Linguistic Prominence of Pitch within the Native Language Determines Accuracy of Tone Processing

Vance Schaefer and Isabelle Darcy
Indiana University

1. Introduction

1.1. Pitch prominence

We can distinguish four groups of languages that differ in the use they make of pitch (or fundamental frequency, abbreviated as f0, i.e., the height of the voice) to distinguish words (ranging from exclusive use to none). We ask whether using pitch to distinguish words in a person's first language (L1), either as tone, stress or pitch-accent, would help perceive tonal contrasts in an unknown language such as Thai. We investigate this question using one methodology to compare four different L1s varying typologically in the usage of pitch prominence (defined as the degree to which pitch is used to distinguish lexical items in the L1, ranging from exclusive to none). This study specifically examines how Thai tones, i.e., the use of pitch to distinguish words, are perceived by naïve listeners (= non-learners of Thai) who speak another tone language (Mandarin Chinese), a pitch accent language (Japanese[1]), a stress language (English), and a language which does not employ pitch in any way to distinguish the meaning of words (Korean[2]). Additionally, Thai was adopted as it features three level and two contour tones, allowing for a better understanding of tone height and direction as compared to a language like Mandarin Chinese which features only one level tone, but three contour tones, allowing for a comparison of performance on specific types of tonal contrasts.

The current theoretical framework of second language (L2) phonology models tends to largely focus on the non-native perception of segmentals, i.e., vowels and consonants (Perceptual Assimilation Model (PAM): Best, 1995; Speech Learning Model (SLM): Flege, 1995). Such research may be applicable as well to the perception of suprasegmentals (i.e., tone, pitch accent, stress, intonation) in that it may follow the same mechanisms that have been put forward for segmentals. For example, the classic case of the L1 shaping perception of non-native segmentals is the /l/~/r/ distinction in English for L1 speakers of Japanese. The lack of a contrast between /l/ and /r/ in their L1 causes Japanese listeners to associate either to the flap in Japanese in perception despite being able to produce the two sounds or to distinguish the two in isolated cases at the level of L1 English speakers (Goto, 1971; Miyawaki, Strange, Verbrugge, Liberman, Jenkins, & Fujimura, 1975; Sheldon & Strange, 1982). Yet, we cannot automatically infer from the perception of segmentals how the perception of suprasegmentals works. That is, the application of perception theory to suprasegmental domains is not straightforward, mostly because linguistic pitch is present in all the L1s in at least the form of intonation, if not to contrast words.

---

* The authors wish to thank the many people who have contributed to the conceptualization of this study and paper with their guidance, comments, and encouragement. We thank Professors Kathleen Bardovi-Harlig, Ken de Jong, Laurent Dekydtspotter, Mariko Kondo, Keiko Kuriyama, Philip LeSourd, Charles Lin, Öner Özçelik, Rex Sprouse, and David Stringer. We are also indebted to Stephanie Dickinson at the Indiana Statistical Consulting Center for help with the statistics. We also thank members of the Second Language Psycholinguistics Lab, audience at SLRF and LabPhon 13, two anonymous reviewers, Thai stimuli recorders and all participants.

[1] While standard Japanese is considered a pitch-accent language, there are dialects of Japanese such as the Fukushima and Kumamoto dialects that are considered to be pitch-accentless. Additionally, other pitch-accent dialects vary in their pitch-accent systems from the Japanese standard.

[2] While standard Korean is considered to be a pitch-accentless language (Kim-Renaud, 2009), we note that the Kyungsang dialect spoken in the region surrounding South Korea's second largest city of Busan, is a pitch-accent dialect. Additionally, we are aware of the Silva study (2006) showing that lexical contrastive usage of pitch may be appearing among younger speakers of the Seoul dialect.

 Selected Proceedings of the 2012 Second Language Research Forum, ed. Ryan T. Miller et al., 1-14. Somerville, MA: Cascadilla Proceedings Project.

Research points to similarities between the segmental and suprasegmental domains in second language phonological processing: suprasegmentals such as word stress have been shown to be difficult to perceive for speakers who do not use this dimension in a lexically contrastive manner (Dupoux, Pallier, Sebastián-Gallés, & Mehler, 1997). This phenomenon extends to the domain of L2 tone in the case of L1 speakers of a non-tonal language (Burnham, Kirkwood, Luksaneeyanawin, & Pansottee, 1992). Similarly, previous studies show that speakers of various L1s differ in the accuracy with which they identify non-native tones due to their varying ability to attend to pitch height and/or direction possibly because of the influence of their L1 (Francis, Ciocca, Ma, & Fenn 2008; Gandour, 1983). L1 tone speakers are better able to attend to these cues and/or to map L2 tones onto their L1 tones and, therefore, outperform speakers of non-tone languages in tone identification tasks (Wayland & Guion, 2004). In contrast, speakers of non-tonal languages have less sensitivity to tonal contrasts than people with previous tonal experience (Hallé, Chang, & Best, 2004, for French listeners; Gandour & Harshman, 1978; Wang, Behne, Jongman, & Sereno, 2004, among others).

However, given that languages differ to the extent and function to which they use f0 variations, (i.e., only some use pitch for distinctions at the word or lexical level while all languages use pitch for intonation at the phrasal or non-lexical level), it is possible that there are differences among non-tonal language speakers in non-native tone perception (Burnham, Francis, Webster, Luksaneeyanawin, Attapaiboon, Lacerda, & Keller, 1996; So, 2006). For example, some studies report that pitch accent language speakers (L1 Japanese) perform at comparable rates to L1 tone language speakers in their perception of L2 tones (Burnham et al., 1996; So, 2006). Attempts to model the mechanisms behind these effects in the same way as for segmentals (i.e., considering both tones and pitch accent to be categories that can be mapped onto each other) suggest that Japanese listeners assimilate Mandarin tones onto "Japanese pitch accent categories" (So, 2010). Also, McGinnis (1997) shows that Japanese speakers show greater improvement compared to L1 English speakers when learning Mandarin tones, possibly suggesting that L1 English speakers have weaker categories or no categories to which they can assimilate tones.

