LEARN Journal: Language Education and Acquisition Research Network
ISSN: 2630-0672 (Print) | ISSN: 2672-9431 (Online)
Volume 19, No: 1, January – June 2026

Language Institute, Thammasat University
https://so04.tci-thaijo.org/index.php/LEARN/index

The Role of Tone Similarities in Thai-Mandarin Bilingual Lexical Access

Pimhathai Sonsuphap[a,*], Theeraporn Ratitamkul[b]

[a] pimhathai.ssp@gmail.com, Department of Linguistics, Faculty of Arts, Chulalongkorn University, Thailand
[b] theereporn.r@chula.ac.th, Department of Linguistics and Applied Linguistics for Language Education Research Unit, Faculty of Arts, Chulalongkorn University, Thailand
[*] Corresponding author, pimhathai.ssp@gmail.com

APA Citation:
Sonsuphap, P., & Ratitamkul, T. (2026). The role of tone similarities in Thai-Mandarin bilingual lexical access. *LEARN Journal: Language Education and Acquisition Research Network, 19*(1), 242-271.
https://doi.org/10.70730/LPUE6327

Received 17/04/2025

Received in revised form 01/09/2025

Accepted 10/09/2025

ABSTRACT

Tone is a crucial factor in lexical access for tonal language speakers and for tonal bilinguals, even when processing a non-tonal language (Lee, 2007; Malins & Joanisse, 2010; Shook & Marian, 2016; Shuai & Malins, 2017; Wang et al., 2020). This study investigated how high-proficiency Thai learners of Mandarin perceived and categorized L2 Mandarin tones and whether tone similarities affected their lexical access. In Experiment 1, 30 Thai learners of Mandarin listened to 200 Mandarin-Thai pairs of segmentally overlapped monosyllabic words with different tones, e.g., /pi-pìː/, and rated the similarity of the items using a 7-point Likert scale. The results showed that participants' ratings were likely based on acoustic similarities of tone contour. In Experiment 2, the same group of participants participated in an auditory lexical decision task with phonological priming to investigate the role of interlingual homophones (IHs) with similar tones in bilingual lexical access. The results revealed that reaction times (RTs) of IHs with similar tone contours were significantly faster than those of IHs with different tone contours and non-IHs, demonstrating facilitation effects. The facilitation effect corresponds to previous studies (Duyck, 2005, among others), indicating that lexicons of both languages may be phonologically integrated (Dijkstra & Van Heuven, 2002). Also, the RTs of IHs with different tone contours were not significantly different from those of non-IHs. Overall, this study highlights the prominent role of F0 properties for bi-tonal speakers, as tone contour similarities can induce a facilitatory phonological priming effect.

Keywords: lexical tones; perceptual assimilation; bilingual lexical access; Thai; Mandarin

Introduction

Studies on bilingual lexical access usually focus on how bilinguals manage both languages to access words in production and perception (van Assche et al., 2020). Do they access words from the target language lexicon alone or from both languages? To investigate this, researchers make use of three types of words which share common features between languages: interlingual homophones (hereafter IHs), which are words that share phonological features; interlingual homographs, which are words whose orthographic symbols overlap; and cognates, which are words sharing the same origin and spelled or pronounced similarly. Several studies have provided evidence which supports the claim that bilinguals activate both languages at the same time, which is known as non-selective lexical access (Dijkstra et al., 1999; Ju & Luce, 2004; Schulpen et al., 2003).

Spoken word recognition is different from visual word recognition in that the linguistic input contains language-specific phonetic information. A number of studies on bilingual spoken word recognition have indicated that IHs can induce a cross-language activation as bilinguals activate lexical candidates from both languages, resulting in either faster or slower reaction to IHs than control words (van Hell & Tanner, 2012). Some works indicate that IHs can induce cross language lexical competition leading to an inhibitory effect (Norris et al., 1995; Lagrou et al., 2011; Wang et al., 2020). To elucidate, when hearing a target word, other words that share similar phonological features will be activated from the listeners' mental lexicon. This process is called lexical competition (Marslen-Wilson, 1993; Mirman, 2016). If bilinguals activate lexical candidates from both languages, resulting in a larger number of candidates, it may take longer to resolve the words. Lagrou et al. (2011)

found inhibitory effects for IH items such as *lief* "sweet" – *leaf* among Dutch-English speakers participating in Dutch and English auditory lexical decision tasks, but there was no effect on English monolinguals. This suggests that phonological overlaps possibly lead to non-selective lexical access in bilinguals. Nevertheless, some studies on bilingual lexical access found facilitation effects, especially in lexical decision task and priming paradigm works (Ando et al., 2014; Duyck, 2005; Haigh & Jared, 2007; Zhou et al., 2010). For example, Haigh and Jared (2007) observed that French-English bilinguals reacted to IHs such as the word *sue* significantly faster and more accurately than to English control words in an English lexical decision task. The facilitation also indicates that the mental lexicon of bilinguals is integrated as similarities facilitate accessing the word (Chen et al., 2025).

There is evidence that lexical tones are as important as other segmental information in lexical access, as it was found that tonal language speakers, such as Mandarin speakers, required matching pitch information to access specific words from the lexicon (Lee, 2007; Malins & Joanisse, 2010). Furthermore, Wang et al. (2020), who conducted English auditory lexical access tasks, found that Mandarin-English speakers reacted to IHs whose pitch was adjusted to be similar to the tone of Mandarin words significantly slower than to non-IHs. However, an inhibitory effect was not found in IHs without pitch adjustment. This indicates that Mandarin-English speakers activate lexical competitors from a non-target language only when suprasegmental information is matched. Hence, it is interesting to investigate the role of lexical tones in lexical access of bilinguals who speak two tonal languages.

The current study aims to investigate how highly proficient Thai learners of Mandarin map Mandarin lexical tones to Thai lexical tones, and how tone similarities affect bilingual tonal speakers' lexical access. The first experiment examined tonal perceptual assimilation of Mandarin and Thai tones by having the participants rate the similarities between tones from each language. The second experiment was an auditory lexical decision task using Thai primes and Mandarin targets.

Literature review

Cross-linguistic perception

Cross-linguistic perception refers to an ability to perceive and discriminate non-native sounds including phonemes and tones. It is well known that L1 categories influence how non-native sounds are perceived (Best, 1995; Reid et al., 2015; So & Best, 2010). Language experience thus affects how non-native speakers perceive non-native language sounds. For example, tonal native speakers are more capable of discriminating non-native tones than non-tonal speakers (Lee et al., 1996; Chow et al., 2018). However, native tone experience may interfere with non-native language perception since tonal speakers are sensitive to pitch variation. When they perceive non-native tones with unfamiliar categorical boundaries, they may be unable to differentiate them (So & Best, 2010; Li, 2016). Also, native tonal density can affect non-native tone perception as well (Zhu et al., 2023).

Previous studies on segmental and suprasegmental perceptual assimilation have shown that non-native listeners categorize non-native sounds according to native phonetic categories, which is referred to as phonetic assimilation (Best et al., 1988; Flege, 1995, 2007; Leung, 2008; So & Best, 2010). For instance, Leung (2008) and Hao (2012) found that Cantonese speakers assimilated Mandarin T2 [35] to Cantonese low-rising tone [23], and Mandarin T1 [55] to Cantonese high-level tone [55], despite the fact that the first pair are not identical. Furthermore, experienced listeners may assimilate sounds that have similar lexical functions to their native categories. For example, French listeners assimilated voiced uvular fricative /ʁ/ to English voiced postalveolar approximants /ɹ/. Although these two sounds are very distinct, they are interchangeable at the lexical level (Best & Tyler, 2007).

Matching Thai and Mandarin Lexical Tones

Lexical tone refers to pitch-level changes of syllables, such as mean pitch and pitch contour, which affect the meaning of words (Singh & Fu, 2016; Yip, 2002), and languages which have these changes are called tonal languages. Both Thai and Mandarin are tonal languages in which pitch changes affect word meaning. Thai lexical tones can be categorized into five tones: Mid tone (sǐaŋ.sǎː.man) [33], Low tone (sǐaŋ.ʔèːk) [21], Falling tone (sǐaŋ.thoː) [51], Rising tone (sǐaŋ.triː) [45], and Falling-rising tone (sǐaŋ.càt.tà.waː) [214]. Mandarin lexical tones consist of four tones: T1, a level tone [55]; T2, a rising tone [35]; T3, a falling-rising tone [214]; and T4, a falling tone [51] (adapted from Ladefoged, 2001).

