Concordia Working Papers in Applied Linguistics

Proceedings of the International Symposium on the Acquisition of Second Language Speech
Concordia Working Papers in Applied Linguistics, 5, 2014

Pitch Prominence Matters: Perception of Thai Tones by Seoul Korean and Kyungsang Korean Speakers

Vance Schaefer
Indiana University

Isabelle Darcy
Indiana University

Abstract

This paper is a follow-up analysis of a previous study on the perception of Thai tones by speakers of various L1s differing in pitch prominence. Varying degrees of pitch prominence (Mandarin, tone = high; Japanese, pitch accent = intermediate; English, word stress = reduced; Korean, none = low) globally resulted in the expected hierarchy of performance: Mandarin (M = 87% correct), Japanese (M = 77%), English and Korean (M = 67% for both). However, the equal performance between the English and Korean participants was not predicted. In this study, we examined whether differences in dialectal exposure among the Koreans influenced results. Three speakers of the Korean Kyungsang dialect featuring pitch accent performed with higher accuracy than Seoul dialect speakers, reaching comparable accuracy levels with the L1 Japanese speakers. Additionally, the Seoul dialect speakers performed less accurately than the L1 English speakers. Both results confirm the originally predicted hierarchy of performance.

This study specifically examines how Thai tones are perceived by naïve listeners (= non-learners of Thai) who speak a range of languages varying in the usage of lexically-contrastive pitch: a tone language (Mandarin Chinese), a pitch accent language (Japanese, Kyungsang Korean dialect), a word-stress language (English), and a language which does not employ pitch in any way to distinguish the meaning of words (standard Korean). Pitch is generally defined as the variations in the height of the voice or fundamental frequency (abbreviated as f0) used to distinguish words.

We ask whether pitch prominence shapes tone perception. Prominence here is based on the function of linguistic pitch to signal lexical contrast (cf. Feature Hypothesis, McAllister, Flege, & Piske, 2002). Our working hypothesis is that the greater the prominence of lexical pitch in the first language (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.

We selected 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 standard Korean does not use pitch to contrast words but uses pitch in a domain larger than a word (i.e., intonational phrase) like all languages that we are aware of. Thus, naïve listeners who encounter a new sound system which utilizes pitch in a new manner do not approach the task without any "tools": their perception will most likely rely on the pitch usage defined by their L1 phonological grammar. For example, the presence in the L1 of lexical pitch contrasts has been shown to facilitate perception of non-native tone (Burnham et al., 1996). To further characterize the role of pitch in the languages used in this study, we could arrange them according to language types on a functional scale, such as one proposed by Van Lancker in 1980, summarized in Table 1. We see that languages use pitch contrasts in different domains (from small to large), and also in the function that the pitch fulfills, whether lexical or non-lexical (i.e., from tonal contrasts to marking focus structure or affect, for example). When the domain and function of pitch overlap between the L1 and L2, we (generally) expect perception of the non-native linguistic pitch system to be more reliable.

Combining the idea of functional use of pitch with the prominence idea, we can establish a scale of prominence for lexically-contrastive pitch going from maximal to low (see Table 1). We also have an additional distinction between English and Japanese based on whether or not pitch can be exclusively used to distinguish words; English word stress has multiple correlates in addition to pitch while Japanese relies on pitch only. 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 low as well. Based on the prominence for lexically-contrastive pitch in a given language, specific predictions for accuracy performance on naïve tone perception by the speakers of this language can be made.

However, languages do not necessarily neatly coincide within these four language types, especially when dialects are considered. For example, Japanese is often cited as a typical pitch-accent language, but features dialects which vary in pitch-accent systems as compared to so-called standard Japanese or are considered to be pitch-accentless, e.g., Fukushima and Kumamoto dialects. Another example is standard Korean which is considered to be a language which does not use pitch to distinguish lexical items (Kim-Renaud, 2009). However, the Kyungsang dialect spoken in the southeast region of South Korea is a pitch-accent dialect. Furthermore, the contrastive usage of pitch may be appearing among younger speakers of the Seoul dialect (Silva, 2006). According to our scale of prominence, we would predict that speakers of a pitch-accentless dialect of Japanese and a pitch-accent dialect of Korean would perform less or more accurately, respectively, than their counterparts on the perception of another lexically-contrastive pitch system, i.e., Thai tones. Hence, care must be taken when describing the use of lexical pitch in a language, highlighting the necessity to narrowly define one's L1 as one's first dialect.