Any attempt to apply segmental models such as the PAM or the SLM (Best, 1995; Flege, 1995) to suprasegmental dimensions therefore faces the challenge of conceiving suprasegmental dimensions as categories in the same way as segmental categories. Indeed, native speakers perceive tones as linguistic categories (Van Lancker & Fromkin, 1973; Wang, Jongman, & Sereno, 2001; Hallé et al., 2004; Francis et al., 2008 for tone) so that tonal information also constrains lexical access (Lee, 2007) and could be referred to as "tonemes," parallel to phonemes. Japanese pitch accents are considered categories by So and Best (2010), but their status as categories is not fully clear yet. Lexical stress, on the other hand, is usually not looked at in terms of categories. In the domain of intonational patterns, several attempts at establishing their categorical nature indeed suggest that some (e.g., contrastive vs. neutral focus) are perceived categorically (Feldhausen, Pešková, Kireva, & Gabriel, 2011). Predictions made by these models presuppose that assimilation (or equivalence classification) applies to both segmental and suprasegmentals in the same way. In short, if lexically-contrastive pitch is not used in the L1 to the same extent, it might be perceived as very dissimilar and acquisition is predicted to be easier (see Dupoux, Sebastián-Gallés, Navarrete, & Peperkamp, 2008, for an illustration of these predictions). Similarly, tone language speakers might experience "perceptual assimilation" from one non-native tone category to their native tone category. Recent evidence suggests that this is the case (So & Best, 2010). In sum, the PAM framework makes clear predictions about assimilation patterns between tone languages, but the predictions for other kinds of suprasegmental mappings are not clear (for instance between a tone and a pitch accent language), and crucially hinge on defining these dimensions as categories. In addition, these predictions are complicated by the possibility that intonational categories can also be represented in this way. Attempts at specifically developing assimilation maps between intonation and tones are underway (White, 1981; Hao, 2008).

On the other hand, the predictions made by the Feature Hypothesis (McAllister, Flege, & Piske, 2002) are more straightforwardly applicable to segmental and suprasegmental dimensions alike. Specifically, the Feature Hypothesis predicts that the more a certain phonetic or phonological dimension is prominent in the L1, the easier it might be to learn to discern and use that dimension for L2 phonological processing (see also Dupoux et al., 2008). This hypothesis is supported by data from the L2 acquisition of lexically-contrastive vowel length in Swedish. In a mispronunciation detection task for Swedish words, L1 speakers of Estonian, a language which distinguishes lexically-contrastive vowel length, outperformed L1 speakers of both English (in which vowel length exists as a correlate of word stress) and Spanish (which does not use vowel length lexically). In turn, the L1 English speakers outperformed the Spanish due to the “intermediate” presence of the same lexical feature. The predictions made by the Feature Hypothesis however, have not yet been tested for non-native tone perception.

Therefore, rather than the existence of pitch categories in the L1, we might consider the prominence of lexically-contrastive pitch in the L1, and the resulting weighting of possible tone features such as pitch height and direction in the naïve perception of lexically-contrastive pitch such as tone (McAllister et al., 2002).

2. Tone features

2.1. Native tone perception

The two features thought to play the largest role in both L1 and L2 perception of tone are pitch height and pitch direction (Gandour, 1983; Wang, Jongman, Sereno, 2006; Vũ, 1981). In terms of pitch height, listeners evaluate whether the tone is high or low and in some tone languages in the middle area of the voice range. Even though height is relative and is influenced by gender and voice quality, Lee (2009) shows that native speakers of Mandarin are able to rapidly and accurately identify f0 height by gauging the voice range of multispeaker stimuli from both males and females with no previous exposure to the speakers and only the beginning of the syllable. In terms of pitch direction, listeners detect whether the tone moves up or down or both or whether it remains relatively flat. Thus, tone perception can be measured in terms of perception of pitch height and/or direction.

Similarities between tone patterns may however cause confusion (Leather, 1990) as in the case of tones with the same directional pattern (i.e., both flat or both rising) or in the same register (i.e., a low flat tone and low-to-mid rising tone). Thus, some tones are easier to perceive than others for both native and non-native speakers of the target tone language (Abramson, 1975, 1978; Burnham et al., 1992). When comparing Thai tones in an AX task, L1 Thai listeners have the most difficulty comparing flat vs. contour tones while the comparison of contour tones with one another is the easiest (Burnham et al., 1992), suggesting that the perception of pitch height, specifically the “absolute initial pitch of the component tones” (p. 555), is the most salient for L1 Thai speakers when perceiving contrasting tones (using stimuli recorded in only one female voice). Yet, L1 Thai speakers frequently confuse the mid and low tones on a tone identification test using many voices (i.e., 10 different voices) but not on the same test using one voice (Abramson, 1976), hinting that L1 Thai speakers require more than pitch height or absolute initial pitch when discriminating tones. In a parallel situation for Mandarin, L1 Mandarin speakers can generally place tones 1, 4 and tones 2, 3 into either the high or low registers respectively but confuse tones sharing similar heights (Lee, Tao, & Bond, 2008).

Across languages, listeners weigh pitch height more than pitch direction. However, the L1 influences the relative weight given to these two features. Gandour (1983) compared listeners of Thai, Mandarin, Cantonese, Taiwanese (and English) on the perception of 19 different synthesized tones, using one voice, reflecting pitch patterns found in tone languages (i.e., 5 level tones, 14 contour tones – 8 unidirectional tones of falling or rising tones and 6 bidirectional tones combining falling with rising tones – where beginning and end points varied within each subset). Participants were asked to differentiate pairs on an 11-point scale (*no difference* to *extreme difference*), after hearing the stimulus set twice. The findings show that every group used height more than direction to judge tone dissimilarity, even though tone language speakers were found to use direction to a larger extent than non-tone-language speakers (English). Variance in perceived differences was attributed to the two features of height and direction. This prominence can be ascribed to the fact that only pitch height allows listeners to discriminate between some tones (Tuc, 2003) such as the mid and high tone in Thai. The relative difference in weighting of pitch height and direction is thus shaped by the L1, but may also depend on task-specific variables such as the number of voices used for the stimuli.

Additional features have been put forth to explain how native speakers perceive various tones, including average f0/pitch height; f0 contour/pitch direction, f0 slope/magnitude of pitch change, extreme endpoints (Gandour, 1983), tone duration (Leather, 1990; Gandour, 1983), turning point (Shen, Lin, & Yan, 1993), peak alignment (Zsiga & Nitisaroj, 2007), and amplitude (Whalen & Xu, 1992). This study, however, focuses on the features of pitch height and pitch direction as these appear to be the most salient in tone perception for Mandarin, Cantonese, (Francis et al., 2008; Gandour, 1983), Taiwanese (Gandour, 1983), Thai (Gandour & Harshman, 1978; Gandour, 1983), and Vietnamese (Vũ, 1981, cited in Tuc, 2003).

In sum, we see that pitch height appears to be the most salient feature of tone for native speakers of tone or non-tonal languages (Gandour, 1983), but pitch direction also appears to be necessary to discriminate tones which have the same initial pitch level (Abramson, 1976) or occur in the same register (i.e., upper or lower range of the voice). When tones have similar pitch heights and directions, they are easily confusable (e.g., Mandarin high-rising and low-dipping tones), especially when the task uses several voices.

2.2. Non-native tone perception

We organize this section according to the four language types described in the introduction.

Tone languages. L1 tone language speakers transfer their L1 tone patterns onto non-native tones or their ability from experience to track the pitch direction and/or height (Wayland & Guion, 2004). Yet, L1 tone language speakers also confuse non-native tones which are similar but not the same as their L1 tones (Gandour, 1983; Wayland & Guion, 2004). In short, having tones in the L1 does not necessarily aid in the perception of L2 tones as they may also impede perception of similar but sufficiently different tones (So & Best, 2010).