Matching lexical tones from the two languages can be complicated because Thai has five lexical tones while Mandarin has four lexical tones, with diverse contours. One of the methods for comparing them is to look into official transcription instructions. In 2007, Thailand's Office of the Prime Minister issued a Thai government gazette regarding transliteration criteria for Mandarin into the Thai alphabet. It recommends that T1 should be transcribed into Mid tone, T2 into Falling-rising tone, T3 into Low tone, and T4 into Falling tone. Nevertheless, this might not be an accurate comparison of lexical tones in the two languages in terms of acoustic properties and perception.

In fact, lexical tones can be divided into two major groups based on their pitch quality. Level tones or register tones are those whose pitches tend to be steady, so they rely on pitch height to convey meaning. Some languages, such as Cantonese, may have multiple level tones. For Thai, some previous work has proposed that the language had three level tones (mid, low, and high) and two contour tones (falling and rising) (Abramson, 1975; Tingsabadh & Abramson, 1993 as cited in Tsukada, 2019). However, more recent work involving acoustic analysis has revealed certain changes that affect tone categorization. To clarify, high tone, which was classified as level tone [44] in Abramson (1975), has become rising tone [334] for speakers under 20 years of age (Teeranon, 2007; Teeranon & Rungrojsuwan, 2009). Moreover, while in the past Abramson (1978) categorized low tone as a level tone because its F0 level carried information, Teeranon (2007) has found that F0 contour is sufficient to identify low tone. Hence, the Thai tones should be re-categorized into one level tone (Mid tone) and four contour tones, namely Low tone, Falling tone, Rising tone, and Falling-rising tone (Teeranon, 2007; Wu, Munro, et al., 2014).[1] Mandarin consists of one level tone, which is T1 or high level tone, and three contour tones: T2 or rising, T3 or falling-rising/dip-rise, and T4 or falling tone (Duanmu, 2007).

Wu, Munro, et al. (2014) conducted the acoustic-phonetic analysis of Thai and Mandarin tones based on 10 monosyllabic words, two non-words, and hums recorded by two native Thai speakers and two native Mandarin speakers. The analysis revealed that there are four tone groups according to their contours: (1) level tones, (2) falling tones, (3) rising tones, and (4) falling-rising tones, as can be seen in Table 1. Both languages contain only one level tone: Thai Mid tone [33] and Mandarin T1 [55]. Interestingly, a comparative study (Leelapornpinit, 2016) and perception experiments with native speakers and learners (Wu, Munro, et al., 2014) have also shown that these two tones are comparable even though T1 is higher in pitch level than Mid tone.

Table 1

Tone groups of Thai and Mandarin[2]

Tone groups	Thai	Mandarin
Level	Mid [33]	T1 [55]
Falling	Falling [51], Low [21]	T4 [51]
Rising	Rising [45]	T2 [35]
Falling-rising	Falling-rising [214]	T3 [214]

The second group is falling tones, which are Thai Falling tone [51], Thai Low tone [21], and Mandarin T4 [51]. Tone contours of the Thai Falling tone and Mandarin T4 are comparable because both start to fall from a very high pitch, while the Thai Low tone starts at a lower point. It is also commonly agreed among researchers and Thai learners of Mandarin that the Thai Falling tone and Mandarin T4 are comparable (Chen et al., 2023; Leelapornpinit, 2016; Rungruang & Mu, 2017; Wu, Munro, et al., 2014).

The third group is rising tones, which contain Thai Rising tone [45] and Mandarin T2 [35]. The contouring patterns of both tones are highly similar although Thai Rising tone starts at a higher point than Mandarin T2. A comparative study by Leelapornpinit (2016) proposes that both tones are one of the most similar tone pairs. Wu, Munro, et al. (2014)'s perception study also revealed that Thai listeners without Mandarin experience perceived T2 as Thai Rising tone. Interestingly, the same study showed that Thai listeners with Mandarin experience tended to perceive T2 as Thai Falling-rising tone rather than Rising tone. This also aligns with how Thai officials formally transcribe T2 words into the Thai alphabet.

Lastly, falling-rising tones consist of Thai Falling-rising tone [214] and T3 [214]. Although their tone contours are comparable, Thai listeners with Mandarin experience in Wu, Munro, et al. (2014)'s study tended to perceive T3 as Thai Low tone while those without experience perceived it as Thai Falling-rising tone. The researchers explained that Thai listeners without Mandarin experience were likely to assimilate tones that shared acoustic properties. In other words, assimilation was at the phonetic level. On the other hand, Thai listeners with Mandarin experience assimilated tones using their knowledge of the allophonic variation of T3, which was at the phonological level (Gandour, 1981; Heung, 2001; Leung, 2008). To clarify, T3 is pronounced as falling-rising [214] only when it appears in isolated words. When it appears with syllables having other tones or in the final position, half of its tone contour is pronounced as a falling tone [21]. As a consequence, T3, which is pronounced as a falling tone [21], seems to appear more frequently than T3 as a falling-rising tone [214]. This may be why T3 words are commonly transcribed with Low tone in Thai. Thus, Thai learners' perception in Wu, Munro, et al. (2014) was influenced in such a way that they perceived T3 as Thai Low tone [21] despite the fact that they were listening to monosyllabic words.

In conclusion, comparing Thai and Mandarin tones is not straightforward. Two of the pairs, T1-Mid tone and T4-Falling tone pairs, appear to be the most distinct, while the comparison is unclear for T2 and T3. T2 is often recognized as Thai Falling-rising tone despite its rising contour. Meanwhile, T3 is more complex since it has various contours depending on its position in the utterance. It can be matched with both Thai Falling-rising tone and Low tone. In the present work, we would like to examine how Thai learners of Mandarin with high proficiency perceive Mandarin lexical tones and assimilate them to Thai lexical tones.

Bilingual lexical access

It has been proposed that bilinguals' lexical access is different from that of monolinguals because of differences in the lexicon. One of the main issues in bilingual lexical access studies is whether lexical representations of the two languages are separately or jointly processed (Ju & Lace, 2004). The present study is interested in bilingual lexical access during listening or spoken word recognition. Unlike visual word recognition, spoken word recognition involves other factors that may affect lexical access, such as acoustic-phonetic information and duration. Wang (2021) state that there are two important questions in bilingual spoken word recognition studies. The first question is whether bilinguals retrieve words from both languages or from only the target language while listening to interlingual homophones. The second question asks whether bilinguals use language-specific phonetic cues in accessing the target language.

For the first question, if the language input only activates the target language, the process is viewed as language selective access. On the other hand, if the language input causes lexical competition of both languages, it is language non-selective access. Most research to date supports the language non-selective view (Lagrou et al., 2011; Persici et al., 2019; Schulpen et al., 2003; Spivey & Marian, 1999). For instance, the results of Lagrou et al. (2011), who investigated lexical decision tasks of Dutch-English and English speakers, revealed that only the reaction times to IHs of Dutch-English speakers, but not those of English monolinguals, were slower than to non-IHs. This suggests that there is an inhibitory effect that may have been caused by lexical competition between the two languages. Similarly, Schulpen et al. (2003) conducted a cross-modal priming lexical decision experiment with Dutch-English speakers. The experiment consisted of two conditions; the auditorily presented target word and the written prime word were Dutch-English IHs, or non-IHs. It was found that participants reacted to the IH stimuli significantly slower than to non-IHs. The results suggest that words from both languages are simultaneously activated.

Concerning the second question, whether bilinguals use language-specific phonetic cues in accessing the target language, several studies show that bilinguals employ language-specific phonetic cues such as Voice Onset Time (VOT) (Ju & Lace, 2004), accent (Lagrou et al., 2013), and pitch (Wang et al., 2020). It has been discovered that pitch or F0 is a critical cue in lexical access for tonal language monolinguals (Lee, 2007; Malins & Joanisse, 2010; Shuai & Malins, 2017). In a similar vein, bilinguals who speak one tonal language and one non-tonal language, such as Mandarin-English bilinguals, have also been found to be sensitive to F0 information even when listening to a non-tonal language (Braun et al., 2014; Ortega-Llebaria et al., 2015; Shook & Marian, 2016; Wang et al., 2020). In Wang et al. (2020)'s lexical decision task, Mandarin-English speakers exhibited an inhibitory effect when listening to IHs in English whose pitch had been adjusted to be similar to the pitch of Mandarin words. The effect was not found when the pitch had not been adjusted. Hence, it can be concluded that segmental information alone is not sufficient for tonal bilinguals to induce cross-language lexical competition. Moreover, this suggests that a similar pitch may lead to language non-selective access. Therefore, it is interesting to investigate whether Thai-Mandarin IHs, which have similar tones, would have any effect on how Thai learners of Mandarin retrieve words. In other words, the researchers would like to explore if tone similarity is sufficient to retrieve lexical competitors from both languages.

