Applied Psycholinguistics **33** (2012), 623–641
doi:10.1017/S0142716411000506

Native and nonnative processing of Japanese pitch accent

XIANGHUA WU
Simon Fraser University

JUNG-YUEH TU
Indiana University

YUE WANG
Simon Fraser University

Received: November 27, 2009 Accepted for publication: December 29, 2010

ADDRESS FOR CORRESPONDENCE
Xianghua Wu, Department of Linguistics, Simon Fraser University, 9212 Robert C. Brown Hall Building, 8888 University Drive, Burnaby, BC V5A 1S6, Canada. E-mail: xianghua_wu@sfu.ca

ABSTRACT
The theoretical framework of this study is based on the prevalent debate of whether prosodic processing is influenced by higher level linguistic-specific circuits or reflects lower level encoding of physical properties. Using the dichotic listening technique, the study investigates the hemispheric processing of Japanese pitch accent by native Japanese listeners and two groups of nonnative listeners with no prior pitch accent experience but differing in their native language experience with linguistic pitch: native listeners of Mandarin (a tone language with higher linguistic functional use of pitch) and native listeners of English (a stress language with lower functional use of pitch). The overall results reveal that, for both native and nonnative listeners, the processing of Japanese pitch accent is less lateralized (compared to lexical tone processing, which has been found to be a left hemisphere property). However, detailed analysis with individual pitch accents across groups shows a right hemisphere preference for processing the high–accent–low (H*L) pattern, a left hemisphere preference for LH*, and no hemisphere dominance for LH, indicating a significant reliance on the acoustic cues. These patterns are particularly prominent with the English listeners who are least experienced with linguistic pitch. Together, the findings suggest an interplay of linguistic and acoustic aspects in the processing of Japanese pitch accent by native and nonnative listeners.

This study examines the role of linguistic experience in the perception and hemispheric processing of Japanese pitch accent by native Japanese listeners and two groups of nonnative listeners differing in their native language (L1) backgrounds with linguistic pitch: Mandarin Chinese and English.

BACKGROUND

Speech prosody functions at various linguistic domains (Baum & Pell, 1999; Van Lancker, 1980). At the lexical level, prosodic features can be superimposed on monosyllabic words to make phonemic contrasts (e.g., lexical tone), or on multisyllabic words to make phonemic (e.g., pitch accent) or grammatical (e.g., stress) contrasts. At the sentential level, they can be used as linguistic intonation to indicate sentence type (e.g., questions versus statements; Sadock & Zwicky, 1985). Prosodic features can also be realized in the paralinguistic domain, such as emotional intonation used to express happiness or anger. Thus, the perception and processing of linguistic prosody may involve multiple and hierarchical stages of acoustic, lexical, and sentential analysis (Cutler & Clifton, 1999; Gandour, Dzemidzic et al., 2003).

Previous studies have revealed complex hemispheric processing patterns for linguistic prosody, in that native prosodic processing may involve right hemisphere dominance[1] (Grimshaw, Kwasny, Covell, & Johnson, 2003; Zatorre & Samson, 1991), left hemisphere dominance (Arciuli & Slowiaczek, 2007; Gandour et al., 2002), or no hemisphere dominance (Gandour, Wong, et al., 2003; Mitchell & Crow, 2005). Research has since attempted to investigate the factors affecting the complex hemispheric processing of linguistic prosody.

Theoretical accounts

Linguistic function has been proposed to account for hemispheric asymmetry in the perception of prosody, as different linguistic features may carry different levels of functional load (e.g., Van Lancker, 1980). Functional load refers to the extent of contrastivity between linguistic units (e.g., distinctive features, phonemic opposition), as well as a measure of the number of minimal pairs for a given contrast, gauging the frequency with which two features contrast (King, 1967; Surendran & Niyogi, 2006). Based on this definition, lexical tone has a higher functional load than pitch accent, because all words in a tone language are contrastive for tone, whereas only approximately 20% of word pairs contrast for pitch accent (e.g., in Japanese; Pierrehumbert & Beckman, 1988). Likewise, tone also has a higher functional load than lexical stress, because stress is used to make grammatical contrasts and thus is less lexically contrastive (Cutler, 1986; Hallé, Chang, & Best, 2004). In contrast, the functional load of intonation is low as it is typically used at a more global level to indicate sentence types or emotional expressions (Cruttenden, 1997). In terms of hemispheric processing, a prosodic feature carrying high functional load (e.g., tone) tends to be lateralized in the left hemisphere, whereas a feature with low linguistic use (e.g., emotional intonation) tends to be lateralized in the right hemisphere. Those features falling somewhere in the middle of the linguistic functional hierarchy (e.g., pitch accent or stress) may involve a lesser degree of hemispheric dominance (Van Lancker, 1980).

