

Better than Native: Tone Language Experience Enhances English Lexical Stress

Discrimination in Cantonese-English Bilingual Listeners

William Choi^{1,2}, Xiuli Tong³, Arthur G. Samuel^{4,5,6}

1. Institute of Cognitive Neuroscience, University College London
2. Deafness, Cognition and Language Research Centre, University College
London
3. Division of Speech and Hearing Sciences, The University of Hong Kong
4. Department of Psychology, Stony Brook University
5. Basque Center on Cognition, Brain, and Language
6. Ikerbasque, Basque Foundation for Science

Abstract

While many second language (L2) listeners are known to struggle when discriminating non-native features absent in their first language (L1), no study has reported that L2 listeners perform *better* than native listeners in this regard. The present study tested whether Cantonese-English bilinguals were better in discriminating English lexical stress in individual words or pseudowords than native English listeners, even though lexical stress is absent in Cantonese. In experiments manipulating acoustic, phonotactic, and lexical cues, Cantonese-English bilingual adults exhibited superior performance in discriminating English lexical stress than native English listeners across all phonotactic/lexical conditions when the fundamental frequency (f_0) cue to lexical stress was present. The findings underscore the facilitative effect of Cantonese tone language experience on English lexical stress discrimination.

Keywords: tone language expertise, second language learning, lexical stress, lexical tone, cross-language speech perception

Better than Native: Tone Language Experience Enhances English Lexical Stress

Discrimination in Cantonese-English Bilingual Listeners

Second language learners are known to struggle when discriminating L2 speech contrasts absent in their L1. For example, Japanese listeners have difficulty discriminating /r/ and /l/, due to the lack of the /r/-/l/ contrast in Japanese (e.g., Iverson et al., 2003). Here, we argue that under specific circumstances, L1 experience can be advantageous for L2 listeners even when the L2 speech contrast is absent in L1. Specifically, we investigated Cantonese-English bilingual adults' ability to discriminate English lexical stress contrasts in individual words (e.g., whether IMport /'im'pɔrt / and imPORT /im'pɔrt / differ). Cantonese listeners were predicted to outperform native English listeners, even though Cantonese does not use lexical stress, due to their tone language expertise.

Lexical stress is the relative prominence given to a syllable within a word (Teschner & Whitley, 2004). In the canonical context¹, acoustic analyses show that a stressed syllable typically has a higher f₀, longer duration, and higher intensity than an unstressed syllable (e.g., see Wang, 2008 for review). Although other phonetic correlates such as vowel quality, hyperarticulation and spectral tilt are also cues to stress (e.g., Beckman & Edwards, 1994; Campbell & Beckman, 1997; De Jong, 1995; Rosner & Pickering, 1994; Sluijter & van Heuven, 1996), we focus on f₀, duration and

intensity because they are important acoustic cues for English lexical stress and are the primary cues manipulated in previous research on lexical stress in the canonical context (e.g., Wang, 2008; Yu & Andruski, 2010).

Several studies of Spanish lexical stress judgements by native French listeners provide a useful context for the current study (e.g., Dupoux, Pallier, Sebastian-Galles, & Mehler, 1997; Dupoux, Peperkamp, & Sebastian-Galles, 2010; Dupoux, Sebastian-Galles, Navarrete, & Peperkamp, 2008). Like Cantonese, French does not use lexical stress; like English, Spanish does. As would be expected, monolingual French listeners were less able to discriminate lexical stress than native Spanish listeners, even though the native French listeners were sensitive to the acoustic correlates of lexical stress (Dupoux et al., 1997). Dupoux and colleagues (2010) tested simultaneous French-Spanish bilinguals, late French-Spanish bilinguals, and Spanish monolinguals on Spanish lexical stress perception using short-term memory encoding and speeded lexical decision tasks. Across the tasks, the late French-Spanish bilinguals consistently performed more poorly than the Spanish monolinguals, consistent with their L1 (French) not using that feature of language. The above findings, usually described as “stress deafness”, suggest that non-native listeners can never be as proficient as native listeners in discriminating or perceptually encoding lexical stress if this feature is not utilized in their L1.