Experiment 1

Thai and Mandarin tones cannot be perfectly matched. Hence, it is interesting to know how Thai learners assimilate Mandarin tones to Thai tones. Experiment 1 investigated whether Thai learners of Mandarin with high language proficiency had perceptual assimilation patterns that depended solely on acoustic similarities or whether they were affected by phonological factors such as allophonic variations of T3 as found in Wu, Munro, et al. (2014)'s work.

Participants

Thirty-two native Thai speakers participated in the experiment. However, two of them were excluded because the accuracy of their responses to filler items in Experiment 2 was below 70%. This resulted in data from 30 participants (25 female and five male) aged 19 to 24 (mean = 21.03, SD = 1.25) being included in the analysis. All participants had learned Mandarin for over six years (mean = 9.43, SD = 3.09) and had never lived in a Mandarin-speaking country for over three months. Furthermore, none of the participants had knowledge of any tonal languages other than Standard Thai and Mandarin, including dialects. Neither did they regularly do musical activities such as singing or playing music. All participants had passed level 5 of the Hanyu Shuiping Kaoshi (HSK) test, indicating that they were highly proficient Mandarin learners. None of them had a problem with hearing or speaking.

Materials

Stimuli for this experiment were 200 pairs of Mandarin and Thai monosyllabic words of the structures CV or CVC. As the production of voiced and aspirated consonants could affect the F0 value, only the sounds /p, t, k, m, l/ were selected for initial consonants. Close vowels such as /i/ and /u/ were used to prevent the effect of intrinsic pitch. However, since voiceless velar plosive /k/ does not occur before high vowels in Mandarin, close-mid vowel /o/ was selected in this case. For the final consonants in the CVC stimuli, nasal endings /n/ and /ŋ/ were used as they are the only two possible final consonants in Mandarin. Consonants and vowels in the two languages are matched based on Leelapornpinit (2016). As a result, there were 10 different segment patterns. Applying Mandarin and Thai tones, there were 40 Mandarin words such as 逼 /pi/, 鼻 /pí/, 比 /pǐ/, 必 /pì/, and 50 Thai words such as ปี /pi:/, ปี่ /pì:/, ปี้ /pî:/, ปี๊ /pí:/, ปี๋ /pǐ:/. The majority of these words had meanings in Mandarin and Thai. Some of the words, however, were possible words in Mandarin and Thai but have no meaning, such as /liŋ/ in Mandarin and /pìŋ/ in Thai.

Forty Mandarin monosyllabic words were recorded by a female native Mandarin speaker who was from Jiangxi but lived in Beijing, China, and 50 Thai words were recorded by a female native Thai speaker who was born and lived in Bangkok, Thailand. Both speakers did not know any other tonal languages. The recording occurred in a closed recording studio. Each word was repeated three times, and only one was selected by a native speaker of each language for use in the experiment.

To create 200 pairs of Mandarin-Thai stimuli, each Mandarin word was paired with its counterpart in Thai, for example, /pi/-/pi:/, /pi/-/pì:/, /pi/-/pî:/, /pi/-/pí:/, /pi/-/pǐ:/. The order of items in the list was randomized using <PermuteBalancedNoDoublets> command in Praat program so that stimuli that had that same segment would not occur next to each other.

There were also 10 pairs of practice items to familiarize participants with the test. Two of the pairs had identical words (same segments, same tones) in Thai or Mandarin: /ma:/-/ma:/ (Thai) and /pa/-/pa/ (Mandarin); two pairs had words that shared segments but not tone: /ma:/-/má:/ (Thai) and /pa/-/pà/ (Mandarin); and, six pairs had Thai and Mandarin words which had overlapping segments: /ma/-/ma:/, /má/-/ma:/, /mǎ/-/ma:/, /mà/-/ma:/, /pà:/-/pǎ/, and /pǎ:/-/pá/.

All stimuli were examined for their validity through Item Objective Congruence (IOC) analysis by three Thai-Mandarin experts. Also, all stimuli were pre-tested by three Thai learners of Mandarin who had the same level of proficiency as the participants.

Procedure

The experiment was conducted in a quiet room at a university in Bangkok, Thailand. Before the experiment, participants were asked to fill in a demographic information form consisting of questions about their linguistic background and language usage. Then, participants wore headphones and sat in front of the computer. The experiment had two phases: the practice phase and the test phase. Before each phase, a written instruction in Thai was shown on the screen and explained by the researcher.

During the practice phase, 10 pairs of auditory stimuli were presented using the Praat program. Participants were asked to use a mouse to click a number button on the screen to rate the similarity of the sounds in each pair on a 7-point Likert scale: 1 as "totally different" and 7 as "very similar".

During the test phase, participants were asked to rate the similarities of 200 pairs of stimuli using the same instruction as in the practice phrase. Participants could take a break anytime they wanted by clicking the "Rest" button. Responses were collected using the Praat script.

Results and Discussion

To observe whether participants rated similarities of tones based on their acoustic characteristics, an acoustic analysis of the audio stimuli in this work was performed. In speech production, physiological factors such as differences in vocal tract configuration due to a speaker's gender and age affect fundamental frequency (hereafter F0). Also, the same speaker might produce the same word differently in different situations (Newman et al., 2001 cited in Tao et al., 2021). However, tonal speakers tend to perceive lexical tones as 'categorical' rather than 'gradient.' In this work, inter-speaker and intra-speaker variations are eliminated by normalizing F0 (Jitwirayanon, 2012). There are various methods to normalize F0 such as using Z-Score (Rose, 1987) or using psychoacoustic scales like Mel, Bark, ERB-rate, and semitone (Nolan, 2003). In this work, semitone is utilized as it is one of the best scales to reflect listeners' pitch and intonation perception (Nolan, 2003).

Firstly, the vocalic part of all stimuli was manually segmented and F0 extracted at every 10%, using the Praat script (Jitwirayanon, 2016). Then, the extracted F0 in hertz of all tokens was transformed to semitone using a reference value of 100.

Semitone = 12*log(transform value/reference value)/log(2) [3]

Figure 1

Tone contour of Thai and Mandarin tones

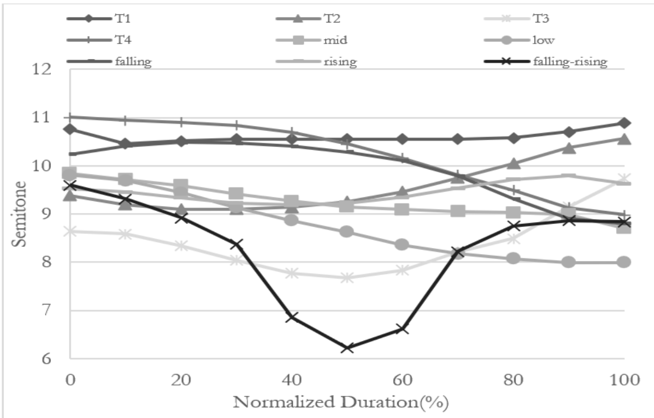


Figure 1 shows the average tone contours of Thai and Mandarin lexical tones. Thai and Mandarin lexical tones can be divided into four groups based on tone contour comparable to Wu, Munro, et al. (2014)'s tone categories, which are level tones (T1 and Mid tone), falling tones (T4, Low tone and Falling tone), rising tones (T2 and Rising tone), and falling-rising tones (T3 and Falling-rising tone).

Table 2 shows the mean and SD of each tone pair, and Figures 2 to 5 are scatter plots of responses for each tone pair. It can be seen that the Thai tone that was rated as most similar to Mandarin T1 was Mid tone (mean =5.03, SD = 2.03), followed by Falling tone, Rising tone, Low tone, and Falling-rising tone, respectively. Also, most participants rated T1-Mid tone stimuli remarkably high compared to other T1 pairs (see Figure 2). The Kruskal-Wallis test was employed for group analysis, showing that participants assimilated T1 to Mid tone significantly more often than to other tones (p-value < 0.001). This is probably because both T1 and Mid tone are the only level tones in Mandarin and Thai even though the pitch level of T1 (55) is higher than that of Mid tone (33) (see Figure 1).

