

# Effects of Arabic Emphatic Environment on the Perception of Plain Consonants in L1 Arabic and L2 English for Arabic Native Learners

Abdulaziz Altamimi

Imam Mohammad Ibn Saud Islamic University (IMSIU), Saudi Arabia

**Abstract**—Although some research has examined the influence of emphatic environments on the perception of plain Arabic segments, little is known about whether this influence extends to the perception of English consonants by Arabic learners of English. Accordingly, the purpose of the present study is to investigate the effects of emphatic consonant environments on the perception of plain consonants in Arabic as a first language and in English as a second language. Fourteen second-language learners of English participated in a forced-choice identification task assessing their perception accuracy of Arabic and English consonants. The results showed that the presence of an emphatic environment (i.e., a plain consonant followed by a low back vowel) led learners to perceive the plain consonants /s/ and /ð/ as emphatic in both Arabic and English stimuli. These findings are discussed in light of the Perceptual Assimilation Model (Best, 1993) and the revised Speech Learning Model (Flege & Bohn, 2021).

**Index Terms**—speech perception, L2 perception, emphatic consonants, plain consonants

## I. INTRODUCTION

One salient property of Arabic phonology that has attracted research attention in both speech production and perception is emphatic spread. Emphatic spread refers to the influence of emphatic consonants on neighbouring segments, whereby an emphatic sound (e.g., /t/) might influence the surrounding sounds by changing its quality to assimilate with the articulatory characteristics of the emphatic sound. In Arabic, emphatic consonants are typically produced with a primary coronal constriction accompanied by a secondary retraction of the tongue body toward the pharyngeal or uvular region. Thus, emphatic consonants do not merely contrast locally with their plain counterparts; rather, they cause more global and gradient modifications in the acoustic and perceptual properties of adjacent segments.

Previous research has documented the phonetic consequences of emphatic spread in Arabic, particularly with respect to its effects on vowel realization. Acoustic studies have consistently shown that vowels occurring in emphatic environments tend to exhibit reduced F2, compared to vowels in non-emphatic contexts, reflecting the retracted tongue body gesture associated with emphasis (Al-Masri, 2009; Card, 1983; Hetzron, 1989). These findings indicate that emphatic consonants exert robust and systematic coarticulatory effects on neighboring segments. Beyond production, perceptual studies indicate that listeners are sensitive to these coarticulatory patterns and can exploit them as cues in segment identification and lexical access (e.g., Al-Tamimi, 2017; Hayes-Harb & Durham, 2016). However, the majority of existing research has focused on native speakers' production and perception within Arabic, with comparatively limited attention to how these coarticulatory cues are interpreted by Arabic native speakers in second language (L2) perception.

Despite growing interest in the interaction between phonological systems and speech perception, little is known about how emphatic spread in Arabic influences Arabic speakers' perception of non-native speech contrasts. In particular, it remains unclear whether long-term exposure to emphatic consonants affects Arabic learners' perceptual processing of English consonants that do not exhibit emphasis but may occur in vowel environments that overlap acoustically with the Arabic emphatic environment. The present study addresses this gap by examining the extent to which properties of the Arabic phonological system, specifically emphatic spread, influence Arabic second-language learners' perceptual sensitivity to English consonants.

## II. LITERATURE REVIEW

Arabic emphatic consonants (/t, s, ð/) are produced with a primary coronal constriction combined with a secondary retraction of the tongue body toward the pharyngeal or oropharyngeal region (Bin-Muqbil, 2006). This secondary articulation is a defining feature of emphatic segments that distinguishes them from their non-emphatic counterparts. Emphatic consonants are known to influence the articulatory and acoustic properties of adjacent vowels, creating a gradient effect that extends beyond the segment itself. Understanding these articulatory properties provides a foundation for examining their perceptual consequences, both within Arabic and in L2 contexts.

The secondary articulatory gestures associated with emphatic consonants result in coarticulatory patterns that listeners may use as perceptual cues. For example, vowels adjacent to emphatic consonants are typically lowered and backed,

assimilating toward the posterior articulatory space (Aldamen & Al-Deaibes, 2023; Alkhudidi et al., 2025; Al-Masri, 2009; Card, 1983; Hetzron, 1989; Khattab et al., 2006). Such acoustic modifications can bias listeners to perceive plain consonants as emphatic. This phenomenon, known as emphatic spread, illustrates the interaction between production and perception in speech processing. It is particularly relevant for speech perception studies because it demonstrates how contextual articulatory cues can systematically influence consonant categorization.

Emphatic spread has important implications for second-language (L2) perception, especially for Arabic learners of English. Because emphatic features can influence adjacent segments, Arabic learners may extend emphatic cues to English consonants, potentially leading to mistakenly perceiving plain consonants as emphatic. This pattern is consistent with theoretical frameworks such as the Perceptual Assimilation Model for L2 learners (PAM-L2), which predicts that non-native listeners assimilate unfamiliar sounds to the most perceptually similar native categories (Best & Tyler, 2007). This suggests that the acoustic consequences of emphatic spread may shape how learners encode and categorize English phonemes, biasing perception toward emphatic-like interpretations. Another account is provided by the revised Speech Learning Model (SLM-r), which proposes that L1 and L2 sounds are represented within a shared phonetic space and are evaluated relative to one another, with L2 speech perception shaped by learners' sensitivity to phonetic detail and cue weighting rather than categorical separation alone. Under this view, creating new L2 categories depends on whether learners detect reliable phonetic distinctions between L2 sounds and existing L1 categories (Flege & Bohn, 2021).