A previous publication of part of this study's results (Schaefer & Darcy, in press) demonstrated that the accuracy rates in the naïve perception of Thai tones as evidenced in an AXB discrimination task indeed appear to be determined by the degree to which lexically-contrastive pitch is used in one's L1. These results, summarized in the following section, are in line with the Feature Hypothesis (McAllister, et al., 2002) which predicts, whether for segmental and suprasegmental dimensions alike, 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). Accuracy scores from the study resulted in a hierarchy of performance for Thai tones discrimination: L1 Mandarin > L1 Japanese > L1 English = L1 Korean (see Table 1 for predictions).

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

Pitch pattern	Prosodic Domain	Lexical status of pitch use/function	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 (Standard Korean)	Intonational phrase, PPh	Non-lexical	Low	Lowest

Note: PPh = Phonological phrase

However, an unexpected finding was the equal performance of the L1 English speakers and L1 Korean speakers. This paper addresses this unexpected result (see Schaefer & Darcy, in press, for a more extensive literature review). In this follow-up to our study, we specifically ask whether the L1 English speakers performed lower than expected or whether the L1 Korean speakers performed higher than expected. That is, we question whether f0 information is less readily accessible for phonological discrimination for the L1 English speakers and more so for the L1 Korean speakers. There are two possible scenarios. The first one is that the word stress correlate of pitch in English is not robust enough for the L1 English speakers to access/transfer to their perception of Thai tones. The second one is that the Koreans were able to use the presence of L1 pitch accent patterns (either from the Kyungsang dialect or "young-generation" Seoul dialect) to aid their perception of Thai tones. In this paper, we explore the latter hypothesis and explicitly investigate whether dialectal exposure has influenced Korean performance.

THE SCHAEFER & DARCY STUDY

Participants, Materials and Procedure

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 pitch-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). However, exposure to Kyungsang dialect speakers might be expected for this speaker.

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 non-words). 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 [ɤ]. 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: $[\text{bi:}]^{L}$ – $[\text{bi:}]^{L}$ – $[\text{bi:}]^{M}$.

The experiment contained two conditions, test and control, with 48 trials each. In the test condition, the syllables within one triplet only differed by tone; the segmental make-up of the syllables remained the same. In the control condition, all syllables in the triplet had the same tone but varied 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. These specific comparisons are not discussed in the present paper. 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 three 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 for the control condition 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 (RT) 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.

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 15-20 minutes in total. The participants also filled out a questionnaire about their demographic and linguistic background.

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 RTs 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 further 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. Analyses of variance (ANOVAs) were run in SPSS 20 by subjects on average accuracy comparing Condition (*Test* vs. *Control*) and Language Groups, to examine 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.

Overall accuracy rates and reaction times on the test vs. control are presented in Table 3. 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 the Thai revealed a significant main effect of condition (F (1, 37) = 7.0, $p < .05$). There is also a significant main effect of group (F (3, 37) = 3.9, $p < .05$), and 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.

Table 3. Accuracy means (%) and RTs (in ms) by language group for test and control conditions.

	Accuracy			Reaction times		
Language group	Test	Control	(SE)	Test	Control	(SE)
Thai (n=2)	95.6	88.5	6.2	844	751	142.5
Mandarin (n=10)	87.0	77.0	2.8	1122	1054	63.7
Japanese (n=11)	77.5	80.0	2.6	1268	1206	60.8
English (n=10)	67.3	75.9	2.8	1327	1137	63.7
Korean (n=10)	67.5	83.5	2.8	1345	1168	63.7

Note. SE = standard error

Thai and Mandarin listeners were faster than all the other non-native groups. Overall, latencies in the test condition were about 120 ms slower than on the control condition (1265 ms vs. 1141 ms). There is a main effect of condition (F (1, 37) = 31.4, p < .001). The main effect of group was not significant (F (3, 37) = 1.7, p > .1). The lack of group effect is likely due to the three non-native groups (Japanese, English and Korean) displaying comparable reaction times to each other within Test and Control conditions, respectively. The interaction was not significant (F (3, 37) = 2.4, p = .08).