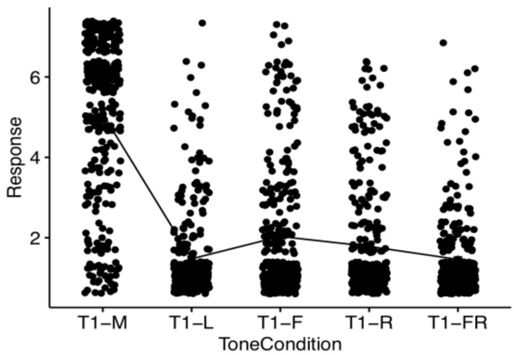
Table 2

Mean and SD (in parentheses) of tone pair similarity rating

Mandarin tones	Thai tones				
	Mid	**Low**	**Falling**	**Rising**	**Falling-rising**
T1	5.03 (2.03)	1.51 (1.11)	2.04 (1.62)	1.82 (1.43)	1.49 (1.06)
T2	1.72 (1.42)	1.58 (1.12)	1.56 (1.15)	5.16 (2.11)	2.48 (1.95)
T3	1.99 (1.65)	3.48 (2.39)	1.52 (1.13)	3.16 (2.32)	5.07 (2.22)
T4	1.37 (0.92)	1.39 (0.95)	5.41 (1.8)	1.58 (1.19)	1.41 (0.92)

Figure 2

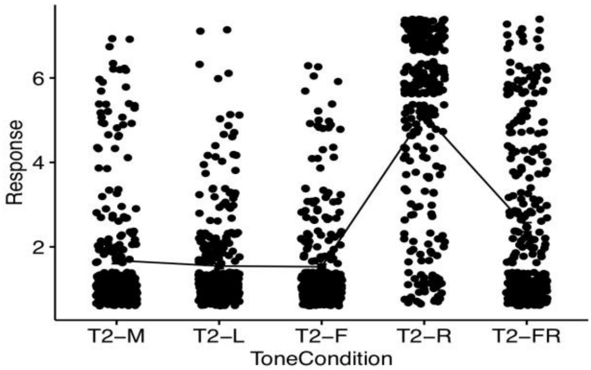
Scatter Plot of T1 Pairs Responses



The results also indicated that T2 was rated as most similar to Rising tone (mean = 5.16, SD = 2.11). The T2-Rising tone similarity score was significantly higher than other pairs (p-value < 0.001), which were rated low in terms of similarity. Both T2 and Rising tone have rising tone contours and are also the only rising tone in their language. Hence, it can be said that participants rated similarities based on acoustic properties. This, however, differs from the results of Wu, Munro, et al. (2014), which indicated that T2 was mapped onto Thai Falling-rising tone more often than Rising tone.

Figure 3

Scatter Plot of T2 Pairs Responses



Participants rated T3 as most similar to Falling-rising tone (mean = 5.07, SD = 2.22). Lower rated pairs were T3-Low tone, T3-Rising tone, T3-Mid tone, T3-Falling tone, respectively. As can be observed, results concerning T3 were somewhat diverse. Ratings of T3-Low tone (mean = 3.48, SD = 2.39) and T3-Rising tone (mean = 3.16, SD = 2.32) were relatively high compared to other pairs. Still, the Kruskal-Wallis test analysis revealed that T3-Falling-rising tone was rated significantly higher than T3-Low tone and T3-Rising tone (p-value < 0.001). Also, the scatter plot in Figure 4 shows that although the average response of T3-Low tone and that of T3-Rising tone pairs were somewhat high, the responses were densely plotted at the bottom of the graph. Again, this result reflects the influence of tone contour similarities in Thai learners' perception as both T3 and Falling-rising tones have a falling-rising contour. This contrasts with Wu, Munro, et al. (2014), who found that Thai learners mapped T3 to Low tones, as the low falling tone [21] is its most frequent realization.

Figure 4

Scatter Plot of T3 Pairs Responses

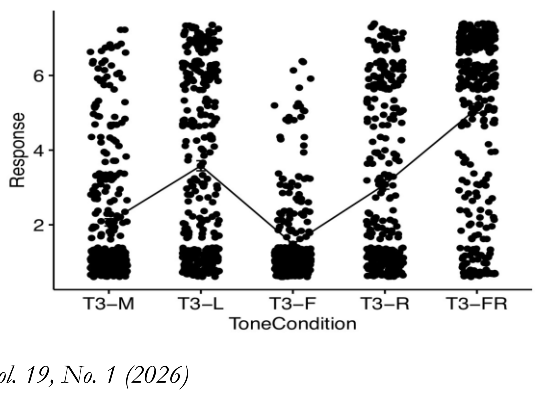
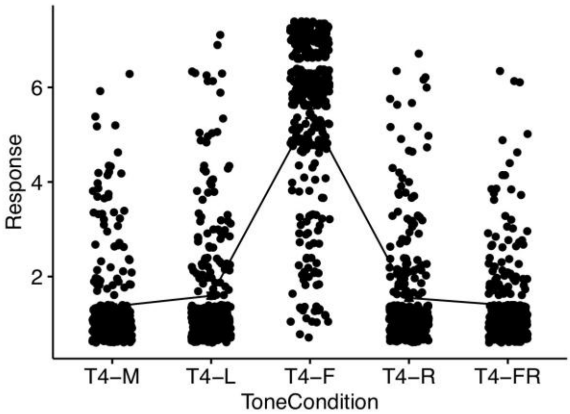


Figure 5

Scatter Plot of T4 Pairs Responses



Finally, participants perceived T4, which is a falling tone, to be most similar to Thai Falling tone. T4-Falling tone pairs were highly rated (mean = 5.41, SD = 1.8) and were also rated significantly higher than others (p-value < 0.001). The pair that was rated as the least similar was T4-Mid tone, followed by T4-Low tone, T4-Falling-rising tone, and T4-Rising tone, respectively. Results from the experiment showed that T4 was assimilated to Falling tone rather than Low tone despite their similar tone contours. This is possibly because T4 and Falling tone are more similar regarding F0 height and contour. Their tone contours start to fall at around 20% of the duration while the tone contour of Low tone immediately falls from the start. Further, the average F0 height of T4 and Falling tone are higher than that of Low tone.

To conclude, the highly proficient Thai learners of Mandarin tended to perceive T1 as similar to Mid tone, T2 to Rising tone, T3 to Falling-rising tone, and T4 to Falling tone. Furthermore, the results also revealed that participants assimilated the tones based on similarities in acoustic properties of the stimuli, which can be called phonetic assimilation according to the Perceptual Assimilation (PAM) model (Best, 1995; Best & Tyler, 2007). Participants might perceive the similarities of these tones because the primary cue to tone identity is tone contour for tonal language speakers (Liu & Samuel, 2004). Our results are not completely in line with those found by Wu, Munro, et al. (2014), which showed that inexperienced listeners relied solely on acoustic properties in tone assimilation whereas experienced listeners also used their phonological knowledge. We speculate that this might result from methodological differences. Our experiment presented the tone pairs and had participants rate them while Wu, Munro, et al. (2014) had participants listen to the sound and identify a similar tone in their language by choosing a labelled key on the screen. As the Mandarin and Thai tones in our experiment were auditorily presented in sequences, this perhaps offered an opportunity for the participants to rely on acoustic cues more heavily than in Wu, Munro, et al. (2014)'s setting.

Experiment 2

The lexical tone contains information crucial to lexical access for speakers of tonal languages such as Mandarin or Thai (Lee, 2007; Malins & Joanisse, 2010). Previous studies have also suggested that native tonal speakers are sensitive to suprasegmental information such as pitch, which influences lexical access of non-tonal target languages as well (Braun et al., 2014; Ortega-Llebaria et al., 2015; Shook & Marian, 2016; Wang et al., 2020). Furthermore, how learners perceive L2 sounds may affect how they access L2 words (Cutler et al., 2006; Weber & Cutler, 2004). This experiment examined whether IHs with assimilated tones based on Wu, Munro, et al. (2014) influenced bilingual lexical access in a Mandarin auditory lexical decision task with Thai primes.

