069-1093. https://doi.org/10.1017/S0272263124000408

Choi, W. (2022). Theorizing positive transfer in cross-linguistic speech perception: The Acoustic-Attentional-Contextual hypothesis. *Journal of Phonetics*, *91*, 101135. https://doi.org/10.1016/j.wocn.2022.101135

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in second language acquisition*, *26*(4), 551-585. https://doi.org/10.1017/S0272263104040021

Flege, J. E., & Bohn, O.-S. (2021). The Revised Speech Learning Model (SLM-r) Applied. In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 84-118). Cambridge University Press.

Georgiou, G. P. (2024). Perception of familiar second language accents and the role of linguistic background. *Applied Sciences*, *14*(24), 11776. https://doi.org/10.3390/app142411776

Huang, Y., & Hashim, A. (2024). Contributing factors to Chinese English learners' perceptions of different English accents. *RELC Journal*, *55*(2), 313-328. https://doi.org/10.1177/00336882221090321

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford university press.

Jeong, H., & Lindemann, S. (2025). Beyond ideologies of nativeness in the intelligibility principle for L2 English pronunciation: A corpus-supported review. *System*, *129*, 103599. https://doi.org/10.1016/j.system.2025.103599

Kafryawan, W., & Zulihi, Z. (2023). Sound production of English dental fricatives by Papuan EFL students. *Tell: Teaching of English Language and Literature Journal, 11*(2). https://doi.org/10.30651/tell.v11i2.19007

Kalaiah, M. K., & Bhat, J. S. (2017). Effect of vowel context on the recognition of initial consonants in Kannada. *Journal of audiology & otology*, *21*(3), 146. https://doi.org/10.7874/jao.2017.00122

Kapadia, A. M., Tin, J. A., & Perrachione, T. K. (2023). Multiple sources of acoustic variation affect speech processing efficiency. *The Journal of the Acoustical Society of America*, *153*(1), 209-223. https://doi.org/10.1121/10.0016611

Kitikanan, P. (2016). *L2 English fricative production by Thai learners* [Doctoral dissertation, Newcastle University].

Kitikanan, P. (2017). The effects of L2 experience and vowel context on the perceptual assimilation of English fricatives by L2 Thai learners. *English Language Teaching*, *10*(12), 72-82. https://doi.org/10.5539/elt.v10n12p72

Kostromitina, M., Sudina, E., & Baghlaf, E. (2025). Study and instrument quality in perception-based L2 pronunciation research: A methodological synthesis. *Studies in Second Language Acquisition, 47*(2), 721–754. https://doi.org/10.1017/S027226312500018X

Krug, M. (2024). English dental fricatives and their substitutes in EFL learners: A study of L1-German university students and implications for EFL teaching. In A. Rosen & K. Beuter (Eds.), *Englische Sprachwissenschaft und Fachdidaktik im Dialog: Chancen zur Stärkung der Lehrkräftebildung* (pp. 65–80). Narr Francke Attempto Verlag. https://doi.org/10.24053/9783381112524

Lenth, R. (2024). emmeans: Estimated Marginal Means, aka Least-Squares Means. *R package version 1.10. 5*. https://CRAN.R-project.org/package=emmeans

Liu, P. B., & Li, M. (2024). The perceptual distinctiveness of the [n-l] contrast in different vowel and tonal contexts. *JASA Express Letters*, *4*(11). https://doi.org/10.1121/10.0034196

Luthra, S. (2024). Why are listeners hindered by talker variability? *Psychon Bull Rev, 31*, 104–121. https://doi.org/10.3758/s13423-023-02355-6

McGillycuddy M, Warton DI, Popovic G, Bolker BM (2025). Parsimoniously fitting large multivariate random effects in glmmTMB. *Journal of Statistical Software*, 112(1), 1–19.

Miao, Y. (2024). Factors affecting listener perception of accented speech: The role of accent familirity and linguistic training. *International Journal of Listening*, 38(3), 203–215. https://doi.org/10.1080/10904018.2023.2252019

Mingxing, L. (2022). Vowel context, place difference, and the perceptual distinctiveness of sibilants. *Chinese Journal of Phonetics*, *17*(1), 96-105.