However, contrary to this proposition, some simultaneous French-Spanish bilinguals were actually *as good as* Spanish monolinguals in attending to Spanish lexical stress (Dupoux et al., 2010). On average, the simultaneous French-Spanish bilingual group performed better than the late French-Spanish bilingual group but worse than the Spanish monolingual group. A more nuanced analysis of individual performance revealed a bimodal distribution: Some simultaneous French-Spanish bilinguals performed as poorly as late French-Spanish bilinguals, while some achieved Spanish-like performance. Nonetheless, none of the simultaneous French-Spanish bilinguals demonstrated superior performance to that of the Spanish monolinguals; bilingualism per se did not lead to an enhanced ability to discriminate differences in lexical stress. In this context, the present study examines a rather bold hypothesis that Cantonese-English bilinguals will achieve *superior* performance to that of native English listeners in discriminating English lexical stress in the canonical context¹, even though Cantonese does not utilize lexical stress.

Our hypothesis arises from the possibility that the Cantonese listeners may harness superior f₀ processing skills in the service of English lexical stress discrimination. F₀ is the primary acoustic correlate of Cantonese lexical tones (e.g., Gandour, 1981; 1983), and a large body of evidence demonstrates that tone language experience induces fundamental changes in how listeners process f₀ information, and facilitated f₀

perception in both linguistic and non-linguistic contexts (Bidelmen, Hutka, & Moreno, 2013; Lee, Vakoch, & Wurm, 1996; Pfordrescher & Brown, 2009; Waylabd & Guion, 2004). For example, Bidelman and colleagues found that Cantonese non-musicians outperformed English non-musicians in pitch memory, musical melody discrimination, and pitch processing speed. Similarly, Lee, Vakoch, and Wurm (1996) showed that Cantonese listeners are better than English listeners in Mandarin tone discrimination.

Given that the variations in f_0 that differentiate the six lexical tones in Cantonese require more precise perceptual tracking than the cues for English lexical stress (Eady, 1982), native Cantonese listeners may have greater f_0 perceptual acuity than native English listeners. Thus, in the present study, we examined whether Cantonese-English bilinguals can leverage their f_0 tracking skill to enhance L2 English lexical stress discrimination in the canonical context.

A recent study of Cantonese-English bilingual listeners could be taken to suggest that learning English lexical stress in the context of a rich tone system (Cantonese) can enhance lexical stress discrimination. Tong, Lee, Lee, and Burnham (2015) tested Cantonese-English bilinguals and native English adults who were living in Hong Kong on a discrimination task for English lexical stress. The Cantonese-English bilinguals were faster and more accurate than the native English listeners in discriminating English lexical stress. The interpretation of this advantage is limited because (a) the

accuracy difference did not reach a statistical significance ($p = .061$) and (b) the test stimuli used consonants and vowels common in both Cantonese and English, leaving open the possibility that the listeners might perceive the stimuli as Cantonese rather than English sound sequences. Thus, despite the preliminary evidence showing that tone language experience provides non-native Cantonese-English bilingual listeners with a specific advantage in discriminating lexical stress, it is counter-intuitive to predict non-native performance being better than native performance. We conducted a proof of concept experiment, which provided preliminary support for the hypothesis; the methods and results are summarized in the Supplementary Materials.

The current study extends previous research to further examine whether Cantonese-English bilingual listeners are better than native English listeners in discriminating English lexical stress. In addition, this study clarifies how bilinguals make their judgments by examining two other key factors that have never been tested in previous research: (1) phonotactic/lexical influences, and (2) acoustic cues to lexical stress. For the first factor, we presented listeners with three stimulus types: items that were phonotactically legal pseudowords in both Cantonese and English; items that were phonotactically legal pseudowords in English, but illegal in Cantonese; and real English words (e.g., PERmit/perMIT). This manipulation allows us to examine

whether any advantage by Cantonese-English bilingual listeners is limited to stimuli that can be processed as Cantonese.

For the second factor, we varied the set of cues (f0, duration and intensity) to lexical stress that were present in the stimuli. Our reasoning is that, if the potential advantage that Cantonese bilinguals have over native English listeners is due to their enhanced sensitivity to f0, the elimination of f0 cues in the English stimuli should significantly diminish the advantage. Thus, for this factor, we included three types of stimuli: (1) naturally produced items, (2) items that preserved only the f0 cue, and (3) items that provided duration and intensity cues without the f0 cue. We expect that the advantage, if any, will only manifest in naturally produced items and items that preserve the f0 cue.