The model has been supported by perception studies showing that learners' accuracy in identifying L2 segments is systematically constrained by the phonetic proximity of those segments to existing L1 categories (e.g., Flege, 1995). Within the SLM-r framework, when L2 sounds overlap substantially with L1 categories, learners are less likely to establish autonomous L2 representations and instead rely on existing L1-based phonetic mappings, resulting in persistent perceptual difficulty (Flege & Bohn, 2021). Together, these theoretical frameworks provide a basis for interpreting how language-specific coarticulatory patterns, such as emphatic spread, may bias L2 perceptual encoding and categorization.

Jaber et al. (2023) investigated the acoustic properties of emphasis spread in Jordanian Arabic, focusing on how emphatic articulation affects surrounding vowels. Speech production data were collected from nine native speakers who produced minimal pairs containing the voiceless emphatic fricative /s/ and its plain counterpart /s/ in both word-initial and word-final positions. The analysis showed that emphasis was associated with consistent acoustic effects on adjacent vowels, specifically raised F1 and F3, and lowered F2 in vowels preceding and following emphatic consonants.

Khattab et al. (2006) examined the phonetic and perceptual correlates of emphasis in Arabic, with particular attention to how emphatic consonants influence adjacent vowel quality. Acoustic analysis demonstrated that emphatic consonants are consistently associated with F2 lowering and F1 raising in neighboring vowels. Importantly, the perception task showed that listeners relied heavily on these vocalic spectral cues when identifying emphatic versus plain consonants. The authors argued that emphasis is best characterized as a property extending over the syllable rather than a segmental feature. This study provides strong empirical support for the claim that vowel cues play a central role in emphasis perception.

Al-Masri (2009) investigated the role of emphatic spread in speech perception. Using a cross-splicing paradigm, Al-Masri tested whether listeners could identify consonants as plain or emphatic based solely on coarticulatory information present in the adjacent vowel. In this paradigm, emphatic consonants were placed in non-emphatic vowel environments, and plain consonants were inserted into vowels originally produced in emphatic contexts. The results demonstrated that when a plain /t/ was followed by a vowel originally produced in an emphatic context (e.g., a vowel from /ṭab/ spliced onto /t/), most participants misidentified the consonant as emphatic (/ṭ/). A similar pattern was observed for /s/, underscoring the importance of vocalic coarticulatory cues for perception.

Despite the theoretical and empirical significance of emphatic spread, research examining its effects on second-language speech perception remains limited. Most studies have focused on Arabic speech perception in L1 contexts, with comparatively little attention to how Arabic emphatic features may affect Arabic learners' perception of English (L2) consonants. Furthermore, previous research has not examined the role of L2 proficiency in modulating perceptual accuracy under conditions of emphatic coarticulation. Addressing these gaps is essential for understanding how language-specific phonetic patterns interact with non-native sound categorization.

The present study addresses these gaps by investigating whether Arabic emphatic spread biases Arabic learners' perception of English non-emphatic consonants (/t, s, ð/). Specifically, the study examines whether these consonants are more likely to be misperceived as their emphatic counterparts (/ṭ, ṣ, ḏ/) when they occur in phonetic environments that, in Arabic, license emphatic realization. In addition, the study explores whether learners' L2 proficiency is associated with increased perceptual accuracy, thereby exploring the extent to which experience with the L2 may override L1-based perceptual biases. The research questions are as follows:

1. When Arabic /t, s, ð/ occur in environments that motivate emphatic realization, are they more likely to be perceived as /ṭ, ṣ, ḏ/?
2. When English /t, s, ð/ occur in environments that would license emphasis in Arabic, do Arabic learners of English as an L2 misperceive them as emphatic sounds?
3. Is perceptual accuracy for English /t, s, ð/ associated with learners' L2 proficiency level?

## III. METHODOLOGY

**Participants**

In order to address the research questions, fourteen participants were recruited. Participants were native speakers of Saudi Arabic, aged between 25 and 30 years. A language background questionnaire was administered to assess L2 English proficiency. Participants were asked to provide self-reported ratings of their proficiency on a 7-point Likert scale and their length of residence in an English-speaking country. Based on these measures, participants were assigned to a high-proficiency group ( $n = 7$ ) and a lower-proficiency group ( $n = 7$ ). The high-proficiency learning group (HPL) had resided in the United States for more than four years and self-rated their overall English proficiency between 5 and 6. The lower-proficiency learning group (LPL) reported residing in the United States for less than one year and rated their English proficiency between 1 and 3.

**Stimuli for the Perceptual Experiment**

In order to establish the phonetic realization of the contrast between Arabic plain and emphatic consonants, production data from a single representative native speaker of Arabic were acoustically analyzed. The acoustic baseline and Arabic recordings were derived from a single native speaker to maximize experimental control and ensure consistency in the production of plain and emphatic consonants across vowel environments.

The speaker was asked to produce minimal or near-minimal monosyllabic Arabic word pairs, once with a plain consonant and once with its emphatic counterpart. The target words contained word-initial plain /t, s, ð/ and their emphatic counterparts, followed by different vowel environments.