Participants

Participants in this experiment were the same group as those in Experiment 1.

Materials

Stimuli in this experiment can be divided into three groups: (1) Mandarin targets, (2) Thai primes, and (3) pseudowords which followed Thai and Mandarin phonotactic constraints. The stimuli had either a CV or CVC structure. For initial consonants, only voiceless plosive, nasal, and approximant consonants were selected. This is because voiced plosive sounds, though present in Thai, do not exist in Mandarin, and fricative and affricate sounds in both languages are too fuzzy to find their counterparts. Thus, initial consonants were /p, ph, t, th, k, kh, m, n, l, w, j, h/. The vowels /aː, iː, uː, oː, əː, ia, ua, a, e, i, o, u/ were used. For CVC, the final consonants were /n, ŋ, j, w/.

Table 3

Examples of prime and target stimuli of each condition in Experiment 2

Condition	Prime	Target
1.1	Mid tone ทา /taː/ "to paint"	T1 他 /ta/ "he"
1.2	Falling-rising tone ผี /pʰǐː/ "ghost"	T2 皮 /pí/ "skin"
1.3	Low tone ข่า /kʰàː/ "galangal"	T3 卡 /kǎ/ "card"
1.4	Falling tone ปิ้ง /pîŋ/ "to grill"	T4 并 /pìŋ/ "sick"
2.1	Contour tone ขัง /kʰǎŋ/ "to imprison"	Level tone 摸 /mo/ "to touch"
2.2	Level tone ธง /tʰoŋ/ "flag"	Contour tone 变 /pìan/ "to change"

There were two main conditions. In Condition I, the prime and the target were IHs, which means that they segmentally overlapped, and they also had similar tones based on the assimilation patterns of Thai learners in Wu, Munro, et al. (2014). Hence, there were four sub-conditions: (1.1) T1-Mid tone, (1.2) T2-Falling-rising tone, (1.3) T3-Low tone, and (1.4) T4-Falling tone. Each sub-condition contained 15 pairs of stimuli, resulting in 60 pairs in total for Condition I. Condition II was the baseline condition of this experiment. In Condition II, there were 40 pairs of stimuli, in each of which the prime and the target share no segment or tone. If the prime had a contour tone (Low tone, Falling tone, Rising tone, or Falling-rising tone), then the tone of the target was a level tone, namely T1. On the other hand, if the prime had a level tone, the target had a contour tone. The tones in each pair were selected in such a way as to make sure that there were no phonological overlaps between the prime and the target. To conclude, there were altogether 100 pairs of stimuli for Condition I and II.

There were also 100 fillers. Fillers were divided into two categories: 1) the prime was a real Thai word while the target was a Mandarin pseudoword (50 pairs), and 2) both the prime and target were pseudowords in Thai and Mandarin (50 pairs). Mandarin pseudowords were derived from the Database of Word-level Statistics for Mandarin Chinese (DoWLS-MAN), and pseudowords that sounded like real Thai words were avoided. In addition to test items, 10 practice pairs were created. Five had real word targets, and the other five had pseudoword targets. All audio files were recorded by the same people as those in Experiment 1.

All stimuli were examined for their validity through Item Objective Congruence (IOC) analysis by three Thai-Mandarin experts. Also, all stimuli were pre-tested by three Thai learners of Mandarin who had the same level of proficiency as the participants.

Procedure

The experiment was conducted in a quiet room at a university in Bangkok, Thailand. Participants were individually tested. Firstly, participants were asked to sit in front of the computer and wear headphones. The experiment was run by the PsychoPy program. For each trial, an auditory prime was present along with a yellow triangle on the screen, followed by a target accompanied by a blue triangle. The visual shapes were presented as a way to keep the participants focused. The interval between the prime and the target was 1500 milliseconds, and the interval between trials was 3000 milliseconds. The participants were asked to decide whether the second sound they heard had a meaning in Mandarin by pressing the "a" button if it was meaningful and the "l" button if it was not. Before the experiment, the instructions were given orally in Thai, and written instructions on the screen were given in Mandarin to prepare the participants to focus on Mandarin targets.

The experiment had two phases: the practice phase and the test phase. In the practice phase, participants were asked to perform on 10 practice trials. Then, 200 pairs of stimuli were presented in the test phase. During the test phase, there was a pause after every 25 pairs of stimuli. Reaction times (RTs) as well as response accuracy were recorded.

Results and Discussion

Data from participants who made more than 30% errors were excluded (n = 2). This resulted in data from 30 participants. Only responses and RTs of the real word conditions (IHs and non-IHs) were analyzed. For error analysis, the Binominal General Linear Mixed model, which was from the glme package in the Rstudio program, was used to analyze the error rate of both conditions. Error analysis indicated that there were no significant differences between IHs and non-IH conditions (p-value = 0.757). Errors were found most frequently in Falling-rising tone-T2 condition (1.95%), followed by Falling tone-T4 condition (1.83%), Low tone-T3 condition (1.82%), Mid tone-T1 condition (1.66%), non-IH contour-level condition (0.56%), and non-IH level-contour condition (0.48%), respectively.

Table 4

Linear Mixed-Effects analysis results of IHs and non-IHs stimulus pairs

Random effects:					
Groups	**Name**	**Variance**	**Std.Dev**		
target	(Intercept)	0.07527	0.2743		
subject	(Intercept)	0.15820	0.3977		
Residual		0.76802	0.8764		
Fixed effects:					
	Estimate	**Std. Error**	**df**	**t value**	**Pr(>|t|)**
(Intercept)	1.58475	0.08354	44.15021	18.969	<2e-16 ***
IHnon-IH	0.13575	0.06534	93.54797	2.078	0.0405 *

Inaccurate responses were then excluded from the analysis, and only accurate responses and RTs were exported to the RStudio program. The Linear Mixed-Effects (LME) model was selected to analyze data using LmerTest functions in the Lme4 Test package in R. The fixed effect factor was Word Type (IHs VS Non-IHs), and subjects and items were random effect factors. The emmeans function in emmeans package was also performed. Results from Linear Mixed-Effects analysis of RTs showed the main effect of Word Type (t = 2.078, p-value = 0.0405). These results suggested that participants reacted to items from Condition I, which were IHs, significantly faster than those from Condition II, which were non-IHs. In other words, the participants responded significantly faster when the prime and target were IHs with similar tones (according to Wu, Munro, et al. (2014)'s tonal assimilation results) than when encountering non-IH stimulus pairs.

Table 5

Linear Mixed-Effects analysis results of assimilated tone pairs vs non-IHs (condition II)

Random effects:					
Groups	**Name**	**Variance**	**Std.Dev**		
target	(Intercept)	0.06873	0.2622		
subject	(Intercept)	0.14602	0.3821		
Residual		0.73364	0.8565		
Fixed effects:					
	Estimate	**Std. Error**	**df**	**t value**	**Pr(>|t|)**
(Intercept)	1.46708	0.08945	57.13545	16.400	< 2e-16 ***

conditionnon-IH	0.25277	0.07417	63.15566	3.408	0.00114 **

However, results from Experiment 1 revealed that participants' patterns of tonal assimilation only corresponded to two out of the four sub-conditions. To explain, only Mid tone-T1 and Falling tone-T4 assimilations were found in Experiment 1. Hence, we would like to investigate further to see whether there was any difference between assimilated tone pairs and non-assimilated tone pairs. To do so, stimulus pairs in Condition I were divided into two sub-groups: (1) assimilated group consisting of Mid tone-T1 pairs and Falling tone-T4 pairs, and (2) non-assimilated group consisting of Falling-rising tone-T2 pairs and Low tone-T3 pairs. Then, RTs of the assimilated group were used as fixed effect factors compared with Condition II (Word Type: Assimilated VS Non-IHs). Besides, subjects and items were random effect factors. The Linear Mixed-Effect analysis showed a main effect of Word Type ($t = 3.408$, $p = 0.00114$). This indicates that the participants responded to IHs with assimilated tone pairs significantly faster than to non-IHs. To illustrate, the mean RT of assimilated pairs was 1.47 s, and that of non-IHs was 1.72 s. There was thus a facilitation effect when the participants were primed with words which phonologically overlapped with and had similar tones to target words.