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 our phonological tone discrimination task. Again we note that the data 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 now examine in more detail the performance of the Korean group.

THE CURRENT STUDY

The group of Koreans was not fully homogenous in terms of dialect as can be seen in Table 4. We see that four out of ten participants were not speakers of the Seoul dialect. Three were speakers of the pitch-accent Kyungsang dialect while one was a speaker from Cholla, a region in the southwest of South Korea abutting the Kyungsang dialect region. We know that Kyungsang listeners show categorical perception of pitch accent patterns (e.g., LH, HL, HH to distinguish minimal word pairs) (Kim & de Jong, 2007; Kim, 2011) and limited advantage in the naïve perception of Japanese pitch accent (Sukegawa, Choi, Maekawa, & Sato, 1995). If the L1 phonological system determines accuracy, then Kyungsang dialect speakers should outperform non-Kyungsang dialect speakers, unless emerging pitch accent patterns in Seoul Korean are robust enough to level performance across the whole group.

Table 4. Background of Korean participants

Code	Age	Dialect region	Languages studied	Field of study	Time abroad
KRF1	35	Seoul (Standard)	E, G	Visiting Ph.D. student (business)	US 6 months, Germany 1 month
KRF2	30	Seoul (Standard)	E, J	Business	US 1 ½ years
KRF3	29	Busan (Kyungsang)	E, G, F, J	SLS	US 5 years
KRF4	32	Seoul (Standard)	G, J	SLS	US 4 ½ years, Canada 4 months
KRF5	30	Jinju (Kyungsang)	E, F	Linguistics	US 3 ½ years
KRF6	27	Seoul (Standard)	E, F, J, M	SLS	US 4 years
KRF7	35	Seoul (Standard)	E, J, R	Slavic	US 6 ½ years, Japan 1 year, Russia 1 year
KRM1	27	Busan (Kyungsang)	E, F, G, J	Undergraduate	US 5 years, Japan 1 ½ years
KRM2	47	Cholla (near Kyungsang)	E, F, G, J	Linguistics	US 5 years
KRM3	30	Seoul (Standard)	E, J	Linguistics	US 5 years

Note. KR = Korean, F = female, M = male, E = English, G = German, J = Japanese, M = Mandarin, R = Russian, F = French, SLS = Second Language Studies

We turn now to examining their performance on the test vs. control condition, in order to see if the people who come from pitch accent regions would perform more accurately than the others. Figure 1 shows the individual performance for each of the Korean participants on the combined test conditions, and Figure 2, on the control condition. Black bars represent Seoul dialect speakers while white bars represent Kyungsang dialect speakers; the lone grey bar represents the speaker from Cholla.