Methods

Participants

Thirty Cantonese-English bilingual listeners (*mean age* = 20.9 years; *SD* = 1.69 year) were recruited from The University of Hong Kong, and 30 native English listeners (*mean age* = 23.7 years; *SD* = 4.16 years) were recruited from Stony Brook University (*N* = 15) and the Massachusetts Institute of Technology (*N* = 15). Their language and musical backgrounds are summarized in the Supplementary Materials.

English Lexical Stress Discrimination Task

This task employed an AX test with an inter-stimulus interval of 600ms. On each trial, listeners heard two items that were segmentally identical, with one item presented in a male voice and the other presented in a female voice (voice order was random). The two items could either have the same lexical stress pattern (e.g., IMport – IMport), or different stress patterns (e.g., IMport – imPORT). Participants pushed one response button for “same” pairs, and a different button for “different” pairs.

There were nine blocks of trials, comprised of three phonotactic/lexical contexts (i.e., pseudowords phonotactically legal in Cantonese and English, pseudowords phonotactically legal in English but not in Cantonese, and English real words) × three acoustic contexts (all-cues, f0-only, duration-and-intensity only). Signal processing software was used to neutralize the duration and intensity cues for the f0-only stimuli, and to neutralize the f0 cue for the duration-and-intensity stimuli; see the Supplementary Materials for a detailed presentation of the stimulus properties and construction. The order of presentation of the nine blocks was randomized. A total of 96 trials (8 stimuli × 2 speaker orders × 2 trial types × 3 repetitions) were included in each block. Accuracy and response time were recorded on each trial. We report analyses of d' scores based on the hits and false alarms for the same and different trials. Reaction time analyses are completely consistent with what we report here (see the Supplementary Materials).

Control Tasks

We assessed participants' short-term memory and non-verbal intelligence using two tasks that Zheng and Samuel (2018) had used. The tasks are described in detail in the Supplementary Materials.

Results

Figures 1, 2 and 3 show the full data set, while Figures 4 and 5 show the results collapsed across phonotactic/lexical context and acoustic condition, respectively. To examine whether Cantonese-English bilingual listeners perform better than native English listeners in discriminating English lexical stress, a $3 \times 3 \times 2$ mixed factorial analysis of covariance (ANCOVA) was conducted, with acoustic condition (all-cues, f0-only, and duration-and-intensity-only) and phonotactic/lexical context (legal pseudoword, illegal pseudoword, and real word) being within-subject factors, and group (Cantonese -English bilinguals and native English listeners) being a between-subjects factor. Non-verbal intelligence was the covariate in the ANCOVA, to control for its effect (see the Supplementary Materials for a converging analysis that used a subset of the participants for which non-verbal intelligence was matched between the groups).

Our main question is whether Cantonese tone language experience enhanced listeners' ability to discriminate English lexical stress. The ANCOVA showed the

significant main effect of group, $F(1, 57) = 9.85, p = .003, \eta_p^2 = .15$: Cantonese-English bilinguals outperformed native English listeners on English lexical stress discrimination (see Figures 1, 2 and 3).

Our second hypothesis is that the advantage Cantonese-English bilingual listeners over native English listeners in English lexical stress discrimination is due to their enhanced sensitivity to f0 cues. To examine this hypothesis, we manipulated acoustic cues in order to isolate the f0 cue, either providing just this cue, or eliminating it and leaving just duration and intensity. As expected, the ANCOVA revealed a significant two-way interaction of acoustic condition \times group, $F(2, 114) = 9.33, p = .001, \eta_p^2 = .14$. As Figure 4 shows, this interaction reflects the non-significant difference between the Cantonese-English bilinguals and the native English listeners in the condition that lacked f0 information (“di” in Figure 4). Simple main effects analyses further revealed that Cantonese-English bilingual listeners outperformed native English listeners in the all-cues condition, $F(1, 57) = 14.59, p < .001, \eta_p^2 = .20, CI [.58, 1.87]$, and in the f0-only condition, $F(1,57) = 6.31, p = .015, \eta_p^2 = .10, CI [.18, 1.56]$ but not in duration-and-intensity-only condition, $F(1, 57) = .73, p = .395$. As predicted, the facilitative effect



Figure 1. Mean d 's of Cantonese-English bilingual and native English listeners across legal, illegal and real word contexts in the all cues condition. Error bars denote 95% confidence intervals.