Aside from the linguistic function, acoustic features, such as the temporal frame length of a prosodic unit, may determine the lateralization pattern of prosodic processing (e.g., Poeppel, 2001, 2003). According to this hypothesis, speech prosodic units over a shorter temporal domain (e.g., tone) tend to be left lateralized, whereas those with a longer temporal domain (e.g., sentential intonation) tend to call for greater right hemisphere participation. Presumably this is because the former mostly involves analytical processing of local information, whereas the latter involves a more holistic processing of global information (Bever, 1975). Likewise, hemispheric processing of pitch may also be affected by relative frequency, with the left hemisphere biased for high-frequency information and the right hemisphere biased for low-frequency information (Ivry & Lebby, 1993).

Moreover, the functional and acoustic aspects may complementarily account for lateralization patterns (Zatorre & Gandour, 2008). For instance, although tone and pitch accent are functionally similar (i.e., both used to make lexical distinctions), tone may involve a greater degree of left hemisphere processing than pitch accent due to its shorter temporal frame length. In contrast, although pitch accent and stress are used in comparable temporal domains, the processing of pitch accent may be more left lateralized than that of stress due to its higher functional load. Research has shown that linguistic and acoustic cues both contribute to the neural mechanisms underlying prosodic processing. Particularly, although shared neural substrates can be involved for different linguistic tasks depending on the amount and kind of acoustic cues available, competition between different linguistic domains may result in cortical competition (Zhao et al., 2008). For example, it has been found that although the perception of speech rhythm and intonation involved certain shared neural mechanisms due to their common acoustic properties, different brain regions were selectively more responsive to specific acoustic features as a function of listeners' linguistic experience (Zhang, Shu, Zhou, Wang, & Li, 2010).

Native processing

Empirical findings have not been consistent with respect to how these theoretical accounts are supported. In general, a left hemisphere superiority has been revealed in the native processing of tone in Mandarin (Klein, Zatorre, Milner, & Zhao, 2001; Wang, Sereno, & Jongman, 2001; Wang, Sereno, Jongman, & Hirsch, 2003) and Thai (Gandour, Dzemidzic, et al., 2003; Van Lancker & Fromkin, 1973). Left hemisphere dominance has also been found in prosodic processing over a multisyllabic domain, such as Norwegian tone (Berker & Reinvange, 2007; Moen, 1993; Moen & Sundet, 1996) and English stress (Arciuli & Slowiaczek, 2007; Baum, 2002; Shah & Baum, 2006). However, the native processing of Japanese pitch accent has shown noncompatible patterns. Using near-infrared spectroscopy, Sato, Sogabe, and Mazuka (2007) showed a left hemisphere dominance. In contrast, a magnetoencephalography study (Hayashi et al., 2001) revealed a significant negativity in the bilateral temporal cortices, when incorrect identification of a pitch accent prevents proper semantic comprehension. Likewise, findings on sentence-level prosodic features have been mixed. Although emotional intonation has revealed a right hemisphere dominance (Chernigovskaya et al., 2000), the processing of linguistic intonation has demonstrated left hemisphere (Chernigovskaya et al., 2000), right hemisphere (Shipley-Brown, Dingwall, Berlin, Yeni-Komshian, & Gordon-Salant, 1988), or bilateral mechanisms (Gandour, Dzemindzic, et al., 2003; Pihan, Tabert, Assuras, & Borod, 2008).