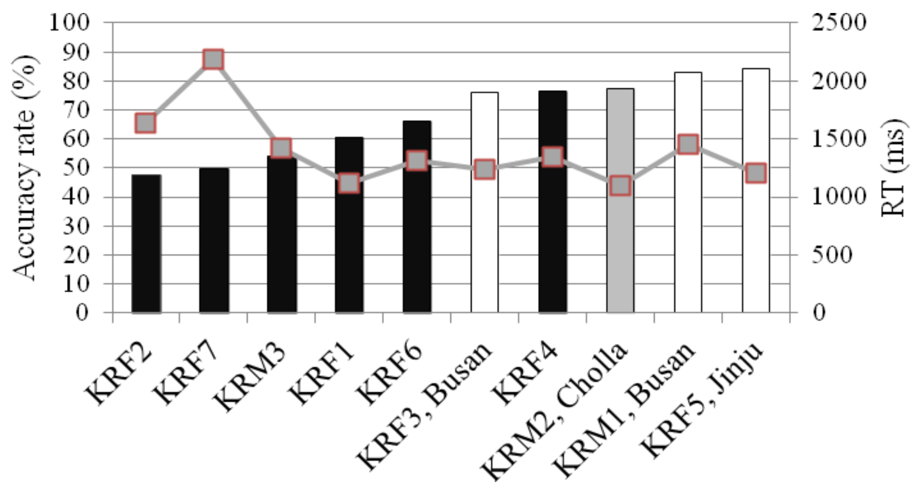

Figure 1. Mean individual accuracy rate (%) and RT (ms) for the Koreans on the combined test conditions.

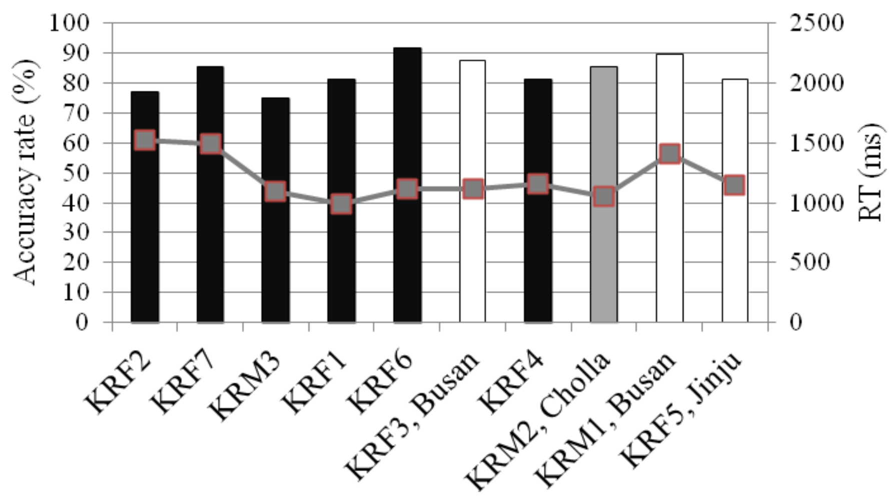

Figure 2. Mean individual accuracy rate (%) and RT (ms) for the Koreans on the control condition.

In Figure 1, all the Kyungsang dialect speakers, i.e., the pitch accent speakers, are on the higher end of the accuracy spectrum, outperforming non-pitch-accent dialect speakers. In Figure 2 for the control condition, however, there is no clear relationship between dialectal group and performance. Figure 3 now compares the performance of the Korean group to the performance obtained by the other groups. The group of Koreans is split by dialect region. In Figure 3, Thai listeners are displayed in black for comparison purposes.

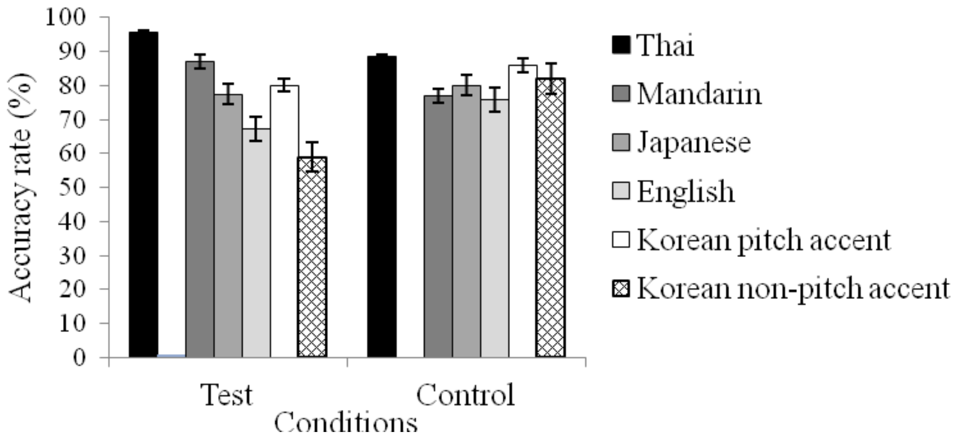

Figure 3. Mean accuracy rate (%) for each language group with the Koreans split into subgroups for the test vs. control condition. Error bars enclose +/- 1 standard error.