The acoustic analysis of F1 and F2 values revealed that while emphatic consonants occurred across all vowel environments, plain consonants were produced in all environments except before low back vowels. Specifically, it was only emphatic consonants that were followed by vowels with F1 values higher than 600 Hz and F2 values lower than 1200 Hz. Based on these acoustic properties, low back vowels were grouped into a single vowel class that was used as a defining environment of emphatic consonants. The purpose of this classification was to examine whether introducing a plain consonant before a vowel class that naturally follows emphatic consonants would bias participants toward identifying the plain consonant as emphatic in perception.

To examine whether low back vowels influence the perception of a preceding plain consonant, stimuli consisting of Arabic and English non-words were compiled. Each stimulus contained one of the plain consonants /t, s, ð/ preceding either high front (/i, æ, u/) or low back vowels (/ɔ, ɑ, ʌ/). All vowels selected occur in both Arabic and English, ensuring cross-linguistic comparability. It was expected that placing a plain consonant before a low back vowel, an environment typically restricted to emphatic consonants, might shift listeners' perceptual categorization of the consonant toward an emphatic interpretation. The total number of experimental stimuli was 36 (18 in Arabic and 18 in English). The stimuli represented each of the three plain consonants appearing before the six vowel environments in Arabic and English.

To reduce predictability, emphatic counterparts of /ʃ, d, z/ were included as fillers in both the Arabic and English stimuli. Table 1 presents a sample of the stimuli used for the consonant /t/. All target consonants occurred in initial positions. A complete list of the stimuli is presented in Appendix.

The Arabic stimuli were recorded by a native speaker of Arabic, and the English stimuli were recorded by a native speaker of English, both in a quiet room using a digital recorder. Because Arabic plain consonants do not naturally occur before low vowels, as indicated by the acoustic analysis, the Arabic speaker—a trained linguist—was explicitly instructed to pronounce the non-words as transcribed in IPA. No production difficulties were observed when producing plain consonants before LB vowels.

TABLE 1  
A SAMPLE OF THE STIMULI USED FOR /t/

Arabic non-word stimuli in IPA	/tib/	/tæɡ/	/tuf/	/tɔk/	/tav/	/tʌt/
English non-word stimuli in IPA	/tip/	/tæk/	/tuv/	/tɔɡ/	/taf/	/tʌl/

**Procedures**

After completing the language background questionnaire, participants were seated in front of a laptop computer. A forced-choice identification task was administered, in which participants were instructed that they would listen to stimuli produced by an English native speaker (for the English stimuli) or by an Arabic native speaker (for the Arabic stimuli), and they needed to identify the most likely initial consonant in each stimulus by selecting one of two response options. One option represented a plain consonant, and the other represented its emphatic counterpart. Arabic orthographic symbols were used as options to select from, as they orthographically capture the plain–emphatic sound contrast. Although one option always corresponded to an emphatic consonant, the experimental trials consisted primarily of stimuli with plain initial consonants. Stimuli with emphatic initials were included only as fillers. Participants were presented with English stimuli first, followed by a short break, after which Arabic stimuli were introduced. This order was intended to minimize potential L1 phonological interference during the perception of non-native stimuli (i.e., cross-linguistic priming effects), as prior exposure to Arabic emphatic contrasts might increase participants' sensitivity to these features and influence subsequent L2 perception. However, as the presentation order was not counterbalanced, potential order effects cannot be entirely ruled out. Following data collection, participants' accuracy rates in identifying plain consonants were calculated. Stimulus presentation and response collection were implemented using Praat (Boersma & Weenink, 2023).

## IV. RESULTS

The study examined the effects of Condition (a within-subjects factor with four experimental conditions) and L2 Proficiency (a between-subjects factor with two levels) on perceptual accuracy. Participants were presented with the target consonants /t, ð, s/ in four conditions: (1) target consonants occurring in plain environments in Arabic stimuli (before high front-back vowels, H-FB); (2) target consonants occurring in emphatic environments in Arabic stimuli (before low back vowels, LB); (3) target consonants occurring in plain environments in English stimuli (before H-FB); and (4) target consonants occurring in emphatic environments in English stimuli (before LB). A mixed-design analysis of variance (ANOVA) was conducted to examine the effects of Proficiency and Condition on accuracy in perceiving the target consonants. Each consonant was analysed separately.

For the /ð/ sound, the analysis revealed a significant effect of Condition on accuracy,  $F(3, 10) = 105.19, p < .05$ , partial  $\eta^2 = .97$ , observed power = 1.00, indicating a large effect. Post hoc comparisons (Bonferroni-adjusted) showed that participants exhibited significantly higher accuracy for Arabic stimuli in the H-FB vowel environment ( $M = 95.24, SD = 12.10$ ) than for Arabic stimuli in the LB vowel environment ( $M = 9.52, SD = 15.60$ ) or English stimuli in the LB vowel environment ( $M = 54.76, SD = 24.83$ ),  $p < .05$ . There was no significant difference in accuracy between Arabic and English stimuli ( $M = 97.62, SD = 8.90$ ) when they occurred in the H-FB vowel environment,  $p > .05$ . A similar pattern was observed for English stimuli, with participants exhibiting significantly higher accuracy in the H-FB environment than in the LB environment,  $p < .05$ . Accuracy was also significantly higher for English stimuli than for Arabic stimuli in the LB environment,  $p < .05$ . Overall, participants exhibited the highest accuracy for English stimuli in the H-FB environment and the lowest accuracy for Arabic stimuli in the LB environment,  $p < .05$ . Figure 1 illustrates accuracy rates across the four experimental conditions.