Pitch-accent languages. The few studies to have looked at the perception of non-native tone by L1 pitch accent speakers have found that pitch accent speakers perform at comparable accuracy levels to L1 speakers of tone languages (Burnham et al., 1996; So, 2006). L1 speakers of pitch accent languages also outperform L1 English speakers in perception accuracy (Burnham et al., 1996) or show greater improvement when learning Mandarin tones (McGinnis, 1996). Their performance appears to hinge on their varying ability to map tones onto pitch accent categories (So, 2010; So & Best, 2010).

Stress-accent languages. Although English does not use lexical tone, we must consider the possibility that L1 English speakers might transfer the pitch used in stress, although they might experience difficulty in extracting pitch only from the other features of English stress (i.e., vowel duration and intensity) as hinted to by their inability to do so in the production of L2 Japanese, for which pitch is the only acoustic cue (Kondo, 2007). Thus, English speakers might not transfer contrastive lexical pitch from their L1, but they do seem to transfer intonation patterns (e.g., question intonation to a rising tone in Mandarin, Francis et al., 2008, see below). L1 English listeners tend to confuse similar tones in a comparable manner to L1 speakers of a tone language (Burnham et al., 1992; Leather, 1983, 1990; Wang, Spence, Jongman & Sereno, 1999).

Languages without lexically-contrastive pitch. Speakers of languages which do not feature lexical pitch are not "deaf" to non-native tones, but such listeners do not perceive non-native tone categorically and therefore, have difficulty mapping tone onto any L1 French category (Hallé et al., 2004). This is in line with findings that L1 speakers of a word-stress language are more sensitive to small acoustic differences in pitch than L1 tone language speakers as only the latter group perceives tone categorically (Stagray & Downs, 1993; Wang, 1976).

Citing Beckman, Hirschberg, & Shattuck-Hufnagel (2005), Francis et al. (2008) note (in response to the Hallé et al., 2004 study) that intonational contours can be considered linguistic categories stored in long-term memory similar to categories for vowels, consonants, and tones, and so, should have the same influence on the perception of non-native tones, influencing in turn the weighting of certain features. Thus, speakers of languages which do not feature lexically-contrastive pitch should be able to access their mental representations of intonational categories when perceiving non-native tones, resulting in their ability to map non-native tones (e.g., Mandarin tone 2 [35] and tone 4 [51]) onto intonational categories for questions and statements, respectively (see Francis et al., 2008; So & Best, 2010).

Thus, naïve listeners who encounter a totally new sound system which utilizes pitch in a new manner do not approach the task without any "tools" but most likely employ the pitch patterns used in their L1. Accordingly, L1 linguistic pitch shapes perception of non-native tone (Wang et al., 2006). Hence, we could arrange the languages used in this study according to language types on a functional scale, such as one proposed by Van Lancker in 1980. In Figure 1, we see that the way a language uses pitch contrasts vary in domain (from small to large), and also in the function that the pitch will fulfill (from tonal contrasts to marking e.g., focus structure or affect).

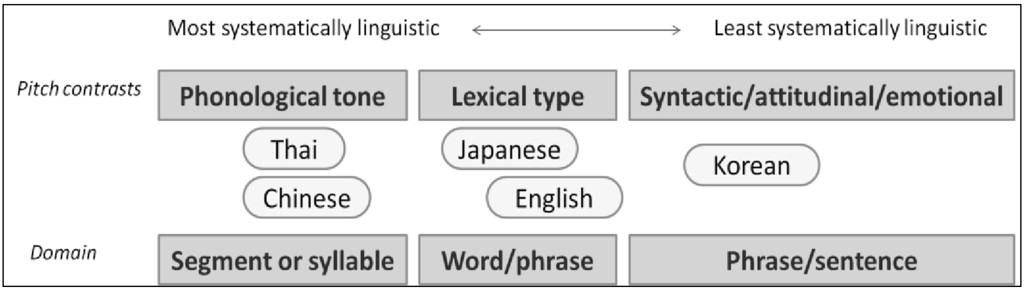

Figure 1. Functional scale of pitch contrasts (Adapted from Van Lancker, 1980: 210)

3. The present study

This study tests predictions derived from the Feature Hypothesis for non-native tone perception. We employ one methodology in order to ascertain how naïve listeners of a wide range of languages with varying levels of prominence and inventory of contrastive linguistic pitch patterns perceive the tones of a non-native tone language. This goal represents a first step toward better understanding the L2 acquisition of tone. Such an approach provides valuable information regarding what non-learners do, allowing us to theorize how learners acquire L2 tone. Additionally, a comparative approach of languages differing in the linguistic use of pitch will more comprehensively illuminate the perception of suprasegmentals by pinpointing whether the L1 inventory (i.e., type of pitch patterns) and the prominence of certain pitch features in the L1 influences the weighting and thereby, the non-native perception of certain tonal features.

This leads us to this study's first research question. Does pitch prominence shape tone perception? Prominence here is based on the function of linguistic pitch to signal lexical contrast (cf. Feature Hypothesis, McAllister et al., 2002). Thus, our working hypothesis is that the greater the prominence of lexical pitch in the L1, the better the perception of non-native tones, resulting in a hierarchy of performance among the various L1s, relative to the prominence of lexical pitch in the L1. Accordingly, we predict accuracy of cross-language speech perception for tones based on prominence of pitch in the L1. That is, we combine the idea of functional use of pitch with the prominence idea.

Specific predictions for each language are presented in Table 1. We organize the languages we use in this study according to the role of pitch for each. The three languages of Mandarin, Japanese, and English use pitch to distinguish words (even though the domain at which they do so varies) whereas Korean uses pitch in a domain larger than a word and does not use pitch to contrast words. Thus, we can establish a scale of prominence going from maximal to low. We also have an additional distinction between English and Japanese based on whether or not pitch can be exclusively used to distinguish words. As a result, we predict that if the prominence of pitch to distinguish words in your first language is high, then your accuracy in tone perception will also be high, and conversely, if it is low, then accuracy will be lower as well.

Table 1. Overview of our pitch prominence typology and predictions for tone perception accuracy

Pitch pattern	Prosodic Domain	Lexical status of pitch use	Prominence for lexical distinction?	Predicted Sensitivity/ Accuracy in tone perception
Tone (Mandarin)	Syllable	Lexical	Maximal	Highest
Pitch-Accent (Japanese)	Word	Lexical	High-Intermediate (pitch is exclusive)	High-intermediate
Word Stress (English)	Word/Foot	Lexical	Low-intermediate (pitch is non-exclusive)	Low-intermediate
Intonation (Korean)	Intonational phrase, PPh	Non-lexical	Low	Lowest

Additionally, our second research question asks whether specific tone contrasts also influence accuracy. That is, do listeners use the feature of height more than direction, as indicated by previous research? Does this prevalence interact with the L1? We know, for example, that the exact type of tonal contrasts examined can influence performance. For example, English speakers process dynamic (contour) f0 variations more accurately than static f0 differences in disyllabic stimuli (Repp & Lin, 1990; Wood, 1974; Lee & Nusbaum, 1993). Also, Dutch listeners attend to f0 information when these correspond to contours having linguistic meaning in Dutch (disyllabic stimuli, question intonation) (Braun & Johnson, 2011). As a result, we ask whether performance in non-native tone perception is affected by specific tonal shapes. Our working hypothesis is that listeners will use height more than direction (e.g., Gandour, 1983), but this could be affected by the fact that we use multiple voices for the stimuli (see below).