Nonnative processing

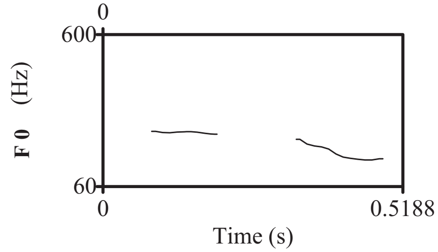
Research on nonnative processing has further addressed the extent to which linguistic and acoustic aspects of prosodic features influence hemispheric dominance. It has been found that nonnative listeners, such as English listeners of Thai or Mandarin, process tone bilaterally (Van Lancker & Fromkin, 1973; Wang et al., 2001), as they may focus more on subtle acoustic differences rather than distinguishing phonemic tone categories (Kaan, Wayland, Bao, & Barkley, 2007). Furthermore, these patterns retain even for nonnative listeners with tonal L1s, such as Mandarin listeners of Thai or Norwegian listeners of Mandarin, indicating that nonnative listeners do not process tone in a nativelike manner (i.e., left lateralized) despite their L1 experience with linguistic tone, presumably because the specific tone contrasts in the target language are not linguistically meaningful in their L1 (Gandour et al., 2002; Wang, Behne, Jongman, & Sereno, 2004). Likewise, native and nonnative intonation processing has also revealed different patterns. For example, the processing of intonation in Mandarin was bilateral for native Mandarin listeners but predominantly in the right hemisphere for English listeners (Gandour, Dzemidzic, et al., 2003).

Thus, it appears that nonnative listeners deviate from the native patterns with a lesser degree of left hemisphere processing, with the nature of deviation influenced by nonnative listeners' experience with their native and target prosody. Nevertheless, like the natives, nonnative listeners also show more left hemisphere involvement when processing prosody with a shorter temporal domain (e.g., tone) while exhibiting more right hemisphere activation in the processing of prosody with a longer temporal domain (e.g., intonation), pointing to the role of universal temporal acoustic factors on prosodic processing.

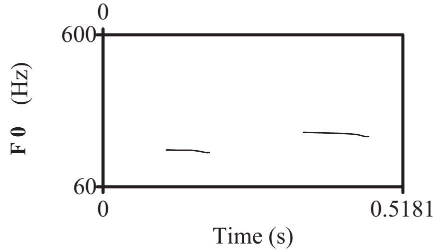
Taken together, the native and nonnative patterns show that the higher the functional load a prosodic feature carries or the shorter the temporal domain it is realized on, the more involvement of the left hemisphere. However, the lateralization of those prosodic patterns realized at an intermediate level along the functional and temporal scales, such as pitch accent, have been inconsistent (e.g., Hayashi et al., 2001; Sato et al., 2007). Furthermore, research has not examined nonnative hemispheric processing of pitch accent. Thus, the current investigation into native and nonnative pitch accent processing contributes to the overall picture of prosodic processing, given that the temporal domain on which pitch accent is realized as well as its functional load falls between the parameters of tone and intonation.

PITCH ACCENT PATTERNS

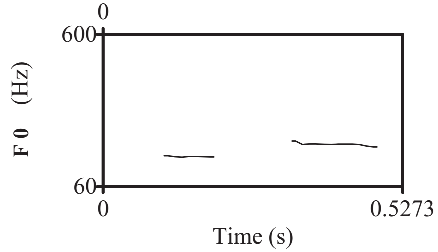
Pitch accent in Japanese is realized at a domain with at least two syllables to differentiate word meanings. Occurrence of high (H) or low (L) pitch as well as that of the accent (*) is predictable for unaccented and accented words. For an unaccented word, the first syllable has a low pitch, whereas the remaining syllables have a high pitch. For an accented word, one syllable is marked for accent. If the accent falls on the first syllable, the syllable has a high pitch and all the following syllables have a low pitch. If the accent falls on the second or later syllable, the first syllable has a low pitch and the syllables from the second until the accented one all have a high pitch (Kitahara, 2001; Sugiyama, 2006). Disyllabic words, therefore, have three pitch accent patterns: accented H^*L, LH^* patterns, and unaccented LH pattern, as shown in Figure 1.



hashi (H*L) ‘chopsticks’



hashi (LH*) ‘bridge’



hashi (LH) ‘edge’

Figure 1. The fundamental frequency (F0) contours of the three pitch accent patterns high–accent–low (H*L), LH*, and LH, exemplified by the syllable *hashi*.