There was no significant effect of Proficiency on accuracy for the /ð/ sound,  $F(1, 12) = 3.00, p > .05$ , partial  $\eta^2 = .20$ , observed power = .36. Post hoc comparisons showed that the HPL ( $M = 67.85, SD = 11.55$ ) and LPL ( $M = 60.71, SD = 15.50$ ) groups did not differ significantly,  $p > .05$ . In addition, there was no significant Proficiency  $\times$  Condition interaction,  $F(3, 10) = 0.31, p > .05$ , partial  $\eta^2 = .09$ , observed power = .09.

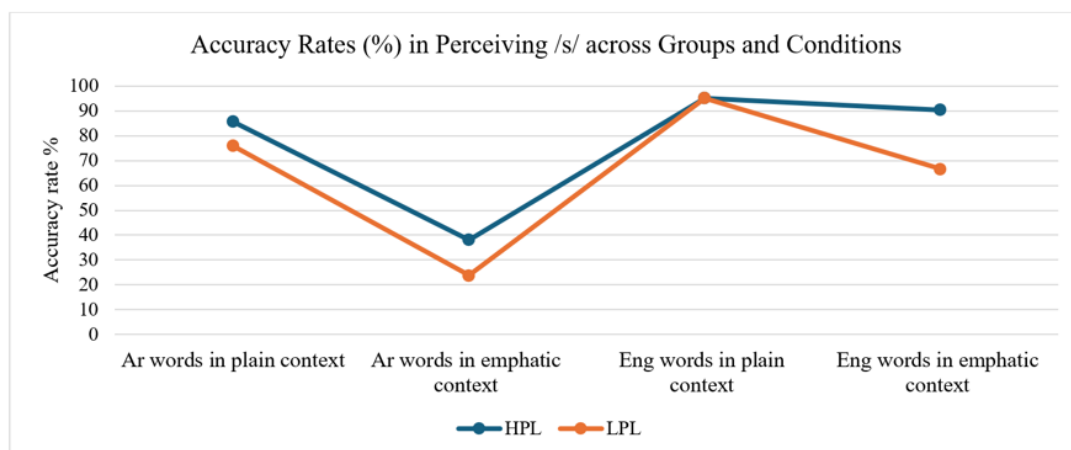


Figure 1.

As for the /t/ sound, the mixed-design ANOVA revealed a significant effect of Condition,  $F(3, 10) = 72.91, p < .05$ , partial  $\eta^2 = .96$ , observed power = 1.00, indicating a large effect. Post hoc comparisons (Bonferroni-adjusted) indicated that participants exhibited significantly higher accuracy for Arabic stimuli in the H-FB vowel environment ( $M = 92.86, SD = 14.10$ ) than for Arabic stimuli in the LB vowel environment ( $M = 36.01, SD = 15.60$ ),  $p < .05$ . There was no significant difference in accuracy between Arabic stimuli in the H-FB environment ( $M = 92.86, SD = 14.10$ ) and English stimuli in the LB environment ( $M = 83.33, SD = 21.68$ ),  $p > .05$ . Similarly, no significant difference was observed between Arabic and English stimuli in the H-FB environment (Arabic:  $M = 92.86, SD = 12.50$ ; English:  $M = 100.00, SD = 0.00$ ),  $p > .05$ . The Bonferroni post hoc test further showed no significant difference between English stimuli in the H-FB environment ( $M = 100.00, SD = 0.00$ ) and those in the LB environment ( $M = 83.33, SD = 21.68$ ),  $p > .05$ . Accuracy was significantly higher for English stimuli in both the H-FB and LB environments than for Arabic stimuli in the LB environment ( $M = 36.01, SD = 15.60$ ),  $p < .05$ . There was no significant effect of Proficiency on accuracy,  $F(1, 12) = 0.84, p > .05$ , partial  $\eta^2 = 0.07$ , observed power = 0.14, indicating that the HPL ( $M = 81.50, SD = 7.15$ ) and LPL ( $M = 77.45, SD = 13.50$ ) groups did not differ significantly overall,  $p > .05$ . However, there was a significant Proficiency  $\times$  Condition interaction,  $F(3, 10) = 4.68, p < .05$ , partial  $\eta^2 = 0.58$ , observed power = 0.74. Figure 2 illustrate this interaction, showing that the magnitude of the accuracy decline in emphatic contexts differed across proficiency groups. Specifically, the HPL group exhibited lower accuracy than the LPL group for Arabic stimuli in emphatic (BL) environments, whereas the HPL group showed higher accuracy than the LPL group for Arabic stimuli in plain (H-FB) environments and for English stimuli in emphatic environment (BL).