4. Methodology

4.1. Participants

Forty-seven participants were recruited from five language groups: Mandarin (n=10; females=6), Japanese (n=12, females=11), English (n=13; females=10), Korean (n=10; females=7) and Thai (n=2; males=2). The Thai speakers were recruited to ensure that the stimuli and AXB task itself were valid for native speakers. The participants were primarily graduate students or former graduate students who were involved in language studies (i.e., language education, linguistics, applied linguistics) with the exception of 11 participants who were undergraduate students (n=3) or not involved in language studies (n=8) (i.e., Mandarin=3, Japanese=3, English=1, Korean=3, Thai=1). However, four individuals were cut from the final analysis as they had significant exposure to one of the other target languages in the study or differed in background from the target group, resulting in a reduction from 47 to 43 participants in total. Three English speakers had exposure to Japanese or a tone language (i.e., Mandarin or Vietnamese). One female Japanese student was an ESL student with lower exposure and proficiency in English as compared to the graduate student participants. As a result, only 11 Japanese-speaking participants and 10 English-speaking participants' data were analyzed. Average ages were 27.1 years for Mandarin speakers (range 24-31), 35.4 years for Japanese (range 25-50), 31 years for English (range 25-45) and 32.2 years for Korean speakers (27-47). The two Thai listeners were 25 and 32 years old. The average time spent in an English-speaking country was 3.5 years for the Mandarin speakers, 6.6 years for the Japanese, 4.5 years for the Koreans, and 2 years for the Thai speakers. The English speakers had spent an average of 1.7 years abroad in a non-English speaking environment.

The speakers of Mandarin included six speakers who also had various degrees of exposure to Taiwanese, another tonal language. Most had been exposed to another Chinese dialect even if they did not consider themselves a fluent speaker of that dialect. The speakers of Japanese, a pitch accent language, were recruited on the basis of speaking a dialect of Japanese which features pitch accent although not necessarily standard Japanese. Two speakers were from Tochigi and Ibaraki prefecture which are close to Fukushima prefecture, known for its accentless dialect. The English speakers were native speakers of American English who had no proficiency in Thai, Mandarin, Japanese, Korean or any tone language. Korean speakers were mainly from the Seoul area, but three were from the Kyungsang region where a pitch-accent dialect of Korean is spoken, and one speaker was from Cholla, an area abutting Kyungsang but with a dialect not featuring pitch accent (self-reported).

4.2. Materials

The test stimuli consisted of 16 open CVV syllables, with a long vowel (VV). Each syllable was recorded with each of the five different Thai tones, resulting in 80 items (41 items being real words and 39 nonwords). Control stimuli were CVV and CVC syllables (all were real Thai words) composed of vowels or consonants similar to those used for the test items. The control condition also included more difficult vowels such as [ɯ], [ɛ], and [ə]. The syllables were then arranged in triplets for the AXB design. In an AXB design, four trials are needed for each comparison: AAB, ABB, BAA, BBA. For example, the two tones Low (L) and Mid (M) would be paired as LLM, LMM, MLL, and MML. If the syllable carrying such a comparison were [bi:], a trial would look like the following: [bi:]^{L} – [bi:]^{L} – [bi:]^{M}.

The experiment contains two conditions, test and control, with 48 trials each. In the test condition, the syllables within one triplet only differ by tone; the segmental make-up of the syllables remains the same. In the control condition, all syllables in the triplet have the same tone but vary in either one consonant or one vowel. Furthermore, within the test condition, we included three subconditions in order to examine specific tonal comparisons: 1) Height, comparing flat tones, 2) Direction, comparing contour tones, and 3) Mixed, comparing flat tones with contour tones. Table 2 presents the overview of the conditions used in the study. Twelve triplets each for the direction and height condition, and 24 triplets in the mixed condition were created. All trials were randomized and put into 3 blocks of 32 items, respectively.

Table 2. Overview of the tonal comparison(s) and number of trials used for each condition

Test Conditions			Control Condition
Direction (n=12)	Height (n=12)	Mixed (n=24)	Control (n=48)
rising-falling	low-mid	low-rising low-falling	consonant
rising-falling	low-high	mid-rising mid-falling	vowel
rising-falling	mid-high	high-rising high-falling	

The AXB stimuli were recorded by two native Thai speakers. Both spoke the Central Thai dialect. Sixteen different words were recorded, with three tokens of each, spoken without a carrier phrase. Another recording of distracters was made, with two tokens of each item.

The female voice was used for the A and B while the male voice was used for the X. The interstimulus interval (ISI) between the A and X and between the X and B tokens was 500 ms. The experiment was timed so that after the presentation of each trial, participants had 3000 milliseconds to make their answer. Reaction times were measured from the onset of the X stimulus.

The warm-up phase of the task consisted of 16 trials with feedback indicating their accuracy and RT. The 16 tokens consisted of two trials comparing flat tones, two trials comparing contour tones, four comparing flat tones with contour tones, and eight control trials. None of these were used in the test phase of the task, which comprised 96 trials. All trials within each phase were randomly ordered.

4.3. Procedure

Participants were tested individually. For each trial, participants heard a triplet of syllables and chose whether the middle one (i.e., X) was more similar to the first (i.e., A) or the third one heard (i.e., B), by pressing two clearly identified keys on the computer keyboard. The task required from 15-20 minutes in total. The task was then followed by a debriefing session. The participants also filled out a questionnaire about their demographic and linguistic background.

5. Results

Accuracy rates and reaction times were obtained for each condition for each group and analyzed as follows. Reaction times shorter or longer than two standard deviations from the RT mean of each participant were replaced by the RT mean of each participant (5.2% of total RTs). Additionally, remaining reaction times lower than 300 ms were replaced by the mean RT of the participant (0.23% of total RTs). Data for three items in the height condition (one each for L vs. M, L vs. H, M vs. H) and for one item in the direction condition (R vs. F) were excluded from analysis as one Thai participant felt that the tones were not ideal models of the targeted tone. Means for individual participants and items were screened for outliers. No item or participant was excluded. Mean accuracy in each condition was computed for each group. Similarly, the mean reaction time for correct items was computed for each condition and each group. Linear mixed models to compare means were run in SPSS 20 by subjects on average accuracy comparing Condition (*Test* vs. *Control*) or Subconditions, and Language Groups, as well as interactions. Unlike the other language groups, the Thai listeners were able to approach the task using lexical knowledge. Therefore, analyses were run excluding this group. We used an alpha level of .05 for all statistical tests.