We see a clear difference between the two Korean subgroups. The non-pitch accent speakers are in fact less accurate than the L1 English speakers, which conforms to our original prediction while the pitch-accent speakers in contrast outperform the English speaker group, and are more comparable to the Japanese speaker group.

DISCUSSION AND CONCLUSIONS

Previous findings by Schaefer and Darcy (in press) offered evidence that performance on the perception of non-native Thai tones varies as a function of the presence and prominence of lexically-contrastive pitch in the L1. The 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. However, the finding that speakers of L1 English (with word stress) and speakers of L1 Korean (no lexically-contrastive pitch) both reached comparable levels of accuracy had not been predicted. Therefore, the goal of this follow-up was to examine one possible reason for this symmetrical pattern: In this current analysis, we specifically asked whether the Korean listeners performed more accurately than expected, focusing on the hypothesis that the effect could be due to exposure to the Kyungsang pitch-accent dialect (Sukegawa et al., 1995).

The results from this follow-up analysis provide further evidence for the interpretation obtained from the Schaefer and Darcy (in press) findings: the prominence of lexically-contrastive pitch use in different L1s shapes the cross-linguistic perception of non-native tone. This interpretation is supported here even more specifically by the effect of the native dialect (D1) in the Korean group.

A speaker's D1 is essentially that speaker's first language. Oftentimes if the D1 is not the standard dialect of a country, then speakers will be exposed to a second dialect (D2) in the form of the standard dialect. One's D1 has been shown to result in D1/D2 asymmetry from a young age constraining lexical access for suprasegmentals (Otake & Cutler, 1999) and segmentals (Sumner & Samuel, 2009), further bolstered for segmentals by some studies on closely related languages (Pallier, Bosch, and Sebastián-Gallés, 1997). Accordingly, we expect the D1 to interact with the D2 just as an L1 and L2 interact, i.e., "Interaction Hypothesis" (Weinreich, 1953; Flege, Frieda, & Nozawa, 1997). This is substantiated by the fact that different D1s of the same L1 influence the pronunciation of an L2 (Cui & van Heuven, 2011), suggesting possible differences in the naïve perception of another language.

A potential alternative explanation had to do with the emergence of lexical pitch among younger speakers of the Seoul dialect (Silva, 2006). In our sample, the ages of the Seoul speakers ranged from 27-35. This may have aided the Korean group in the perception of Thai tones altogether. Yet, our results appear to not bear this out, suggesting that this development is not as robust as lexical pitch in the Kyungsang dialect.

It has also been pointed out that English listeners may have performed less accurately due to the fact that f0 is rarely used alone to distinguish words in English and may be only a weak cue to word stress (Sluijter & van Heuven, 1996). Yet, our study appears to demonstrate that pitch in English word stress may indeed be robust enough to differentiate these speakers' performance in Thai tonal perception accuracy from that of the L1 standard Korean speakers. We, therefore, may have also answered our question as to whether the English performed less accurately than expected: it appears that they have not, as they outperform the speakers of L1 standard Korean. Additionally, it appears that exposure to the pitch correlate of L2 English word stress by the speakers of standard Korean was not sufficient enough to boost their performance to the levels of the L1 speakers of English.

The present study allows us to further refine the picture put forward in Schaefer and Darcy (in press). That study established a baseline for tone perception focusing on the functional use of linguistic pitch in four language types. Within the framework of the Feature Hypothesis, the *Pitch Prominence Hypothesis* was proposed to model the perception of lexically-contrastive pitch. The findings globally supported the model in the case of a non-native tone system, but were not fully borne out given the equal performance of L1 English and L1 Korean speakers. After taking the role of native dialect exposure into account, the present reanalysis bolsters the model further by confirming the accuracy of the originally predicted hierarchy. An additional consideration for the *Pitch Prominence Hypothesis,* or for any study of naïve or L2 perception, is the necessity to define one's L1 in the narrow sense of one's dialect as suggested by the results of the Korean speakers in this study.