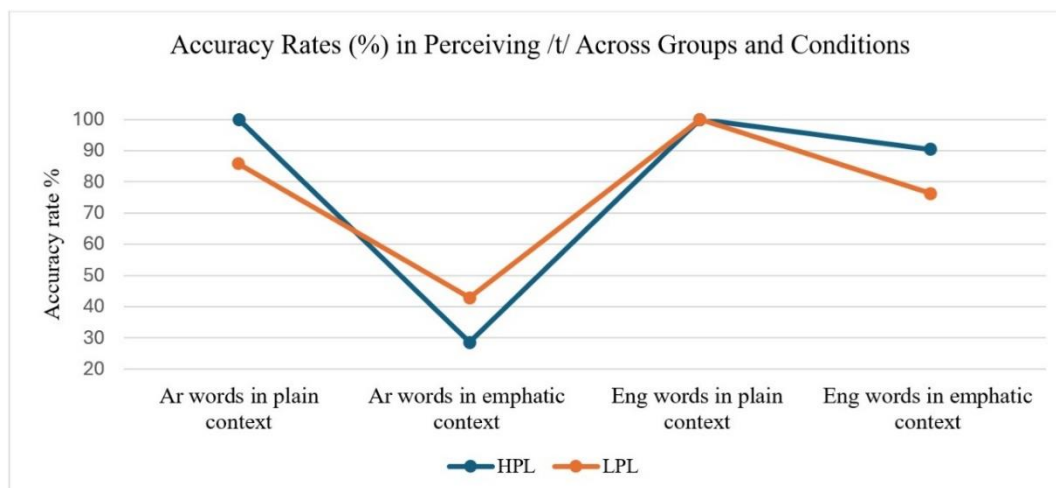


Figure 2.

Finally, the ANOVA was performed on the /s/ sound. The test revealed a significant main effect of Condition,  $F(3, 10) = 21.33$ ,  $p < .05$ , partial  $\eta^2 = .87$ , observed power = 1.00, indicating a large effect. Bonferroni-adjusted post hoc comparisons showed that participants exhibited significantly higher accuracy for Arabic stimuli in the H-FB vowel environment ( $M = 81.0$ ,  $SD = 17.1$ ) than for Arabic stimuli in the LB vowel environment ( $M = 30.1$ ,  $SD = 29.7$ ),  $p < .05$ . There was no significant difference in accuracy between Arabic stimuli in the H-FB environment ( $M = 81.0$ ,  $SD = 17.1$ ) and English stimuli in the LB environment ( $M = 78.5$ ,  $SD = 17.12$ ),  $p > .05$ . Likewise, no significant difference was observed between Arabic stimuli in the H-FB environment and English stimuli in the H-FB environment ( $M = 95.2$ ,  $SD = 12.11$ ),  $p > .05$ . Accuracy was significantly higher for English stimuli in the H-FB environment than for English stimuli in the LB environment,  $p < .05$ . In addition, accuracy for English stimuli in the LB environment was significantly higher than for Arabic stimuli in the LB environment,  $p < .05$ . The Bonferroni post hoc test further showed that participants exhibited significantly higher accuracy for English stimuli in the H-FB environment than for Arabic stimuli in the LB environment,  $p < .05$ . Figure 3 summarizes participants' accuracy rates across the experimental conditions. There was no significant main effect of Proficiency on accuracy for /s/,  $F(1, 12) = 0.99$ ,  $p > .05$ , partial  $\eta^2 = .08$ , observed power = .15, indicating that the HPL ( $M = 77.5$ ,  $SD = 20.05$ ) and LPL ( $M = 65.5$ ,  $SD = 10.7$ ) groups did not differ significantly,  $p > .05$ . There was also no significant Proficiency  $\times$  Condition interaction,  $F(3, 10) = 1.97$ ,  $p > .05$ , partial  $\eta^2 = .37$ , observed power = .36.

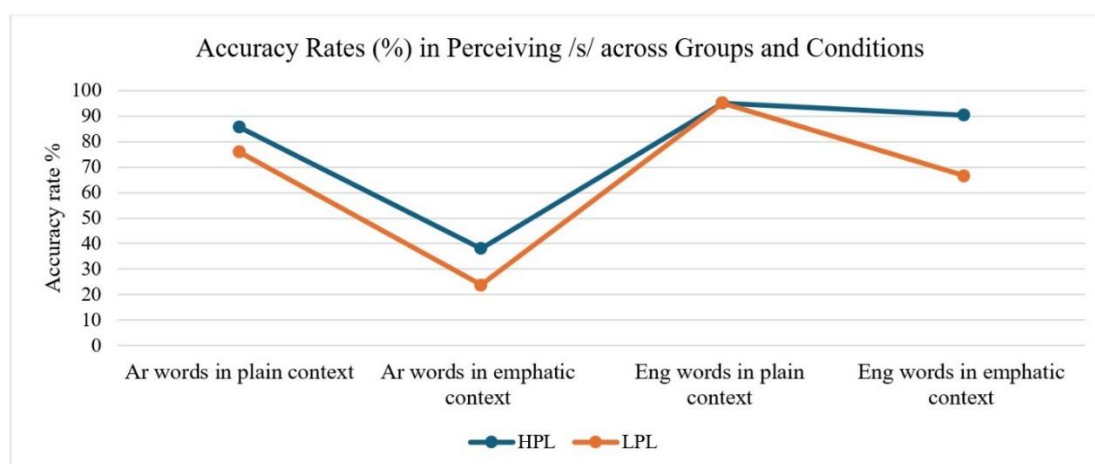


Figure 3.

## V. DISCUSSION

The first objective of this study was to examine whether Arabic learners of English accurately perceive the Arabic consonants /t, s, ð/ when these sounds are introduced in phonological environments that, in Arabic, would license emphatic realizations (/t̤, s̤, ð̤/) in production. The second objective was to determine whether perceptual cues derived from Arabic stimuli extend to English stimuli, and whether L2 proficiency modulates the perceptual accuracy for English consonants.