Overall accuracy rates in each condition are presented in Figure 2. Results show that overall accuracy in the test condition was slightly lower than on the control condition (74.8% vs. 79.1% correct). The analysis omitting Thai revealed a significant interaction ($F(3, 37) = 11.3$, $p < .001$) between group and condition. On the test condition, we observed that Mandarin participants outperform other non-native groups (87% accuracy), followed by Japanese participants (77%), and by English and Korean (both at 67% accuracy). Whereas accuracy of all groups was comparable on the control condition (no effect of “group” on the control condition: $F(3, 67.3) = 1.5$, $p > .1$), there was a significant effect of group on the test condition, as suggested by the difference in accuracy rates ($F(3, 67.3) = 11.3$, $p < .001$). Mandarin listeners discriminated tonal contrasts with higher accuracy than the other groups, significantly outperforming both Korean and English participants ($p < .001$) but only marginally more accurate than the Japanese group ($p = .093$). Notably, Korean and English participants were not significantly different from each other ($p = 1$) on the test condition, as is visible in Figure 2.

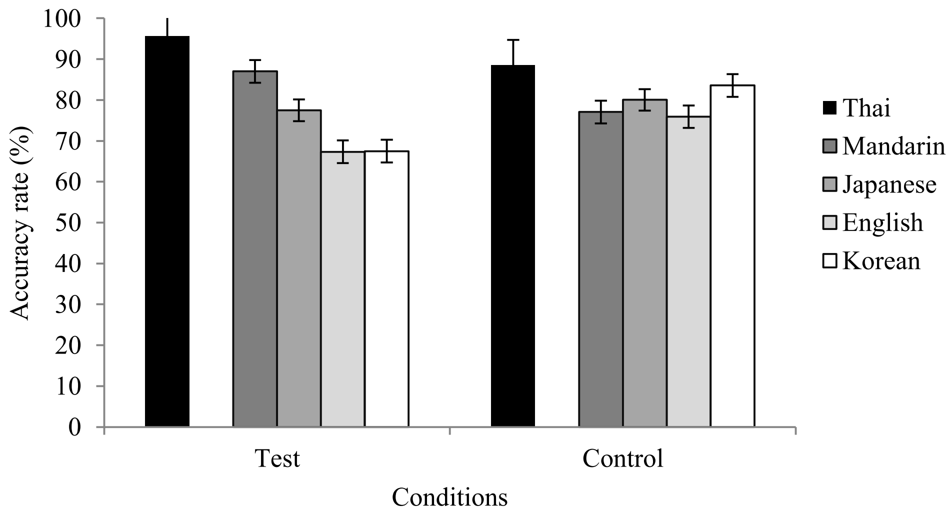

Figure 2. Accuracy rate (%) for each language group in the test vs. control condition. Error bars enclose +/- 1 SE. (Thai listeners are displayed in black for comparison purposes)

Reaction times on the test vs. control condition are displayed in Figure 3 below. There is a main effect of condition ($F(1, 37) = 31.4$, $p < .001$). Overall, latencies in the test condition are about 120 ms slower than on the control condition (1265 ms vs. 1141 ms). Thai and Mandarin listeners are faster than all the other non-native groups but the main effect of group was not significant ($F(3, 37) = 1.7$, $p > .1$). The interaction was not significant ($F(3, 37) = 2.4$, $p = .08$).

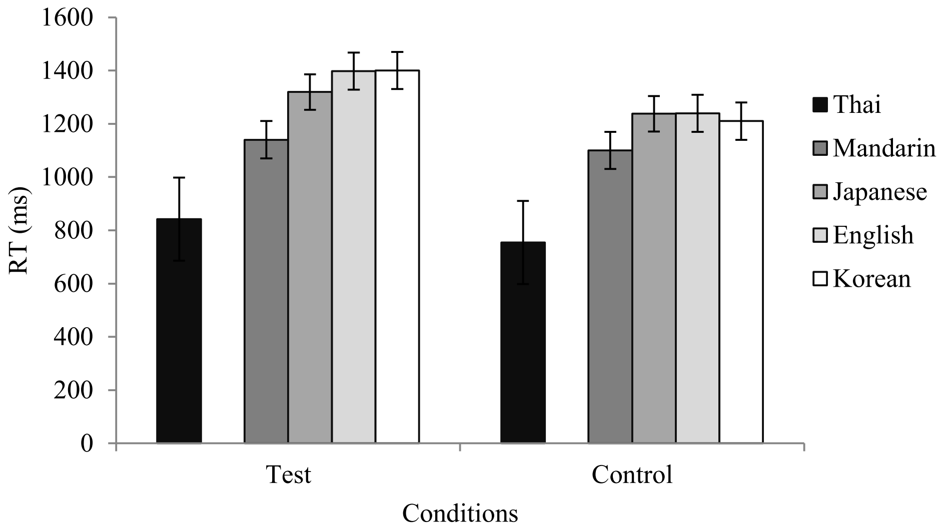

Figure 3. Reaction times (ms) for each language group in the test vs. control condition. Error bars enclose +/- 1 SE. (Thai listeners are displayed in black for comparison purposes)

The overall accuracy pattern that emerged from these accuracy and RT data confirms in large part the predicted hierarchy, according to which the functional prominence of pitch in the L1 determines accuracy in a phonological discrimination task. Against our prediction, however, the data also revealed that English and Korean participants pattern identically, perhaps suggesting that f0 information is less readily accessible for phonological discrimination in these two groups.

We turn now to the performance in each subcondition. Accuracy and RT data for each group and subcondition are shown in Table 3.

Table 3. Accuracy means (%) and reaction times (in ms) by language groups for each subcondition

	Accuracy				Reaction times			
Language group	height	direction	mixed	(SE)	height	direction	mixed	(SE)
Thai (n=2)	*88.9*	*100*	*97.9*	8.5	*887*	*846*	*800*	*164.8*
Mandarin (n=10)	76.7	97.3	87.1	3.8	1169	1041	1155	73.7
Japanese (n=11)	66.4	87.6	78.4	3.6	1303	1155	1347	70.3
English (n=10)	64.4	65.5	72.1	3.8	1288	1306	1389	73.7
Korean (n=10)	58.9	78.2	65.4	3.8	1370	1363	1301	73.7
Average (non-n.)	66.6	82.1	75.7		1282	1216	1298	

Note: SE = standard error; non-n. = non-native groups only

To make comparisons among the language groups, Linear Mixed Models with Repeated Measures for the four conditions within each subject were run. The dependent variables were accuracy rates and reaction times. The independent variables were language group (Thai, Mandarin, Japanese, English, Korean) and condition (height, direction, mixed).