THE CURRENT STUDY

In the current study, pitch accent processing patterns are compared between native listeners of Japanese and two nonnative groups: Mandarin listeners whose L1 uses pitch (tone) with a higher functional load than pitch accent, and English listeners with the functional use of pitch in their L1 (e.g., stress) being lower than pitch accent. Thus, the gradation of functional use of pitch in these L1s provides a useful testing ground to examine the effect of linguistic experience on pitch accent processing.

To assess hemispheric processing patterns, this study employs the dichotic listening paradigm, in which different stimuli in a pair are simultaneously presented to the left ear and the right ear (Kimura, 1961, 1967; Wang et al., 2001). During dichotic stimulation, a stimulus presented to the right ear can be more effectively processed in the left hemisphere than in the right hemisphere (and vice versa), because the information conveyed by the contralateral auditory pathways typically suppresses that by the ipsilateral pathways (Gazzaniga, 1984, p. 97). Consequently, a right ear advantage (REA) indicating left hemisphere dominance is often found in association with the processing of linguistic stimuli (e.g., Bryden & Murray, 1985; Dwyer, Blumstein, & Ryalls, 1982). Thus in this study, an REA would be expected if pitch accent were processed predominantly in the left hemisphere as linguistic information.

Based on the previous findings of native and nonnative prosodic processing (e.g., Gandour et al., 2002; Gandour, Ddzemidzic, et al., 2003; Wang et al., 2004), we hypothesize the following: (a) for pitch accent processing by native Japanese listeners, we expect a lesser degree of left hemisphere dominance as compared to native processing of lexical tone because pitch accent has a lower functional load and larger temporal frame length than tone; (b) compared to the native Japanese patterns, Mandarin listeners are expected to show an even lesser degree of left hemisphere dominance or bilateral processing as nonnative pitch contrasts are involved. However, compared to the English listeners, the Mandarin group may reveal a greater degree of left hemisphere involvement because of their experience with tone; and (c) English listeners are expected to demonstrate bilateral or even predominantly right hemisphere processing due to their lack of experience with phonemic pitch contrast with a higher functional load. Comparing across these native groups, the differences and commonalities in the processing patterns may reflect the extent to which linguistic and acoustic factors contribute to pitch accent processing.

METHOD

Participants

A total of 48 young adults participated in this experiment, including 16 native Japanese listeners (6 males, 10 females; mean age = 25 years, range = 19–46), 16 native Mandarin listeners (5 males, 11 females; mean age = 25 years, range = 20–33), and 16 native English listeners (7 males, 9 females; mean age = 25 years, range = 18–43). A one-factor analysis of variance (ANOVA) showed no significant difference in age variation for the three groups, F (2, 45) = .009, p = .991. The Japanese listeners had no tone language experience. The Mandarin and English listeners had no knowledge of Japanese or other pitch accent languages, and the latter group had no previous background with tone languages. Two participants in the Mandarin group reported speaking another tone language (Taiwanese or Cantonese), and the rest reported having no tone language background other than Mandarin. All participants were right-handed based on the Edinburgh Handedness Inventory (Oldfield, 1971) that they were required to complete prior to testing. None of the participants had any formal linguistic or musical training (i.e., less than 6 years' musical training; see Wong, Skoe, Russo, Dees, & Kraus, 2007). All reported normal speech and hearing. Participants were paid for their participation.

Stimuli

The testing stimuli included a total of 21 Japanese disyllabic words, consisting seven minimal triplets: three pitch accent patterns (H^*L, LH^*, LH) × seven syllables (*aki, hana, kaki, nami, take, tama, yuki*). Nine additional words (three from an additional triplet and six from three minimal pairs) were used as practice stimuli. All the words are commonly used in Japanese. Eighteen of them were adapted from Sugiyama (2006) as they had been selected due to a relatively high familiarity rate determined by a computerized dictionary (Amano & Kondo, 1999). The remaining 12 words were rated as high frequency by an online Japanese dictionary, Denshi Jisho (http://www.jisho.org/).