The overall findings indicate that Arabic ESL learners' perception of English consonants is systematically shaped by Arabic phonological constraints. Specifically, Arabic ESL learners accurately perceived Arabic plain initial consonants

(/t, s, ð/) when these sounds occurred in their canonical Arabic phonological environments, namely before high front and high back vowels (H-FB). In contrast, the same plain consonants were frequently misperceived as emphatic (/t̤, s̤, ð̤/) when they were introduced before low vowels (LB). Importantly, this vowel-driven perceptual pattern extended to English stimuli. English initial consonants were more likely to be perceived as emphatic before LB vowels, and as plain before H-FB vowels, mirroring the same perceptual pattern observed in Arabic. An exception emerged for the consonant /t/, which was correctly identified as plain in English stimuli regardless of vowel environment. Finally, the findings indicate that the influence of L1 phonology on L2 perception was not affected by increased L2 experience or proficiency. Although the high-proficiency learner (HPL) group reported longer residence in an English-speaking environment (over four years), they did not significantly outperform the lower-proficiency learner (LPL) group in identifying English plain consonants. These results suggest that prolonged exposure to L2 and higher proficiency alone were insufficient to override L1 influence on perception.

One explanation for the systematic association between the perception of Arabic and English plain consonants as emphatic and the presence of LB vowels is related to the production of emphatic sounds. The production of emphatic sounds is characterized by a tongue-body retraction in the vocal tract. This may have biased listeners to perceive a plain consonant as emphatic when followed by a low back vowel. This is because Arabic speakers strongly associate emphatic consonants with a low back vowel environment, due to shared articulatory and acoustic properties. This tendency is consistent with previous findings demonstrating that emphatic consonants are systematically associated with lowered F2 and raised F1 values in adjacent vowels, reflecting the articulatory retraction characteristic of emphasis (Al-Masri, 2009; Card, 1983; Hetzron, 1989; Khattab et al., 2006).

The overall findings, which suggest that the perception of non-native sounds may be mediated through a native-language phonetic filter, can be most plausibly interpreted within Best's (1993, 1995) Perceptual Assimilation Model (PAM) and Flege and Bohn's (2021) revised Speech Learning Model (SLM-r). PAM predicts that non-native sounds are perceptually assimilated to the closest native phonetic category, particularly when those sounds are perceived as acceptable exemplars of an existing native phoneme. The model also posits a category-goodness difference, which predicts that when two non-native sounds are assimilated to a single native category but differ in their degree of similarity to that category, listeners will perceive both sounds as belonging to the same native phonemic category, despite differences in their goodness of fit. For example, AlMahmoud (2013) found that American learners of Arabic failed to reliably discriminate between Arabic plain /t/ and emphatic /t̤/, suggesting that these learners assimilated both sounds to the single English phoneme /t/, with differences perceived only in terms of category goodness rather than phonemic contrast. The present study revealed a reversed perceptual pattern, demonstrating that a single non-native English plain consonant (/s/ or /ð/) was mapped onto two native categories (/s, s̤ and /ð, ð̤/, respectively). Given that language learners are predicted to perceive non-native sounds based on their articulatory similarity to native categories, these findings are compatible with PAM, suggesting that participants assimilated the English plain consonants to two native phonemic realizations conditioned by phonological context. This interpretation is consistent with Best and Tyler's (2007) claim that non-native speech perception is systematically biased by perceived similarity between native and non-native phones.

The findings are also compatible with PAM's prediction that L1 allophonic, phonetic, and phonemic patterns may interfere with L2 perception, thereby decreasing perceptual accuracy. Specifically, if participants assimilated English /s/ and /ð/ to their closest native plain counterparts, they may have transferred Arabic-specific phonetic adaptations associated with these sounds to English, resulting in the categorization of emphatic realizations as variants of the corresponding English plain phonemes. In other words, while PAM typically predicts that two non-native sounds may be perceived as instances of a single native category, the present findings suggest that a single non-native plain sound may be perceived as two environment-dependent native realizations (plain vs. emphatic), depending on the environment in which it occurs. That is, when a non-native plain consonant—assimilated to a native plain category—is presented in environments that, in Arabic, are restricted to emphatic consonants, listeners may interpret it as an emphatic native sound.

The findings suggest that the influence of L1 phonology on L1 perception is more robust than its influence on L2 perception. This pattern was evident in the results, which showed that although emphatic environments reduced perceptual accuracy for both Arabic and English stimuli, that effect was greater for Arabic stimuli, suggesting stronger phonological conditioning within the native language system.

In addition to PAM, the revised Speech Learning Model (SLM-r) proposes that the establishment of new L2 phonetic categories depends on the perceived phonetic distance between L1 and L2 sounds, as well as on learners' ability to perceive phonetic contrast between the sounds (Flege & Bohn, 2021). SLM-r predicts that when L1 and L2 sounds are phonetically similar, learners may continue to rely on L1-based cue weighting strategies even with increased experience. Thus, they are less likely to form autonomous L2 categories and instead tend to map L2 sounds onto existing L1 representations. Because Arabic and English share closely overlapping articulatory and acoustic properties for the consonants /s/ and /ð/, participants in the present study may have failed to establish autonomous L2 categories, and they may also have retained Arabic-specific patterns of cue weighting during L2 perception. Consequently, English consonants were processed through an L1 phonetic filter, with vowel-based coarticulatory cues exerting a strong influence on consonant identification. This account explains why the English consonants were perceived in a similar manner to the Arabic consonants, with perception modulated by vowel environment.