For the analysis on accuracy rates (not including Thai listeners), the analysis of accuracy rates showed an interaction between condition and language group ($F(6, 74) = 2.7$, $p < .05$). Univariate tests for the simple effect of condition within each group show that condition significantly impacts accuracy rates for all groups except for the English participants (English: $F(2, 74) = 1.58$, $p > .1$; for the other three, $p < .001$).

Additionally, the analysis on RTs reveals a main effect of condition ($F(2, 74) = 3.8$, $p < .05$). Overall, the direction condition (e.g., raising-falling) is responded to faster in two out of four groups (average RT: height, 1282 ms; direction, 1216 ms; mixed, 1298 ms), significantly faster than the mixed condition. The main effect of group is not significant ($F(3, 37) = 2.1$, $p = .12$). The interaction between condition and group is marginal ($F(6, 74) = 2.0$, $p = .075$), indicating that groups display similar latencies to the same conditions. This and the lack of main effect of group is likely due to the fact that performance among groups is highly similar for the height and the mixed condition, especially for the Japanese, English and Korean groups. Univariate tests for the simple effect of group within each condition show that latencies differ by group only for the direction condition ($p < .05$), but not for the two other conditions ($p > .1$).

These results are suggestive of the predicted trends regarding overall sensitivity to tonal contrasts. They also indicate, against our working hypothesis, that the height condition is overall the most difficult, and the direction condition is the most accurate (and responded to fastest as well). The one exception to this pattern is found in the English group, for whom accuracy is the same in all conditions.

6. Discussion and Conclusions

Results strongly indicate that the functional prominence of lexically-contrastive pitch use in different L1s shapes the cross-linguistic perception of non-native tone, indicated by the effect of the L1. The L1 Mandarin group performed significantly more accurately than the English and Korean groups overall and also the Japanese (but not significantly). The Japanese were significantly more accurate than the Korean group overall and the English listeners in terms of raw scores, suggesting that a higher L1 pitch prominence aids in non-native tone perception. That is, we clearly see performance on the perception of non-native Thai tones varies as a function of the presence and prominence of L1 lexically-contrastive pitch in the L1. Thus, our predicted hierarchy of performance (of more accurate to less accurate) was generally confirmed as follows: L1 tone > L1 pitch accent > L1 stress = L1 without lexically-contrastive pitch. The finding that the word stress L1 English and the L1 which does not feature lexically-contrastive pitch Korean yielded comparable levels of accuracy was not predicted. We attribute this result to several possibilities. First, f0 is rarely used alone to distinguish words in English. This fact appears to yield the same performance in tone discrimination as if f0 was not used at all to signal lexical contrast (i.e., English=Korean). Additionally, the fact that f0 can be used exclusively to distinguish words in Japanese plays an important role as suggested by the different patterns obtained by the Japanese and English listeners. The flat performance by the English speakers is also consistent with findings showing that stress constrains lexical access only to a limited extent in English (Cooper, Cutler, & Wales, 2002). Another possible reason may be that Koreans are more accurate because of their exposure to and acquisition of L2 English word stress. Alternatively, the effect could be due to exposure to a pitch-accent dialect (cf., Sukekawa, Choi, Maekawa, & Sato, 1995). Both these facts require us to retest with a control group speaking only the Seoul dialect and who are also somewhat older, since lexically-contrastive pitch usage is appearing among younger speakers of the Seoul dialect (Silva, 2006). Globally however, our findings confirm previous results obtained across studies and add strength by allowing a direct comparison among four language groups with the same methodology. The degree of prominence of pitch to signal lexical contrast appears to determine accuracy on our tonal categorization task.

Concerning the research question as to whether different L1 language groups resort to different features in their L1 to perceive non-native tones (height or direction), we find the following results. Overall, most groups reflect a general trend that indicates a more robust tendency to rely on pitch direction, as observed by the fact that performance was less accurate on the height condition compared to other conditions. This result seems to contradict previous research that showed Japanese and English groups focusing more on pitch height (Guion & Pederson, 2007). This result also conflicts somewhat with the previous results showing the mixed condition was the most difficult followed by the height condition (Burnham et al., 1992), requiring this issue to be re-examined. However, it is possible that these results are in part explained by the fact that we used two voices with different gender (male and female), thus making the pitch height comparison more difficult than in other studies using only one voice, which while not confusing for L1 perception (Lee, 2009), may be for non-native, naïve perception. Another possible explanation could also be the fewer number of comparisons in the direction condition (only Rising=Falling) compared to height (Low=Mid, Low=High, Mid=High) or to mixed (see Bohn, 1995). Yet, this did not affect the English (who showed a flat performance in all conditions). One possible explanation for their performance is that the use of monosyllabic stimuli may have prevented them from applying intonational contours to tonal comparisons varying in tone contour, as suggested by similar findings obtained with disyllabic stimuli by Braun and Johnson (2011).

The current study derived specific predictions based on the prominence of a phonetic dimension (Feature Hypothesis, McAllister et al., 2002) for the naïve perception of tone. Our study established a baseline for tone perception focusing on the functional use of linguistic pitch in four language types. We propose the *Pitch Prominence Hypothesis* to model our data. Accordingly, the degree to which pitch differentiates lexical items in the L1 (i.e., lexical prominence) shapes the naïve (= non-learner) perception of a non-native lexically-contrastive pitch system, as in this case, of a non-native tone system. Unlike the Feature Hypothesis, the *Pitch Prominence Hypothesis* takes into account several constraints specific to lexically-contrastive pitch (see below), which go beyond defining prominence merely as “greater usage” or “degree of usage.” This is due to the fact that lexically-contrastive pitch is a complex phenomenon encompassing a large typological variety of lexically-contrastive pitch usages and exhibiting a wide range of pitch pattern inventories, perhaps analogous to vowel/consonant inventories, thus differing in nature from suprasegmentals such as vowel length covered in the Feature Hypothesis study by McAllister et al. (2002). The *Pitch Prominence Hypothesis* works within the Feature Hypothesis but allows for a more streamlined, comprehensive approach to guide future study of the phenomenon of L2 perception of lexically-contrastive pitch across typologically-diverse usages of lexically-contrastive pitch[3].

As for the first constraint, within this framework, the definition of pitch prominence in the L1 takes into consideration the prosodic domain of pitch contrasts. Indeed, our findings suggest that the specific prosodic domain in which pitch differentiates lexical items also constrains performance: domain overlap yields most accurate performance. For instance, Mandarin uses pitch to signal lexical contrast at the syllable domain, which is also the case for Thai, whereas in Japanese, the prosodic domain of pitch contrasts is rather the prosodic word. When domains do not overlap, it appears more difficult to map L1-pitch usage to the non-native pitch contrasts. Similarly, in the case of English, if word stress is not aiding tone perception, the question is whether intonational categories which typically require a phrasal domain can be mapped or associated with tonal contrasts in a smaller prosodic domain such as the syllable.