To investigate whether the participants reacted to assimilated tone pairs differently from non-assimilated tone pairs, Assimilation (Assimilated VS Non-assimilated) was a fixed effect factor, and subject and items were random effect factors. The Linear Mixed-Effects analysis of RTs showed a main effect of Assimilation ($t = 3.153$, $p = 0.0026$). Compared to the non-assimilated group, participants reacted to the assimilated group significantly faster (mean = 1.47 s for the assimilated group and mean = 1.71 s for the non-assimilated group). This analysis suggests that tone similarities which are based on acoustic properties could produce a stronger facilitation effect than non-assimilated tones.

Table 6

Linear Mixed-Effects analysis results of assimilated and non-assimilated stimulus pairs

Random effects:			
Groups	**Name**	**Variance**	**Std.Dev**
target	(Intercept)	0.06254	0.2501
subject	(Intercept)	0.16762	0.4094
Residual		0.68149	0.8255

Fixed effects:					
	Estimate	**Std. Error**	**df**	**t value**	**Pr(>|t|)**
(Intercept)	1.46642	0.09197	50.82988	15.945	<2e-16 ***
conditionnon-as	0.23921	0.07587	55.76427	3.153	0.0026 **

Table 7

Linear Mixed-Effects analysis results of non-assimilated and non-IHs stimulus pairs

Random effects:					
Groups	**Name**	**Variance**	**Std.Dev**		
target	(Intercept)	0.06849	0.2617		
subject	(Intercept)	0.16459	0.4057		
Residual		0.87693	0.9364		
Fixed effects:					
	Estimate	**Std. Error**	**df**	**t value**	**Pr(>|t|)**
(Intercept)	1.70333	0.09391	54.83702	18.139	<2e-16 ***
conditionnon-IH	0.01507	0.07628	63.43690	0.198	0.844

Furthermore, it is interesting to see whether participants would react differently to IHs containing acoustically less similar tones compared to non-IHs. In this analysis, Word Type (Non-assimilated VS Non-IHs) was a fixed effect factor, and subject and items were random effect factors. Results from the Linear Mixed-Effects analysis did not show any main effect. This indicates that RTs of non-assimilated IHs pairs were not significantly different from those of non-IHs (mean = 1.70 s for non-assimilated IHs, and mean = 1.72 s for non-IHs).

To conclude, the results from Experiment 2 suggest that acoustic-phonetic similarities indeed affect how participants access their lexicon. Specifically, participants responded significantly faster to IH than to non-IH stimulus pairs. This indicates that the facilitation effect is exhibited when the prime and target share segmental and tonal similarities based on tonal perceptual assimilation in Wu, Munro, et al. (2014). Moreover, it was revealed that RTs were significantly faster only when the tones of the prime and the target were similar and in accord with participants' assimilation patterns. As can be seen, the RTs of Mid tone-T1 and Falling tone-T4 pairs were significantly faster than those of Falling-rising tone-T2 and Low tone-T3 pairs as well as non-IHs. Nonetheless, the RTs of Falling-rising tone-T2 and Low tone-T3 pairs were not significantly different from those of non-IHs. Despite segmental overlaps in these stimulus pairs, it is still not adequate to lead to the facilitation effect. The observed facilitation may be attributed to cross-language activation, whereby hearing a Thai prime also activates a Mandarin

IH with similar tone contour. This co-activation could expedite lexical judgment for targets containing acoustically similar tones, whereas targets without such similarity would not benefit from prior activation during prime processing. An alternative explanation is that the effect reflects a joint representation of lexical items with corresponding tones in both languages. In this view, IHs with similar tones are encoded as a single lexical representation encompassing meanings from both languages. In any case, this points to the conclusion that without tone similarities which are based on acoustic properties, segmental overlaps alone are insufficient to produce the facilitation effect for bi-tonal bilinguals in accessing the target word which is in the other language. Thus, tone similarities are a crucial cue in bilingual lexical access. Also, the facilitation effect found in this work is in line with previous studies, which also suggest that lexicons in the two languages are integrated (Dijkstra & Van Heuven, 2002; Duyck, 2005; Zhou et al., 2010; Ando et al., 2014).

General Discussion and Conclusion

The present study aims to investigate tonal perceptual assimilation of Thai and Mandarin tones by highly proficient Thai learners of Mandarin and to observe whether the assimilation patterns affect how they access the lexicon. In Experiment 1, Thai learners of Mandarin were assigned to listen to Mandarin-Thai stimulus pairs whose members shared segments but not tones and rated them using a 7-point Likert scale. The results showed that the participants tended to assimilate tones based on their acoustic characteristics. In short, they perceived tones which were not identical but had similar tone contours to be the most similar. Four tone pairs found to be similar were 1) T1 and Mid tone, 2) T2 and Rising tone, 3) T3 and Falling-rising tone, and 4) T4 and Falling tone. The only Thai tone not perceived as similar to any Mandarin tone was Low tone. Even though the tone contour of Low tone could be categorized as a falling tone along with Falling tone and T4, the tone contour of Falling tone is more similar to T4 since their starting points are a little higher than Low tone. In other words, participants had phonetic assimilation patterns because they matched Thai and Mandarin tones based on their acoustic phonetic similarities (Flege, 1995, 2007; So & Best, 2010). These results are different from Thai learners' tonal perceptual assimilation in Wu, Munro, et al. (2014), who claimed that assimilation was also based on tone allophone knowledge which is phonological level knowledge. The disparity might be due to the methodology as participants in Wu, Munro, et al. (2014) were asked to choose a Thai tone that they thought most similar to what they heard by pressing one of the five keys. On the other hand, the current study presented stimuli in pairs for the participants to perform similarity ratings. This may have made participants focus more on tone similarities than on their categorization.

In Experiment 2, an auditory priming lexical decision task was conducted. The test pairs were either (1) Thai prime and Mandarin target which were IHs with assimilated tone patterns from Wu, Munro, et al. (2014), or (2) Thai prime and Mandarin target which were non-IHs. It was found that there was a facilitation effect for IH stimulus pairs compared to non-IHs. However, based on the findings in Experiment 1, the T1-Mid tone and T4-Falling tone pairs were the only pairs containing IHs with acoustically similar tones while the tone contours of T2-Falling-rising tone and T3-Low tone barely shared similarities of F0s. Hence, although participants responded to IH stimulus pairs faster than to non-IH stimulus pairs, we further investigated whether there was any difference between assimilated and non-assimilated tone pairs. Statistical analysis showed that participants judged lexical items in T1-Mid tone and T4-Falling tone IH pairs significantly faster than those in T2-Falling-rising tone and T3-Low tone pairs. This suggests that the degree of tone similarities affects how the L2 lexicon is retrieved. Findings from this study correspond with previous studies on lexical access of tonal language, which point out that pitch information is a crucial cue (Lee, 2007; Shuai & Malins, 2017).

Moreover, this study confirms that segmental overlap combined with tone contour similarities can influence lexical access in tonal languages. That is, the current study indicates that IH pairs with assimilated tones exhibit a facilitation effect, but not those with non-assimilated tones. These results support the idea of non-selective language access as bilinguals cannot disregard non-target language while processing the target language, resulting in cross-language activation (e.g., Spivey & Marian, 1999). In fact, homophone effects on cross-language activation are varied (van Hell & Tanner, 2012). Some previous studies found that bilinguals reacted to IHs slower than to non-IHs, exhibiting an inhibitory effect (Weber & Cutler, 2004; Lagrou et al., 2011, 2013; Wang et al., 2020). Tonal bilingual lexical access studies like Wang et al. (2020) showed that Mandarin-English speakers exhibited an inhibitory effect for IH stimuli which had a similar pitch to an existing lexical tone. The explanation is that lexical access takes longer when bilinguals listen to IHs because the candidates are activated from both languages, causing a greater number of competitors, which leads to cross-language lexical competition. However, homophone facilitation effects were also found in other studies (Ando et al., 2014; Duyck, 2005; Haigh & Jared, 2007; Lemhofer et al., 2004; Zhou et al., 2010).