The results showed that while the perception of /s/ and /ð/ was conditioned by emphatic environments in both Arabic and English stimuli, the perception of /t/ in English stimuli remained relatively accurate in the emphatic environment (specifically for HPL). That is, subjects' judgment of /t/ was not influenced by emphatic environments. One possible explanation for this pattern relates to differences in the acoustic and articulatory properties of stops versus fricatives. Fricatives such as /s/ and /ð/ are featured by relatively long-duration spectral cues, which are more susceptible to coarticulatory influences from adjacent emphatic environments. In contrast, the stop consonant /t/ is characterized by brief, temporally localized cues, which might make it less affected by emphatic spread. Another account for why perceptual cues associated with Arabic /t/ were not transferred to English /t/ may relate to the magnitude of the articulatory contrast between the plain /t/ and its emphatic counterpart /t̤/ compared with the contrast between /s, ð/ and their emphatic counterparts. The greater articulatory and acoustic contrast between plain /t/ and emphatic /t̤/—notably in tongue-body retraction and voice onset time—has been shown to yield a particularly robust perceptual contrast relative to other emphatic pairs (AlMahmoud, 2013; Hayes-Harb & Durham, 2016; Kulikov & Hazi, 2025). Because Arabic /t̤/ is produced with the tongue tip positioned further back than in /t/, and because it exhibits shorter voice onset time (VOT) than the plain consonant (20–30 ms vs. 40–60 ms) (AlMahmoud, 2013), participants' ability to discriminate /t/ from /t̤/ even in emphatic environments may have been reinforced. Thus, the effect of emphatic vowel environments observed for the Arabic /t̤/ did not generalize to the English /t/, suggesting that the strong segmental cues associated with the sound /t/ may have mitigated the influence of contextual coarticulation in L2 perception.

The third research question addressed the effects of proficiency. The results revealed no clear effects of proficiency level on perception; the HPL group was overall as inaccurate in perceiving English plain sounds as the LPL group. This finding contrasts with previous research showing that reaching a highly advanced L2 level usually reflects enhanced perceptual performance (Best & Tyler, 2007; Flege & Bohn, 2021).

The current finding suggests that the LPL group's inaccuracy may not be attributed solely to limited experience in the L2. If the HPL group had outperformed the LPL group, this would indicate that the LPL learners had not yet reached a proficiency threshold enabling them to perceive these sounds accurately. Instead, the finding suggests that speech-perceptual cues may play a stronger role in shaping L2 perception than the level of L2 proficiency. This is consistent with SLM-r's claim that L1 influence can persist even at relatively advanced stages of L2 learning. An alternative explanation is that all subjects were late learners of English. It could be speculated that early learners with greater L2 experience should exhibit reduced interference from L1 phonetic categories. Consistent with this view, learning a second language after puberty may result in reduced accuracy in identifying some non-native sounds due to a loss of neural plasticity essential for establishing L2 phonetic representations (Flege et al., 1995). However, these explanations are merely speculative and should be taken with caution because the study did not administer an objective measure of L2 proficiency.

#### Limitations of the study and future directions

While the present study provides useful insights into the role of emphatic environments in L2 speech perception, several methodological limitations should be acknowledged. First, the small sample size may limit generalizability and reduce statistical power, particularly for detecting smaller effects and interactions. However, it is important to note that the study employed a repeated-measures design with multiple observations per participant, which enhances statistical sensitivity by reducing between-subject variability. In addition, the large effect sizes observed across conditions suggest that the main findings are robust despite the small sample. Nevertheless, future research should replicate the study with larger sample sizes to increase statistical power and improve the reliability of proficiency-related effects.

Second, L2 proficiency was assessed through self-reported measures and length of residence, without the use of standardized or objective proficiency tests. Although self-ratings are commonly used in L2 research, they may not accurately reflect perceptual competence in the L2. As a result, proficiency-based findings in the present study should be interpreted with caution. Future research should incorporate objective proficiency measures (e.g., standardized proficiency tests) to better assess learners' phonetic knowledge.

Third, another limitation concerns stimulus representativeness. The Arabic and baseline recordings were produced by a single native speaker, which, while ensuring high experimental control and consistency across conditions, may limit generalizability to the broader population of speakers.

Finally, the experimental stimuli consisted of non-words produced in controlled laboratory conditions. Although this design was necessary to isolate the effects of vowel environment on consonant perception, it limits ecological validity. Naturalistic speech contains greater phonetic variability, prosodic structure, and lexical information, all of which may modulate the influence of emphatic spread. Future studies should therefore extend this paradigm to more naturalistic listening conditions.

Taken together, these limitations suggest that the present findings should be interpreted with caution. Nonetheless, the study represents an important initial step toward understanding how language-specific phonological patterns—such as emphatic spread—influence L2 speech perception. Addressing these limitations in future research will allow for a more comprehensive account of the interaction between L1 phonological structure and non-native speech perception.

## VI. CONCLUSION

The present study assessed the interaction between plain and emphatic consonantal environments in speech perception. The findings indicate that participants' L1 Arabic phonetic knowledge shaped the perception of English sounds.