Three additional constraints should also be taken into account when defining pitch prominence: 1) *exclusivity* to signal lexical contrast, 2) *functional load*, and 3) *inventory of pitch patterns*. *Exclusivity* refers to whether lexically-contrastive pitch is used solely to differentiate words in the L1. For example, in Japanese this appears to be the case while word stress in English includes the other correlates of vowel length, spectral quality and intensity in addition to pitch. *Functional load* refers to the extent and/or number of minimal pairs differentiated in the L1. That is, tone languages such as Mandarin require pitch to distinguish a far larger number of lexical items in comparison to pitch accent (Pierrehumbert & Beckman, 1988) and word stress languages. Finally, *inventory of pitch patterns* refers to the number and type of patterns possible. For example, Mandarin has four tones, one level and three contour tones, while Thai has five tones, three level and two contour tones, implying a possible bias toward pitch height and/or direction. Furthermore, differences exist between tone types, i.e., the rising tones in Thai and Mandarin vary in shape.

Moreover, we posit that within the Pitch Prominence Hypothesis, tone-to-tone mapping may be applied where there is domain overlap, and where pitch categories are robust as seems to be the case for L1 tone language speakers. In this case, a model like PAM-L2 (Best & Tyler, 2007) might be applied to make straightforward predictions for cross-linguistic tone-to-tone perception although such an approach needs to be more clearly elaborated upon and subsequently tested. Current approaches are

---

[3] Another difference between the Feature Hypothesis and the *Pitch Prominence Hypothesis* is that our study looks at the baseline for perception in non-learners, and does not address learners’ phonological systems, unlike the Feature Hypothesis. This difference prevents the direct application of the Feature Hypothesis to our data.

examining this possibility (Hao, 2012, or So & Best, 2010). However, the question remains as to the extent to which lexically-contrastive pitch categories can be defined for non-tonal language speakers (e.g., Stagray & Downs, 1993). A clearer case for the existence of lexically-contrastive pitch categories could be made for pitch-accent languages such as Swedish or Japanese as such languages are defined as a subclass of tone languages (Yip, 2002: 2) and therefore, allow the application of a tone-to-tone mapping approach as well.

To conclude, this study was conducted to further our understanding of cross-linguistic perception of tonal contrasts and expand current models of L2 phonology by advancing the *Pitch Prominence Hypothesis* as a first step in defining naïve perception of lexically-contrastive pitch and a baseline for L2 tone acquisition.

References

Abramson, Arthur S. (1975). The tones of Central Thai: Some perceptual experiments. In Jimmy G. Harris & James R. Chamberlain (Eds.), *Studies in Thai linguistics in honor of William J. Gedney* (pp. 1-16). Bangkok: Central Institute of English Language.

Abramson, Arthur S. (1976). Thai tones as a reference system. In T. W. Gething, Jimmy G. Harris, & Pranee. Kullavanijaya (Eds.), *Thai linguistics in honor of Fang-Kuei Li*. Bangkok: Chulalongkorn University Press.

Abramson, Arthur S. (1978). Static and dynamic acoustic cues in distinctive tones. *Language and Speech*, *21*, 319-325.

Beckman, Mary, Hirschberg, Julia, & Shattuck-Hufnagel, Stefanie. (2005). The original ToBI system and the evolution of the ToBI framework. In Sun-Ah. Jun (Ed.) *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford University Press

Best, Catherine. (1995). A direct realist view of cross-language speech perception. In Winifred Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Baltimore, MD: York Press.

Best, Catherine & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In Ocke-Schwen Bohn & Murray Munro (Eds.), *Second-language speech learning: The role of language experience in speech perception and production. A festschrift in honor of James E. Flege* (pp. 13-34). Amsterdam: John Benjamins.

Bohn, Ocke-Schwen. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In Winifred Strange (Ed.), *Speech perception and linguistic experience. Issues in cross-language research* (pp. 273-304). Baltimore: York Press.

Braun, Bettina & Johnson, Elizabeth K. (2011). Question or tone 2? How language experience and linguistic function guide pitch processing. *Journal of Phonetics, 39*, 585-594.

Burnham, Denis & Francis, Elizabeth. (1997). The role of linguistic experience in the perception of Thai tones. In Arthur S. Abramson (Ed.), *Southeast Asian linguistic studies in honour of Vichin Panupong* (pp. 29–46). Bangkok: Chulalongkorn University Press.

Burnham, Denis, Kirkwood, Kathryn, Luksaneeyanawin, Sudaporn, & Pansottee, Supatra. (1992). Perception of Central Thai tones and segments by Thai and Australian adults. In *Pan-Asiatic Linguistics: Proceedings of the Third International Symposium on Language and Linguistics* (pp. 546-560). Bangkok: Chulalongkorn University Press.

Cooper, Nicole, Cutler, Anne, & Wales, Roger. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech, 45,* 207-228.

Dupoux, Emmanuel, Pallier, Christophe, Sebastián-Gallés, Nuria, & Mehler, Jacques. (1997). A destressing "deafness" in French? *Journal of Memory and Language*, *36*, 406-421.

Dupoux, Emmanuel, Sebastián-Gallés, Nuria, Navarrete, Eduardo, & Peperkamp, Sharon. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition, 106*, 682-706.

Feldhausen, Ingo, Pešková, Andrea, Kireva, Elena, & Gabriel, Christoph. (2011). Categorical perception of Porteño nuclear accents. *Proceedings of the 17^{th} International Congress of Phonetic Sciences* (ICPhS XVII), Hong Kong, 17-21 August 2011 (pp. 116-119). Hong Kong, China.

Flege, James. (1995). Second language speech learning: Theory, findings and problems. In Winifred Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Baltimore, MD: York Press.

Francis, Alexander. L., Ciocca, Valter, Ma, Lian, & Fenn, Kimberly M. (2008). Perceptual learning of Cantonese lexical tones by tonal and non-tonal language speakers. *Journal of Phonetics*, *36*, 268-294.

Gandour, Jackson. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics, 11*, 49-175.

Gandour, Jackson & Harshman, Richard. (1978). Crosslanguage differences in tone perception: a multidimensional scaling investigation. *Language and Speech, 21*, 1-33.

Goto, Hiromu. (1971). Auditory perception by normal Japanese adults of the sounds "l" and "r". *Neuropsychologia, 9*, 317-323.

Guion, Susan G. & Pederson, Eric. (2007). Investigating the role of attention in phonetic learning. In Ocke-Schwen Bohn & Murray Munro (Eds.), *Language experience in second language speech learning* (pp. 57-77). Amsterdam: John Benjamins.

Hallé, Pierre A., Chang, Yueh-Chin, & Best, Catherine T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics, 32*, 395-421.

Hao, Yen-Chen. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics, 40,* 269-279.

Hao, Yen-Chen. (2008). Intonational interference on the perception and production of Mandarin Chinese tones by English-speaking second language learners. *Journal of the Acoustical Society of America, 124*, 2595 (A).

Kim-Renaud, Young-Key. (2009). *Korean: An essential grammar*. New York: Routledge.