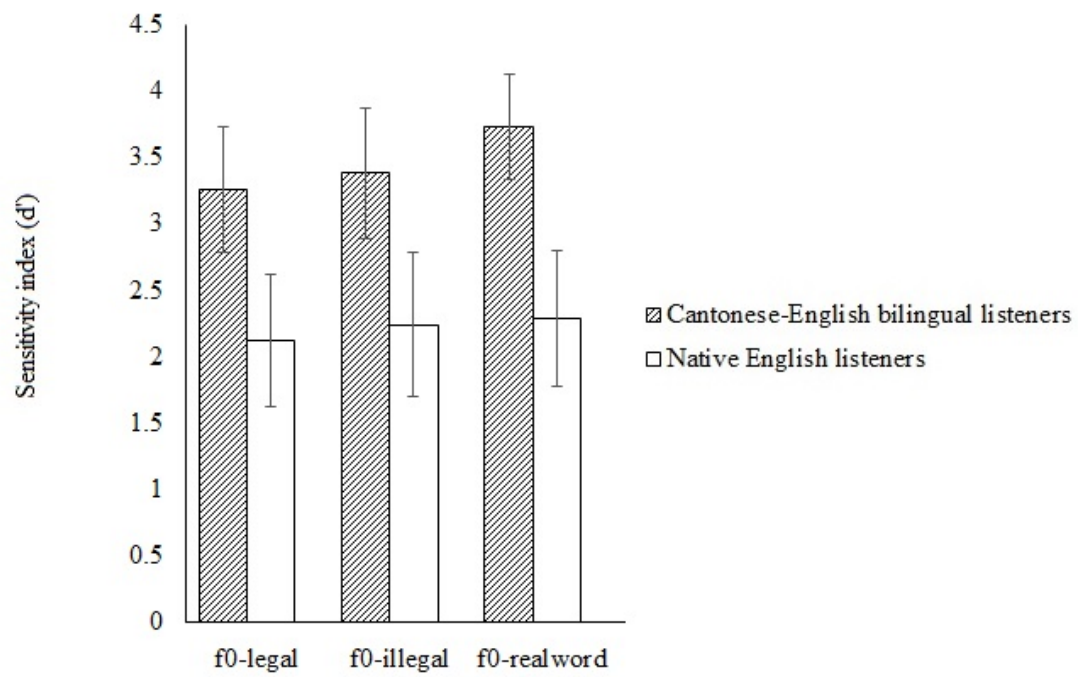


Figure 2. Mean d 's of Cantonese-English bilingual and native English listeners across legal, illegal and real word contexts in the fundamental frequency condition.