Specifically, the majority of the English plain consonants examined in this study were perceived in a similar manner to their Arabic counterparts, such that they were identified as emphatic when occurring before low back vowels. This pattern can be interpreted within the framework of the Perceptual Assimilation Model (PAM), which predicts that phonetic similarity between native and non-native sound categories may result in the assimilation of non-native sounds to existing native categories. In the present study, English plain consonants appear to have been assimilated to Arabic plain or emphatic categories depending on the vocalic context, reflecting the influence of language-specific coarticulatory patterns on L2 perception. The findings are also compatible with the revised Speech Learning Model (SLM-r), which proposes that L2 sounds are often perceived through the same phonetic system as L1 sounds, and that learners rely on familiar L1 phonetic cues when there is no evidence for new L2 categories. However, given the study's limited sample size and reliance on self-reported proficiency measures, any conclusions drawn from the findings should be regarded as tentative. Thus, the present findings should be viewed as only preliminary evidence motivating further investigation from future research.

#### APPENDIX. EXPERIMENTAL STIMULI

Arabic non-word stimuli in IPA	English non-word stimuli in IPA	Arabic non-word stimuli in IPA	English non-word stimuli in IPA	Arabic non-word stimuli in IPA	English non-word stimuli in IPA
/tib/	/tip/	/sik/	/sig/	/ðim/	/ðib/
/tæg/	/tæk/	/sæl/	/sær/	/ðæk/	/ðæg/
/tuf/	/tuv/	/suf/	/suv/	/dib/	/ðun/
/tøk/	/tøg/	/soq/	/soq/	/dib/	/ðol/
/tav/	/taf/	/sab/	/sam/	/ðak/	/ðag/
/tax/	/tal/	/sʌn/	/sʌd/	/ðʌt/	/ðʌn/

#### ACKNOWLEDGEMENTS

This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP2604).

#### REFERENCES

- [1] Aldamen, H., & Al-Deaibes, M. (2023). Arabic emphatic consonants as produced by English speakers: An acoustic study. *Heliyon*, 9(2).
- [2] Alkhudidi, A., Holt, R., Szalay, T., Demuth, K., & Benders, T. (2025). The acquisition of plain–emphatic consonant contrasts by Arabic-speaking children: An acoustic study. *Journal of Child Language*, 1-29.
- [3] AlMahmoud, M. S. (2013). Discrimination of Arabic contrasts by American learners. *Studies in Second Language Learning and Teaching*, 3(2), 261–291.
- [4] Al-Masri, M. S. (2009). *The acoustic and perceptual correlates of emphasis in Urban Jordanian Arabic* [Doctoral dissertation, University of Kansas]. ProQuest Dissertations & Theses Global. (Publication No. 3396411)
- [5] Al-Tamimi, J. (2017). Revisiting acoustic correlates of pharyngealization in Jordanian and Moroccan Arabic: Implications for formal representations. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 8(1), 1–40. <https://doi.org/10.5334/labphon.19>
- [6] Best, C. T. (1993). Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In J. C. Goodman & H. C. Nusbaum (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 289–303). Lawrence Erlbaum Associates.
- [7] Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). John Benjamins.
- [8] Bin-Muqbil, M. S. (2006). *Phonetic and phonological aspects of Arabic emphatics and gutturals* [Doctoral dissertation, University of Wisconsin-Madison].
- [9] Card, E. A. (1983). A phonetic and phonological study of Arabic emphatics. *Phonetica*, 40(1), 1–19. <https://doi.org/10.1159/000261678>
- [10] Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). York Press.
- [11] Flege, J. E., & Bohn, O.-S. (2021). The revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83). Cambridge University Press.
- [12] Flege, J. E., Munro, M. J., & MacKay, I. R. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*, 97(5), 3125-3134.
- [13] Hayes-Harb, R., & Durham, K. (2016). Native English speakers' perception of Arabic emphatic consonants and the influence of vowel context. *Foreign Language Annals*, 49(3), 557-572.
- [14] Hetzron, R. (1989). The Semitic emphatics: Facts and fictions. *Journal of Semitic Studies*, 34(2), 201–223. <https://doi.org/10.1093/jss/34.2.201>
- [15] Jaber, A., Omari, O., & Al-Jarrah, R. (2023). Emphasis spread: The domain and trigger. *Heliyon*, 9(7), e17476. <https://doi.org/10.1016/j.heliyon.2023.e17476>

- [16] Khattab, G., Al-Tamimi, F., & Heselwood, B. (2006). Acoustic and auditory differences in the /t-/ṭ/ opposition in male and female speakers of Jordanian Arabic. In *Perspectives on Arabic Linguistics XVI: Papers from the sixteenth annual symposium on Arabic linguistics* (pp. 131–160). John Benjamins.
- [17] Kulikov, V., & Hazi, R. (2025). Phonetic realization of the emphasis contrast in voiceless stops by speakers of four Arabic varieties: The case of F2 and VOT. *Arabic Linguistics*, 1(1), 1–33. <https://doi.org/10.1075/arli.00001.kul>

**Abdulaziz Altamimi** is Assistant Professor at Imam Mohammad Ibn Saud Islamic University (IMSIU). He holds a PhD in Applied Linguistics from the University of Nottingham (UK). His research interests centre around second language acquisition, psycholinguistics and speech perception. A major focus of his research is on the application of psycholinguistic methods to explore language processing and learning.