A female linguistically trained native speaker of Tokyo Japanese (aged 32) recorded four repetitions of all 30 words in a sound-attenuated recording booth in the Language and Brain Lab at Simon Fraser University, using Presonus Digital Audio 24 B27/96K Firewire Recording Interface and a Shure KSM 109 microphone. Each word was recorded followed by a monosyllabic particle, including が (-ga), を (-o), に (-ni), と (-to), and の (-no). This was to provide a phrasal context for the native speaker to naturally and accurately produce the distinctions among the pitch accent patterns, especially those between the accented and unaccented patterns (Maniwa, 2002).

Forty-two dichotic pairs (7 triplets × 6 pairing patterns) were created such that in each pair, the two words had the same segmental components but differed only in the pitch accent pattern, for example, H*L and LH* pairs such as *hana* (H*L) “a female name” and *hana* (LH*) “flower,” or LH* and LH pairs such as *hana* (LH*) “flower” and *hana* (LH) “nose.” These dichotic pairs were constructed and edited using Audacity 1.2.6 where one word in each pair was imported into the left channel and the other into the right channel. Each pair was normalized for intensity using Sound Forge 6.0 (Sonic Foundry, Inc., Madison, WI). The dichotic pairs were also selected (from the four repetitions) to have similar length, with the durational difference between each pair being under 10% (the just noticeable difference; Lehiste, 1970). The duration of the stimuli ranged from 444 to 581 ms (average = 533 ms).

Procedure

The experiment was conducted in a sound-attenuated perception booth in the Language and Brain Lab at Simon Fraser University, including three sections that were created and run separately on E-prime 1.0 (Psychology Software Tools, Inc., Sharpsburg, PA): a familiarization task, an identification test, and a dichotic listening test.

Following oral instructions, participants were familiarized for approximately 10 min with the three different pitch accent patterns using nine words, including one triplet and three pairs not used in the dichotic test. They were asked to identify these binaurally presented training stimuli and were given feedback for each response. After the familiarization section, the listeners took the identification test, which lasted 5 min, to ensure that they were able to distinguish and identify the three pitch accent patterns in the subsequent dichotic test. The participants were requested to identify the pitch accent patterns for 21 disyllabic words (3 pitch accent patterns × 7 disyllables), which were used in the dichotic test. These words were presented binaurally with no feedback given after each response. These 48 participants whose response accuracy was higher than 60%

(well above the chance level, 33%) continued on and took the dichotic listening test.

The dichotic listening test procedures were modeled after similar previous studies (e.g., Wang et al., 2001, 2004) using the two-response paradigm (Millay, Roeser, & Godfrey, 1977). The stimuli were randomized into four blocks (i.e., four repetitions), with 42 dichotic pairs in each block, resulting in 168 trials in total. Each pair was presented to the participants with one word in the left ear and the other in the right ear simultaneously. The participants were asked to identify both stimuli. To eliminate channel effects, the participants were requested to reverse the headphones across blocks, and thus headphone channels were counterbalanced. In addition, to avoid response order bias within participants, right ear and left ear responses were counterbalanced after two blocks. There were two versions of the dichotic listening test to further avoid order bias between participants: half of the participants in each group were asked to respond to the stimulus in their left ear first followed by that in their right ear, and then respond to the stimulus in their right ear first followed by that in their left ear, whereas for the other half of the participants, the order was reversed. The dichotic test for each participant lasted approximately 30 min.

RESULTS

Perceptual accuracy

Listeners' percent correct identification of pitch accent was analyzed using a four-factor repeated-measures ANOVA with group (Japanese, Mandarin, English) as the between-subjects variable, and ear (left, right), pitch accent pattern (H*L, LH*, LH) and syllable (*aki*, *hana*, *kaki*, *nami*, *take*, *tama*, *yuki*) as the within-subjects variables.