Kondo, Mariko. (2007). Acoustic realization of lexical accent and its effects on phrase intonation in English speakers' Japanese. *Proceedings of the 17th International Congress of Phonetic Sciences* (ICPhS XVII), Hong Kong, 17-21 August 2011 (pp. 1649-1652). Hong Kong, China.

Leather, Jonathan. (1983). Speaker normalization in perception of lexical tone. *Journal of Phonetics, 11*, 373-382.

Leather, Jonathan. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In Jonathan Leather & Allan James (Eds.), *New Sounds 90: Proceedings of the Amsterdam Symposium on the Acquisition of Second-Language Speech* (pp. 72-89). Amsterdam: University of Amsterdam.

Lee, Chao-Yang. (2007). Does Horse Activate Mother? Processing Lexical Tone in Form Priming. *Language and Speech*, *50*, 101-123.

Lee, Chao-Yang. (2009). Identifying isolated, multispeaker Mandarin tones from brief acoustic input: A perceptual and acoustic study. *Journal of the Acoustical Society of America, 125*, 1125-1137.

Lee, Chao-Yang, Tao, Liang, & Bond, Z. S. (2008). Identification of acoustically modified Mandarin tones by native listeners. *Journal of Phonetics, 36*, 537-563.

Lee, Lisa & Nusbaum, Howard C. (1993). Processing interactions between segmental and suprasegmental information in native speakers of English and Mandarin Chinese. *Perception & Psychophysics, 53,* 157-165.

McAllister, Robert, Flege, James E., & Piske, Thorsten. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics, 30*, 229-258.

McGinnis, Scott. (1997). Tonal distinction errors by beginning Chinese language students: A comparative study of American English and Japanese native speakers. In Scott McGinnis (Ed.), *Chinese pedagogy: An emerging field* (pp. 81-91). Columbus, OH: Ohio State University Foreign Language Publications.

Miyawaki, Kanika, Strange, Winifred, Verbrugge, Robert R., Liberman, Alvin M., Jenkins, James J., & Fujimura, Osamu. (1975). An effect of linguistic experience: the discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics, 18*, 331-340.

Pierrehumbert, Jane, & Beckman, Mary. (1988). *Japanese tone structure. Linguistic inquiry monograph 15*. Cambridge: MIT Press.

Repp, Bruno H. & Lin, Hwei-Bing. (1990). Integration of segmental and tonal information in speech perception: A cross-linguistic study. *Journal of Phonetics, 18*, 481-495.

Sheldon, Amy & Strange, Winifred. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, *3,* 243-261.

Shen, Xiaonan, Lin, Mascan, & Yan, Jingzhu. (1993). F0 turning point as an F0 cue to tonal contrast: A case study of Mandarin tones 2 and 3. *Journal of the Acoustical Society of America, 93*, 2241-2243.

Silva, David J. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology, 23*, 287-308.

So, Connie K. (2006). Perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. In Paul Warren & Catherine I. Watson (Eds.), *Proceedings of the 11th Australian International Conference on Speech Science & Technology*. Auckland, New Zealand: University of Auckland.

So, Connie K. (2010). Categorizing Mandarin tones in Japanese pitch-accent categories: The role of phonetic properties. Paper presented at *Interspeech 2010 Second language studies: Acquisition, learning, education and technology.* Waseda University, Tokyo, Japan.

So, Connie K. & Best, Catherine T. (2010). Cross-language perception of non-native tonal contrasts: effects of native phonological and phonetic influences. *Language and Speech, 53*, 273-293.

Stagray, James R. & Downs, David. (1993). Differential sensitivity for frequency among speakers of a tone and a nontone language. *Journal of Chinese Linguistics, 21*, 143-163.

Sukekawa, Yasuhiko, Choi, Hyuncheol, Maekawa, Kikuo, & Sato, Shigeru. (1995). Perception of pitch accent by Korean learners of Japanese and its implications. *Denshi Joho Tsushin Gakkai gijutsu kenkyu hokoku: shingaku giho, 95*(41)/19950518, 61-66.

Tuc, Ho-Dac. (2003). *Vietnamese-English Bilingualism: Patterns of Code-switching*. London: Routledge-Curzon.

Van Lancker, Diana. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *Papers in Linguistics: International Journal of Human Communication, 13,* 201-277.

Van Lancker, Diana. & Fromkin, Victoria A. (1973). Hemispheric specialization for pitch and “tone”: Evidence from Thai. *Journal of Phonetics, 1,* 101-109.

Vũ, Thanh P. (1981). *The acoustic and perceptual nature of tone in Vietnamese*. PhD thesis, Australia National University, Canberra.

Wang, William S-Y. (1976). Language change. In Stevan R. Hamad, Horst D. Steklis, & James Lancaster (Eds.), *Origins and Evolution of Language and Speech. Annals of the New York Academy of Sciences, Vol. 280* (pp. 61-72).

Wang, Yue, Behne, Dawn M., Jongman, Allard, & Sereno, Joan A. (2004). The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Linguistics, 25*, 449-466.

Wang, Yue, Jongman, Allard, & Sereno, Joan A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language, 78*, 332-348.

Wang, Yue, Jongman, Allard, & Sereno, Joan A. (2006). L2 acquisition and processing of Mandarin Chinese tones. In P. Li, E. Bates, L.H. Tan, & O. Tseng (Eds.), *The Handbook of East Asian Psycholinguistics* (pp. 250-256). Cambridge: Cambridge University Press.

Wang, Yue, Spence, Michelle M., Jongman, Allard, & Sereno, Joan A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America, 106*, 3649-3658.

Wayland, Ratree P. & Guion, Susan G. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning, 54*, 681-712.

Whalen, Douglas H. & Xu, Yi. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica, 49,* 25-47.

White, C. M. (1981). Tonal perception errors and interference from English intonation. *Journal of Chinese Language Teachers Association, 16*, 27-56.

Wood, Charles C. (1974). Parallel processing of auditory and phonetic information in speech discrimination. *Perception & Psychophysics, 15*, 501-508.

Yip, Moira. (2002). *Tone*. Cambridge: Cambridge University Press.

Zsiga, Elizabeth & Nitisaroj, Rattima. (2007). Tone features, tone perception, and peak alignment in Thai. *Language and Speech, 50*, 343-383.

Selected Proceedings of the 2012 Second Language Research Forum: Building Bridges between Disciplines

edited by
Ryan T. Miller, Katherine I. Martin, Chelsea M. Eddington, Ashlie Henery, Nausica Marcos Miguel, Alison M. Tseng, Alba Tuninetti, and Daniel Walter

Cascadilla Proceedings Project   Somerville, MA   2014

Copyright information

ISBN 978-1-57473-464-5 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Schaefer, Vance and Isabelle Darcy. 2014. Linguistic Prominence of Pitch within the Native Language Determines Accuracy of Tone Processing. In *Selected Proceedings of the 2012 Second Language Research Forum*, ed. Ryan T. Miller et al., 1-14. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #3081.