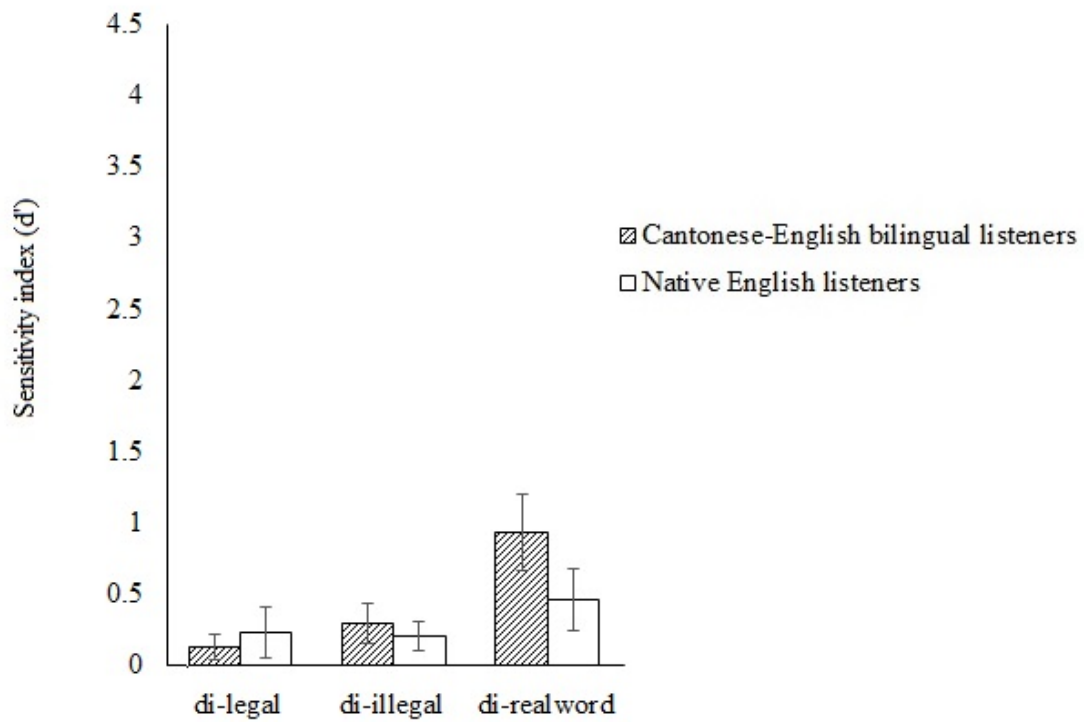


Figure 3. Mean d's of Cantonese-English bilingual and native English listeners across legal, illegal and real word contexts in the duration-and-intensity condition.

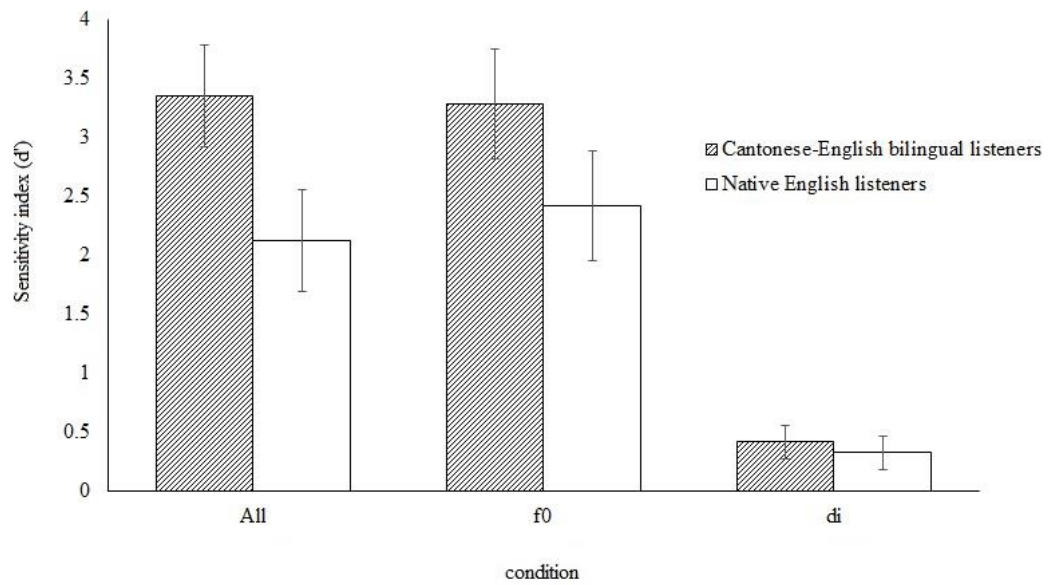


Figure 4. Mean d' 's of Cantonese-English bilingual and native English listeners across all acoustic conditions. All, f0 and di denote all cues, f0-only, and duration-and-intensity-only conditions respectively.