While our findings seem to lend support to this hypothesis, the small sample size accessible in this study does not yet allow for definite conclusions and alternative explanations cannot be fully ruled out. For example, perhaps differences in English proficiency (or in another L2) among participants indirectly contribute to the observed differences independently of the dialectal origin.

To conclude, these findings highlight the need to more thoroughly examine the L1 of participants in terms of dialect and possibly other factors. We also encourage the continued examination of the influence of the acoustic correlates of word stress in English on the naïve and/or L2 perception of other linguistic pitch systems.

ACKNOWLEDGEMENTS

We thank members of the audience at New Sounds 2013.

REFERENCES

Burnham, D., Francis, E., Webster, D., Luksaneeyanawin, S., Lacerda, F., & Attapaiboon, C. (1996). Facilitation or attenuation in the development of speech mode processing? Tone perception over linguistic contexts. In P. McCormack, & A. Russell (Eds.), *Proceedings of the Sixth Australian International Conference on Speech Science and Technology* (pp. 587-592). Canberra: Australian Speech Science and Technology Association.

Cui, R., & van Heuven, V. J. (2011). Mutual intelligibility of English vowels by Chinese dialect speakers. In W.-S. Lee, & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences 2011 [ICPhS XVII]* (pp. 544-547). Hong Kong: Department of Chinese, Translation and Linguistics, City University of Hong Kong.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkcamp, S. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition, 106*, 682-706.

Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics, 25*, 169-186.

Kim, J.-S. (2011). Perception of lexical pitch accent by Kyungsang and Cholla Korean listeners. In W.-S. Lee, & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences 2011 [ICPhS XVII]* (pp. 1070-1073). Hong Kong: Department of Chinese, Translation and Linguistics, City University of Hong Kong.

Kim, J.-S., & de Jong, K. J. (2007). Perception and production in the pitch accent system of Korean. In J. Trouvain, & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences 2007 [ICPhS XVI]* (pp. 1273-1277). Dudweiler: Pirrot.

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

McAllister, R., Flege, J. E., & Piske, T. (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.

Otake, T., & Cutler, A. (1999). Perception of suprasegmental structure in a nonnative dialect. *Journal of Phonetics, 27*, 229-253.

Pallier, C., Bosch, L., & Sebastian-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition*, *64*, B9-B17.

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

Sluijter, A. M. C., & van Heuven, V. J. (1996). Acoustic correlates of linguistic stress and accent in Dutch and American English. In H. T. Bunnell, & W. Idsardi (Eds.), *Proceedings of the Fourth International Conference on Spoken Language Processing [ICSLP 96]* (pp. 630-633). Wilmington, Delaware: University of Delaware, I. DuPont Institute.

Sukegawa, Y., Choi, H., Maekawa, K., & Sato, S. (1995). Perception of pitch accent by Korean learners of Japanese and its implications. *The Institute of Electronics, Information and Communication Engineers (IEICE) technical report: Speech, 95*(41), 61-66.

Sumner, M., & Samuel, A. G. (2009). The role of experience in the processing of cross-dialectal variation. *Journal of Memory and Language*, *60*, 487-501.

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

Weinreich, U. (1953). *Languages in contact.* The Hague: Mouton Publishers.

Schaefer, V., & Darcy, I., (in press). Linguistic Prominence of Pitch within the Native Language Determines Accuracy of Tone Processing. In R. T. Miller, K. I. Martin, C. M. Eddington, A. Henery, N. Marcos Miguel, A. Tseng, A. Tuninetti, & D. Walter (Eds.), *Selected Proceedings of the 2012 Second Language Research Forum: Building Bridges Between Disciplines: SLA in Many Contexts*. Somerville, MA: Cascadilla Proceedings Project.