One possible account for the facilitation effect is cross-language activation; exposure to a Thai prime may simultaneously activate Mandarin IHs with matching tonal patterns. Such co-activation could speed recognition of targets that share these tonal features, while targets lacking similarity would not receive the same processing advantage. We also speculate that our findings of facilitation effects result from a joint representation of both languages' tones. To elucidate, tones from Thai and Mandarin which have similar acoustic properties share a joint representation in learners' lexicon, resulting in one representation of the sound having meanings in both languages. This is perhaps similar to Wu et al. (2014), who claim that a priming effect occurs during the stage when the representation of two forms of tone variations is integrated. Wu et al. (2014) found that tone contour variation that was not contrastive in the language had a very similar effect as the 'actual' tone. They ran an auditory priming lexical decision task for Jinan speakers with three conditions: (1) prime and target were identical, (2) prime and target had slight differences in the shape of the pitch contour but were not contrastive, and (3) prime and target had different tones. The result revealed that both Conditions 1 and 2 showed a facilitation effect with minimal differences, while there was an inhibitory effect for Condition 3. Thus, it can be concluded that the results of this study reflect non-selective language access. The facilitation effect supports the Bilingual Interaction Activation (BIA+) model (Dijkstra & van Heuven, 2002), which proposes that any orthographic, phonological, and semantic similarities should benefit lexical access because both languages are activated in parallel, and their lexicons are integrated (Frances et al., 2021). This also supports the idea of the Bilingual Language Interaction Network for Comprehension of Speech (BLINCS) model (Shook & Marian, 2013), which posits that bilinguals have a shared phonological level that enables co-activation.

Understanding how high proficiency Thai learners of Mandarin perceive Mandarin tones can assist tone perception and pronunciation practice for Thai learners in early stages of Mandarin learning. Also, since Thai learners associate IHs with similar contour tones in both languages, teaching IHs with similar contour tones may facilitate vocabulary acquisition for Thai learners.

Nevertheless, the present study has certain limitations. Both experiments were conducted with a high-proficiency learner group only, precluding comparisons with other populations such as native Mandarin speakers or lower-proficiency Thai learners. In addition, the study did not examine IH primes across the full range of Thai lexical tones, which would have allowed a more comprehensive assessment of their influence on participants' responses. Furthermore, the stimuli design did not account for phonological neighborhood density or homophone frequency effects. Future research should incorporate these factors to better understand their influence on lexical access.

About the Authors

Pimhathai Sonsuphap: A master's degree student in the Department of Linguistics at Chulalongkorn University. She is interested in second language acquisition, psycholinguistics, and phonetics. Currently, she is a language instructor at the Center for Thai as a Foreign Language, Faculty of Arts, Chulalongkorn University.

Theeraporn Ratitamkul: An associate professor in the Department of Linguistics and a member of the Applied Linguistics for Language Education Research Unit, Faculty of Arts, Chulalongkorn University. Her areas of interest include language acquisition, psycholinguistics, and cognitive linguistics.

Endnotes

[1] The names of the tones are different in each work. Rising tone and Falling rising tone in this work are 'high tone' and 'rising tone' in Teeranon(2007).

[2] The numbers in the brackets are the Chao tone numerals indicating pitch and contour on a five-point scale. These numbers were adapted from Ladefoged, 2001, pp. 237-238.

[3] This formula was adopted from Jitwirayanon (2012).

References

Abramson, A. S. (1975). The tones of central Thai: Some perceptual experiments. In J. G. Harris and J. R. Chamberlain (ed.), *Studies in Tai linguistics in honor of William J' Gedney* (pp. 1-16). Central Institute of English Language.

Abramson A. S. (1978). Static and dynamic acoustic cues in distinctive tones. *Language and Speech, 21*(4), 319–325. https://doi.org/10.1177/002383097802100406

Ando, E., Jared, D., Nakayama, M., & Hino, Y. (2014). Cross-script phonological priming with Japanese Kanji primes and English targets. *Journal of Cognitive Psychology, 26*(8), 853–870. https://doi.org/10.1080/20445911.2014.971026

Best, C. T. (1995). A direct realist view of cross-language speech perception. In Winifredstrange (Ed.), *Speech perception and linguistic experience: Issue in cross language research* (pp. 171-204). York Press.

Best, C. T., McRoberts, G. W., & Sithole, N. N. (1988). The phonological basis of perceptual loss for nonnative contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*. https://doi.org/345-360.10.1037/0096-1523.14.3.345

Best, C. T. & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In Murray J. M. & Ocke-Schwen B. (Ed.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13-34). John Benjamin. https://doi.org/10.1075/lllt.17.07bes?locatt=mode:legacy

Braun, B., Galts, T., & Kabak, B. (2014). Lexical encoding of L2 tones: The role of L1 stress, pitch accent and intonation. *Second Language Research*, *30*(3), 323-350. https://doi.org/10.1177/0267658313510926

Chen, J., Best, C. T. & Antoniou, M., (2023). Phonological and phonetic contributions to Thai-naïve Mandarin and Vietnamese speakers' imitation of Thai lexical tones: Effects of memory load and stimulus variability. *Laboratory Phonology, 14*(1). https://doi.org/10.16995/labphon.6435

Chen, Y., Lion, Y. & Nakayama, M. (2025). The structure of bilingual memory: What have we learned over the past 30 years?. *SSRN eJournal.* 1-31. http://dx.doi.org/10.2139/ssrn.5133234

Chow, C., Liu, Y., & Ning, J. H. (2018). The perception of Mandarin tones by Thai and Indonesian speakers. *Proceeding of 6th International Symposium on Tonal Aspects of Languages (TAL 2018)*, 197-201. https://doi.org/10.21437/TAL.2018-40

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representation in second language listening. *Journal of Phonetics, 34*(2), 269-284. https://doi.org/10.1016/j.wocn.2005.06.002

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*(4), 496–518. https://doi.org/10.1006/jmla.1999.2654

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Duanmu, S. (2007). *The phonology of Standard Chinese* (2nd ed.). Oxford

University Press. https://doi.org/10.1093/oso/9780199215782.001.0001

Duyck W. (2005). Translation and associative priming with cross-lingual pseudohomophones: evidence for nonselective phonological activation in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(6), 1340–1359. https://doi.org/10.1037/0278-7393.31.6.1340

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.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interaction. *Laboratory Phonology, 9*, 353-382.

Frances, C., Navarra-Barindelli, E., & Martin, C. D. (2021). Inhibitory and facilitatory effects of phonological and orthographic similarity on L2 word recognition across modalities in bilinguals. *Scientific Reports, 11*, 1-11. https://doi.org/10.1038/s41598-021-92259-z

Gandour, J. (1981). Perceptual dimensions of tone: Evidence from Cantonese. *Journal of Chinese Linguistics, 9*(1), 20–36. https://www.jstor.org/stable/23753516

Haigh, C. A., & Jared, D. (2007). The activation of phonological representations by bilinguals while reading silently: Evidence from interlingual homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(4), 623–644. https://doi.org/10.1037/0278-7393.33.4.623

Hao, Y. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics, 40*(2), 269-279. https://doi.org/10.1016/j.wocn.2011.11.001

Heung, T. (2001, January). The interplay of perception and phonology in tone 3 sandhi in Chinese Putonghua (Ohio State University Working Papers in Linguistics No. 55). Ohio State University. https://core.ac.uk/reader/159585980

Jitwirayanon, S. (2012). การปรับค่าความถี่มูลฐานโดยการแปลงค่าเฮิรตซ์เป็นเซมิโทน: แนวทางในการเสนอผลการวิเคราะห์วรรณยุกต์ [Fundamental frequency normalization by converting hertz to semitone: Lexical tone analysis presentation guideline]. *The Journal: Journal of the Faculty of Arts, 8*(2), 19-45.

Jitwirayanon, S. (2016). Pitch measurement 11 point [Praat script].

Ju, M., & Luce, P. A. (2004). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science, 15*(5), 314-318. https://doi.org/10.1111/j.0956-7976.2004.00675.x

Ladefoged, P. (2001). *A course in phonetics*. Harcourt College Publishers.

Lagrou, E., Hartsuiker, R.J., & Duyck, W. (2011). Knowledge of a second language influences auditory word recognition in the native language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(4), 952–965. https://doi.org/10.1037/a0023217

Lagrou, E., Hartsuiker, R.J., & Duyck, W. (2013). The influence of sentence context and accented speech on lexical access in second-language auditory word recognition. *Bilingualism: Language and Cognition, 16*(03), 508–517. https://doi.org/10.1017/S1366728912000508

Lee, Y. S., Vakoch, D. A., & Wurm, L. H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research, 25*(5), 527–542. https://doi.org/10.1007/BF01758181

Lee C. Y. (2007). Does horse activate mother? Processing lexical tone in form priming. *Language and Speech, 50*(Pt 1), 101–123. https://doi.org/10.1177/00238309070500010501

Leelapornpinit, S. (2016). A comparative study of Thai and Chinese phonology for the use of basic Thai language teaching as a foreign language. *Suthiparithat, 30*(93). 33-46.