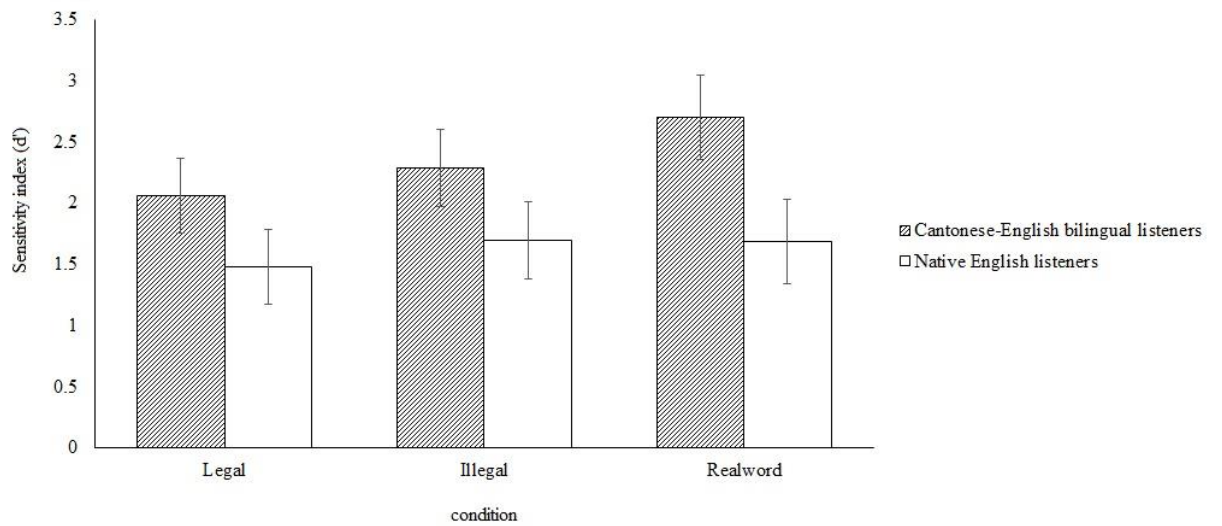


Figure 5. Mean d' 's of Cantonese-English bilingual and native English listeners across all phonotactic/lexical contexts.

of tone language experience on English lexical stress discrimination is absent when the f_0 cue is removed.

One potential concern is that the advantage for the Cantonese listeners might arise if they were treating the English stimuli as Cantonese syllables with tones. Anticipating this possibility, we included one set of items that were not phonotactically legal in Cantonese, and another set that were real English words. If the Cantonese listeners were “cheating”, this strategy should be very difficult or impossible to employ in these cases.

The ANCOVA produced a significant two-way interaction of context \times group, $F(2, 114) = 11.50, p < .001, \eta_p^2 = .17$, but not because these conditions showed a reduced advantage. In fact, as Figure 5 shows, the advantage for the Cantonese listeners was actually largest for the unambiguously English items – the real English words.

Simple main effects analyses indicated that Cantonese-English bilinguals outperformed native English listeners across legal $F(1, 57) = 6.48, p = .014, \eta_p^2 = .10, CI [.12, 1.04]$, illegal, $F(1, 57) = 6.25, p = .015, \eta_p^2 = .10, CI [.12, 1.06]$, and real word contexts, $F(1, 57) = 15.91, p < .001, \eta_p^2 = .22, CI [.51, 1.52]$. Thus, if anything, forcing listeners to process the speech as English enhances the advantage for the Cantonese-English bilingual listeners.

Collectively, the results demonstrate a robust advantage for Cantonese-English bilingual listeners relative to native English listeners in discriminating English lexical

stress. This advantage disappeared when the f₀ cue was removed, confirming that the enhanced discrimination was rooted in the Cantonese-English bilingual listeners' greater sensitivity to f₀ information. Importantly, the advantage was most pronounced in the real word context, a condition designed to force listeners to process the items as English.

Discussion

There are an endless number of situations in which non-native listeners are worse than native listeners in discriminating L2 speech contrasts. Occasionally, one can find a case in which non-natives achieve native-like perception (e.g., Dupoux et al., 2010). Our core finding is that Cantonese-English bilinguals were superior in discriminating English lexical stress than native English listeners, even though lexical stress is not used in Cantonese. This is the first case we are aware of in which non-native listeners significantly outperform native listeners in terms of both d' and response latency, despite the absence of the language feature (lexical stress) in their L1. It is often suggested that features absent in L1 are attenuated in non-native speech perception (e.g., Flege, 1995), such as the case of /r-/l/ for Japanese listeners. The current finding suggests the need to consider features which are not identical across L1 and L2, but which share a critical acoustic-phonetic cue, as in the case of Cantonese lexical tone and English lexical stress.

Our finding may be understood in terms of a theory designed to account for positive transfer between music and speech perception, Patel's OPERA hypothesis (Patel, 2011; 2014). The OPERA (Overlap, Precision, Emotion, Repetition and Attention) hypothesis posits that music and speech processing employ overlapping neural networks, such that the enhanced plasticity and neural precision resulting from musical training or language experience are transferrable across language/music domains. In OPERA, the "Precision" piece is most relevant – positive transfer occurs when one domain requires more perceptual precision than another. Because understanding Cantonese words requires precise f0 tracking of its lexical tones, the tone language experience may enhance the perceptual mechanisms that process f0 information, leading to better sensitivity to English lexical stress. In fact, it has been shown that tone language listeners have better f0 tracking accuracy than non-tone language listeners at the brainstem level (Krishnan, Gandour, & Bidelman, 2010). Our results indicate that the perceptual advantage for f0 information is not only transferrable across the same feature of different languages (Cantonese tones and Thai tones in the Krishnan et al. study), but also across related features of different languages, e.g., Cantonese lexical tones and English lexical stress.