Mohan, D., & Maruthy, S. (2021). Vowel context effect on the perception of stop consonants in Malayalam and its role in determining syllable frequency. *Journal of audiology & otology*, *25*(3), 124-130. https://doi.org/10.7874/jao.2021.00087

Öztürk, C. Y. (2025). Using a Speech Corpus to Raise Phonemic Awareness Among L2 English Speakers: Production and Substitution of Dental Fricatives. In *Exploration of the Intersection of Corpus Linguistics and Language Science* (pp. 133-156). IGI Global Scientific Publishing.

Pan, S., Yang, Y., Kuang, C., & Huang, Z. (2026). Contextual similarity beyond segments: Interlanguage vocalic distance affects Chinese learners' perception of English alveolar–velar nasal contrasts. *The Journal of the Acoustical Society of America*, *159*, 2741-2753. https://doi.org/10.1121/10.0042986

Pérez-Ramón, R. (2024). Discrimination of Degrees of Foreign Accent across Different Speakers. *Languages*, *9*(3), 72. https://doi.org/10.3390/languages9030072

Pickering, L., & Huang, M. (2022). Teaching pronunciation in the context of multiple varieties of English. *Second Language Pronunciation: Bridging the Gap Between Research and Teaching*, 273-293. https://doi.org/10.1002/9781394259663.ch14

Rose, H., McKinley, J., & Galloway, N. (2021). Global Englishes and language teaching: A review of pedagogical research. *Language teaching*, *54*(2), 157-189.

RStudio Team. (2026). *RStudio: Integrated Development for R (Version 2026.01.1)*. In RStudio, Inc. https://www.rstudio.com/

Saito, K., Hanzawa, K., Petrova, K., Kachlicka, M., Suzukida, Y., & Tierney, A. (2022). Incidental and multimodal high variability phonetic training: Potential, limits, and future directions. *Language Learning*, *72*(4), 1049-1091. https://doi.org/10.1111/lang.12503

Schneider, E. W. (2011). *English around the world: An introduction*. Cambridge: Cambridge University Press.

Shejaeya, G. A., Roon, K. D., & Whalen, D. H. (2024). Talker variability versus variability of vowel context in training naïve learners on an unfamiliar class of foreign language contrasts. *Journal of Phonetics, 107*, 101369. https://doi.org/10.1016/j.wocn.2024.101369

Sridhanyarat, K. (2017). The Acquisition of L2 Fricatives in Thai Learners' Interlanguage. *3L: Southeast Asian Journal of English Language Studies*, *23*(1). http://doi.org/10.17576/3L-2017-2301-02

Uchihara, T., Karas, M., & Thomson, R. I. (2025). High variability phonetic training (HVPT): A meta-analysis of L2 perceptual training studies. *Studies in Second Language Acquisition*, *47*(3), 794-827. https://doi.org/10.1017/S0272263125100879

Uchihara, T., Webb, S., Saito, K., & Trofimovich, P. (2022). The effects of talker variability and frequency of exposure on the acquisition of spoken word knowledge. *Studies in Second Language Acquisition, 44*(2), 357–380. https://doi.org/10.1017/S0272263121000218

Ulla, M. B., Perales, W. F., & Paulino, F. B. (2025). Language teachers' awareness of their accents at a Thai university: Global Englishes (GE) in English language teaching (ELT). *Australian Review of Applied Linguistics, 48*(2), 177–198. https://doi.org/10.1075/aral.24037.ull

Van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in psychology*, *6*, 1000. https://doi.org/10.3389/fpsyg.2015.01000

Yeni–Komshian, G. H., & Soli, S. D. (1981). Recognition of vowels from information in fricatives: Perceptual evidence of fricative-vowel coarticulation. *The Journal of the Acoustical Society of America*, *70*(4), 966-975.

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.

Zhang, X., Cheng, B., Zou, Y., & Zhang, Y. (2025). Determining optimal talker variability for nonnative speech training: A systematic review and Bayesian network meta-analysis. *Journal of Speech, Language, and Hearing Research, 68*(3), 1006–1023. https://doi.org/10.1044/2024_JSLHR-24-0059

Zhang, K., Li, D., & Peng, G. (2024). Achieving perceptual constancy with context cues in second language speech perception. *Journal of Phonetics*, *103*, 101299.