Lemhöfer, K., Dijkstra, T., & Michel, M. C. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes, 19*(5), 585–611. https://doi.org/10.1080/01690960444000007

Leung, A. S. (2008). Tonal assimilation patterns of Cantonese L2 speakers of Mandarin in the perception and production of Mandarin tones. *Proceedings of the 2008 CLA annual conference*. https://cla-acl.ca/pdfs/actes-2008/CLA2008_Leung.pdf

Li, Y. (2016). English and Thai speakers' perception of Mandarin tones. *English Language Teaching, 9*(1), 122-132. https://doi.org/10.5539/elt.v9n1p122

Liu, S., & Samuel, A.G. (2004). Perception of Mandarin lexical tones when f0 information is neutralized. *Language and Speech, 47*, 109-138. https://doi.org/10.1177/00238309040470020101

Malins, J. G., & Joanisse, M. F. (2010). The roles of tonal and segmental information in Mandarin spoken word recognition: An eyetracking study. *Journal of Memory and Language, 62*, 407-420. https://doi.org/10.1016/j.jml.2010.02.004.

Marslen-Wilson, W. (1993). Issues of process and representation in lexical access. In G. Altmann & R. Shillcock (Eds.), *Cognitive models of language* *processes: Second sperlonga meeting* (pp. 187-210). Lawrence Erlbaum Associates Publishers.

Mirman, D. (2016). Zones of proximal development for models of spoken word recognition. In T. Harley (Series Ed.) & G.M. Gaskell & J. Mirkovic (Vol. Eds.), *Current issues in the psychology of language: Speech perception and spoken word recognition* (pp. 97-115). Psychology Press.

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

Nolan, F. (2003). Intonational equivalence: An experimental evaluation of pitch scales. In M. J. Solé, D. Recasens, & J. Romero (Eds.), Proceedings of the 15th international congress of phonetic sciences (pp. 771–774). Universitat Autònoma de Barcelona.

Norris, D., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(5), 1209–1228. https://doi.org/10.1037/0278-7393.21.5.1209

Ortega-Llebaria, M., Nemoga, M., & Presson, N. (2015). Long-term experience with a tonal language shapes the perception of intonation in English words: How Chinese-English bilinguals perceive "Rose?" vs. "Rose". *Bilingualism: Language and Cognition. 20*(2), 367-383. https://doi.org/10.1017/S1366728915000723

Persici, V., Vihman, M., Burro, R., & Majorano, M. (2019). Lexical access and competition in bilingual children: The role of proficiency and the lexical similarity of the two languages. *Journal of Experimental Child Psychology, 179*, 103–125. https://doi.org/10.1016/j.jecp.2018.10.002

Reid, A., Burnham, D., Kasisopa, B., Reilly, R., Attina, V., Rattanasone, N. X., & Best, C. T. (2015). Perceptual assimilation of lexical tone: The roles of language experience and visual information. *Attention, Perception, & Psychophysics, 77*(2), 571–591. https://doi.org/10.3758/s13414-014-0791-3

Rose, P. (1987). Considerations in the normalisation of the fundamental frequency of linguistic tone. *Speech Communication, 6*(4), 343-352. https://doi.org/10.1016/0167-6393(87)90009-4

Rungruang, A., & Mu, Y. (2017) Mandarin Chinese tonal acquisition by Thai speakers. *Asian Social Science, 13*, 107-115. https://doi.org/10.5539/ass.v13n5p107

Schulpen, B., Dijkstra, T., Schriefers, H. J., & Hasper, M. (2003). Recognition of interlingual homophones in bilingual auditory word recognition. *Journal of Experimental Psychology Human Perception and*

Performance, 29(6), 1155–1178. https://doi.org/10.1037/0096-1523.29.6.1155

Shook, A., & Marian, V. (2013). The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition, 16*(2), 304–324. https://doi.org/10.1017/S1366728912000466

Shook, A. & Marian, V. (2016). The influence of native-language tones on lexical access in the second language. *Acoustical Society of America, 136*(6). 3102-3109. https://doi.org/10.1121/1.4953692

Shuai, L., & Malins, J. G. (2017). Encoding lexical tones in jTRACE: A simulation of monosyllabic spoken word recognition in Mandarin Chinese. *Behavior Research Methods, 49*(1), 230–241. https://doi.org/10.3758/s13428-015-0690-0

Singh, L., & Fu, C. S. (2016). A new view of language development: The acquisition of lexical tone. *Child Development*, *87*(3), 834–854. https://doi.org/10.1111/cdev.12512

So, C. K., & Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: effects of native phonological and phonetic influences. *Language and Speech, 53*(2), 273–293. https://doi.org/10.1177/0023830909357156

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10*(3), 281–284. https://doi.org/10.1111/1467-9280.00151

Tao, R., Zhang, K., & Peng. (2021). Music does not facilitate lexical tone normalization: A speech-specific perceptual process. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.717110

Teeranon, P. (2007). The change of Standard Thai high tone: An acoustic study and a perceptual experiment. *SKASE Journal of Theoretical Linguistics [online] 4*(3). http://www.skase.sk/Volumes/JTL10/pdf_doc/1.pdf

Teeranon, P., & Rungrojsuwan, R. (2009). Change in the standard Thai high tone: An acoustic study, *Manusya Journal of Humanities, 12* (17), 34-44. http://www.manusya.journals.chula.ac.th/files/essay/Phanintra_p.34-44.pdf

Tingsabadh, M. R. K., & Abramson, A. S. (1993). Thai. *Journal of the International Phonetic Association*, *23*(1), 24–28. https://doi.org/10.1017/S0025100300004746

Tsukada, K. (2019). Are Asian language speakers similar or different? the perception of mandarin lexical tones by naïve listeners from tonal language backgrounds: A preliminary comparison of Thai and Vietnamese listeners. *Australian Journal of Linguistics, 39*(3), 329–346. https://doi.org/10.1080/07268602.2019.1620681 van Assche, E., Brysbaert, M., & Duyck, W. (2020). Bilingual lexical access. In R. R. Heredia & A. B. Cieślicka (Eds.), *Bilingual lexical ambiguity resolution* (pp. 42–67). Cambridge University Press. https://doi.org/10.1017/9781316535967.004

van Hell, J. G., & Tanner, D. (2012). Second language proficiency and cross-language lexical activation. *Language Learning, 62*(Suppl 2), 148–171. https://doi.org/10.1111/j.1467-9922.2012.00710.x

Wang, X., Hui, B., & Chen, S. (2020). Language selective or non-selective in bilingual lexical access? It depends on lexical tones!. *PloS One, 15*(3), e0230412. https://doi.org/10.1371/journal.pone.0230412

Wang, X. (2021). Beyond segments: Towards a lexical model for tonal bilinguals. *Journal of Second Language Studies, 4*(2), 245-267. https://doi.org/10.1075/jsls.21011.wan

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*(1), 1–25. https://doi.org/10.1016/S0749-596X(03)00105-0

Wu, J., Chen, Y., van Heuven, V. J., & Schiller, N. O. (2014). Tonal variability in lexical access. *Language, Cognition and Neuroscience, 29*(10), 1317–1324. https://doi.org/10.1080/23273798.2014.915977

Wu, X., Munro, M. J., & Wang, Y. (2014). Tone assimilation by Mandarin and Thai listeners with and without L2 experience. *Journal of Phonetics, 46*, 86-100. https://doi.org/10.1016/j.wocn.2014.06.005

Yip, M. (2002). *Tone*. Cambridge University Press. https://doi.org/10.1017/CBO9781139164559

Zhou, H., Chen, B., Yang, M., & Dunlap, S. (2010). Language nonselective access to phonological representations: Evidence from Chinese-English bilinguals. *Quarterly Journal of Experimental Psychology (2006), 63*(10), 2051–2066. https://doi.org/10.1080/17470211003718705

Zhu, M., Chen, F., Chen, X., & Yang, Y. (2023). The more the better? Effects of L1 tonal density and typology on the perception of non-native tones. *PLoS One, 18*(9), e0291828. https://doi.org/10.1371/journal.pone.0291828