The syllable-timed nature of Cantonese may have contributed to Cantonese listeners' advantage by leading them to orient their attention to f0 rather than duration

and intensity. Unlike English (Roach, 2003; Underhill, 1998), Cantonese is a syllable-timed language (Mok, 2008) and native speakers may not attend as much to syllable duration. If so, Cantonese listeners may focus on f_0 rather than rhythmic (duration and intensity) cues, which could synergistically combine with their enhanced f_0 sensitivity to facilitate English lexical stress discrimination. This explanation is consistent with cross-linguistic evidence showing that Mandarin (also a syllable-timed language) listeners attended to f_0 more than duration and intensity when perceiving English lexical stress (e.g., Chrabaszc et al., 2014; Wang, 2008; Yu & Andruski, 2010; Zhang & Francis, 2010).

Although direct tests remain to be done, the advantage observed here for Cantonese-English bilinguals appears to be rooted in Cantonese tone language experience, rather than bilingualism, *per se*. Recall that in Dupoux and colleagues' (2010) study of Spanish lexical stress perception, even the better sub-group of simultaneous French-Spanish bilinguals did not outperform Spanish monolinguals. Thus, it is unlikely that "generic" bilingual experience can account for the advantage found here for Cantonese-English bilingual listeners.

The current study demonstrates that Cantonese tone language experience facilitates Cantonese-English bilingual adults' discrimination of English lexical stress in the canonical context to the point that their ability exceeds that of native English

listeners. With this non-native advantage established, it would be interesting to investigate how general this advantage is. For example, would Cantonese-English bilinguals still show better-than-native ability when lexical stress is influenced by the pitch accent patterns at the sentence level? The current findings open up an intriguing set of new questions.

Footnote

1. In connected speech, the phonetic realization of a stressed syllable is affected by the intonational pitch-accent pattern superimposed to it (Beckman & Pierrehumbert, 1986).

For example, in the marked pitch accent context, a stressed syllable typically has a lower f_0 than the unstressed syllable (Ladefoged, 2006). In the present study, we focus on the canonical context—a word's simplest form in which a stressed syllable is typically associated with higher f_0 , longer duration and higher intensity relative to the unstressed syllable. The current study adds to a substantial lexical stress literature focused on discrimination/perception in the canonical context (e.g., Dupoux et al., 1997; 2008; 2010; Fry 1958; Wang, 2008; Yu & Andruski, 2010).

Acknowledgement

This article is, in part, based on the fourth chapter of the PhD thesis submitted by William Choi to The University of Hong Kong. This research was supported, in part, by the Language Learning Dissertation Grant from *Language Learning* to William Choi. It was also supported by the Pilot Scheme on International Experiences for Research Postgraduate Students from The University of Hong Kong to William Choi, and the Early Career Scheme (27402514), General Research Fund (17673216), and General Research Fund (17609518) from the HKSAR Research Grant Council to Xiuli Tong. Support was also provided by Ministerio de Ciencia E Innovacion, Grant PSI2014-53277, Centro de Excelencia Severo Ochoa, Grant SEV-2015-0490, and by the National Science Foundation under Grant IBSS-1519908 to Arthur Samuel. We thank Benjamin Munson for his useful suggestion about the syllable-timed nature of Cantonese and the four anonymous reviewers for comments that have helped us to develop our ideas and presentation more clearly.

References

- Battistella, E. L. (1990). *Markedness: The Evaluative Superstructure of Language*.
Albany: State University of New York Press
- Beckman, M. E., & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In P.A. Keating (Ed.), *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*, pp. 7-33. Cambridge University Press.
- Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology*, 3, 255-309.
- Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians shared enhanced perceptual and cognitive abilities for musical pitch: Evidence for Bidirectionality between the domains of language and music. *PLoS ONE*, 8(4), e60676.
- Campbell, N., & Beckman, M. E. (1997). Accent, stress, and spectral tilt. *Journal of the Acoustical Society of America*, 101(5), 3195.
- Choi, W., Tong, X., Gu, F., Tong, X., & Wong, L. (2017). On the early neural perceptual integrality of tones and vowels. *Journal of Neurolinguistics*, 41, 11-23.
- Chrabaszczyk, A., Winn, M., Lin, C. Y., & Idsardi, W. J. (2014). Acoustic cues to

perception of word stress by English, Mandarin, and Russian speakers. *Journal of Speech, Language and Hearing Research*, 57, 1468-1479.

de Jong, K. J. (1995). The supraglottal articulation of prominence in English:

Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America*, 97(1), 491-504.

Dupoux, E., Peperkamp, S., & Sebastian-Galles, N. (2010). Limits on bilingualism

revisited: Stress 'deafness' in simultaneous French-Spanish bilinguals. *Cognition*, 114, 266-275.

Dupoux, E., Pallier, C., Sebastian-Galles, N., & Mehler, J. (1997). A Destressing

"Deafness" in French? *Journal of Memory and Language*, 36(3), 406-421.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008).

Persistent stress 'deafness': The case of French learners of Spanish. *Cognition*, 106(2), 682-706.

Eady, S. (1982). Differences in the F0 patterns of speech: Tone language versus stress

language. *Language and Speech*, 25, 29-42.

Flege, J. (1995). Second-language speech learning: Theory, findings, and problems. In

W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp.229-273). Timonium, MD: York Press.

Fry, D. (1958). Experiments in the perception of stress. *Language and Speech*, 1, 205-

213.

Gandour, J. (1981). Perceptual dimensions of tone: Evidence from Cantonese. *Journal of Chinese Linguistics*, 9, 20-36.

Gandour, J. T. (1983). Tone perception in far eastern languages. *Journal of Phonetics*, 11(2), 149-175.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, 87, B47-B57.

Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics*, 23, 81-95.

Ladefoged, P. (2006). *A Course in Phonetics*. California: Thomson Wadsworth Corporation.

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

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Mok, P. P. K. (2008). On the syllable-timing of Cantonese and Beijing Mandarin.

Chinese Journal of Phonetics, 2, 148-154.

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology*, 2, 1-13.

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research*, 308, 98-108.

Roach, P. (2003). *Phonetics*. Shanghai: Shanghai Foreign Language Education Press.

Rosner, B. S., & Pickering, J. B. (1994). *Vowel Perception and Production*. Oxford: Oxford Science Publications.

Sluijter, A, & van Heuven, V. (1996). *Acoustic correlates of linguistic stress and accent in Dutch and American English*. Paper presented at the 4th International Conference on Spoken Language Processing, Philadelphia, USA, October 3-6.

Teschner, R. V., & Whitley, S., M. (2004). *Pronouncing English: A stress-based approach with CD-ROM*. Washington DC: Georgetown University Press.

Tong, X., Lee, S. M. K., Lee, M. M. L., & Burnham, D. (2015). A tale of two features: Perception of Cantonese lexical tone and English lexical stress in Cantonese-English bilinguals. *PLoS ONE*, 10(11): e0142896.

Traunmüller, H., & Eriksson, A. (1994). *The Frequency Range of the Voice*

Fundamental in the Speech of Male and Female Adults. Technical Report.

Linguistics Department, University of Stockholm. Stockholm, Sweden.

Underhill, A. (1998). *Sound Foundation: Living Phonology*. Oxford: Heinemann.

Wang, Q. (2008). *Perception of English Stress by Mandarin Chinese Learners of*

English: An Acoustic Study (Unpublished doctoral dissertation). British

Columbia: University of Victoria.

Yu, V. Y., & Andruski, J. E. (2010). A cross-language study of perception of lexical

stress in English. *Journal of Psycholinguistic Research*, 39, 323–344.

Zhang, Y., & Francis, A. (2010). The weighting of vowel quality in native and non-

native listeners' perception of English lexical stress. *Journal of Phonetics*, 38,

260-271.

Zheng, Y., & Samuel, A. G. (2018). The effects of ethnicity, musicianship, and tone

language experience on pitch perception. *Quarterly Journal of Experimental*

Psychology, 71(12), 2627-2642.