For the main effects, first, a marginally significant ear difference was obtained, F (1, 45) = 4.05, $p = .05$, with slightly greater identification accuracy in the left ear (55%) than in the right ear (53%), indicating a negligible left ear advantage (i.e., less-lateralized or bilateral processing). Moreover, a group difference did not reach significance, F (2, 45) = 1.61, $p = .211$. However, a significant effect was observed for pitch accent pattern, F (2, 90) = 26.89, $p < .0001$, with Bonferroni post hoc analyses indicating that the H*L pattern (66%) was more accurately identified than the LH* (50%, $p < .0001$) and LH (47%, $p < .0001$) patterns across groups, but there was no difference between the latter two pitch accent patterns ($p = 1$). The results also showed a significant main effect of syllable, F (6, 270) = 4.51, $p < .0001$, with Bonferroni post hoc analyses showing that pitch accent on syllable *aki* (50%) was more poorly identified than on the other syllables (*take*: 56%, $p = .002$; *hana*: 56%, $p = .003$; *tama*: 56%, $p = .023$; *kaki*: 53%; *yuki*: 54%; *name*: 53%).

Moreover, the results revealed significant interactions of ear and pitch accent pattern, F (4, 90) = 11.78, $p < .0001$, group and syllable, F (12, 270) = 2.24, $p = .011$, as well as group, pitch accent pattern, and syllable, F (24, 540) = 1.99, $p = .004$. No other significant interactions involving the main factor of ear or group were observed. Thus, further analyses were performed on the basis of the exhibited significant interactions between the independent variables.

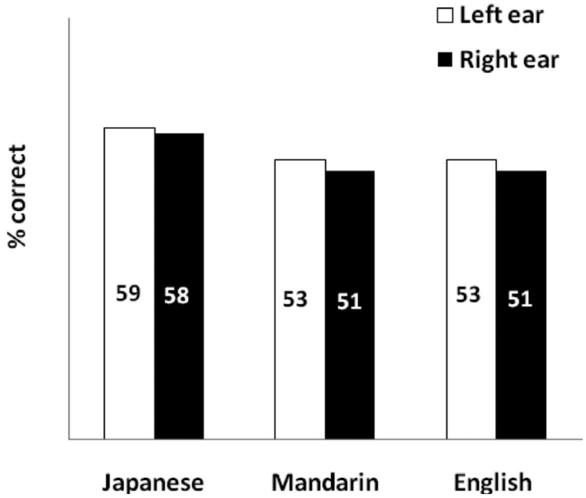


Figure 2. The percentage of correct identification in the left ear and right ear for Japanese, Mandarin. and English listeners.

Ear effect and native group. Because no significant interaction was obtained for ear and group, F (2, 45) = .359, p = .7, or ear, group and pitch accent, F (4, 90) = 1.33, p = .265, or ear, group, pitch accent, and syllable, F (24, 540) = .98, p = .496, no further analysis was performed for each of these interactions. These results revealed that the three groups did not differ in ear advantage in the perception of pitch accent (Figure 2).

Ear effect and pitch accent pattern. Based on the significant interaction of ear and pitch accent pattern reported above, sets of one-factor repeated-measures ANOVAs were conducted to investigate ear effect on the processing of each pitch accent pattern. The results showed a significant effect of ear for the H^*L pattern, F (1, 45) = 20.27, $p < .0001$, with the perception in the left ear (71%) being more accurate than in the right ear (61%) across groups. In contrast, for the LH^* pattern, the perception in the right ear (52%) was significantly more accurate than that in the left ear (48%), F (1, 45) = 4.61, p = .037. For LH, no difference in ear advantage was observed, F (1, 45) = .58, p = .45. Figure 3 illustrates the patterns of ear advantage for each pitch accent pattern. Furthermore, one-factor repeated-measures ANOVAs were performed for each ear using pitch accent pattern as the within-subjects factor. This analysis indicated significant effects of pitch accent pattern for the left ear, F (2, 94) = 42.14, $p < .0001$, and the right ear, F (2, 94) = 6.81, p = .002. Consistent with the across-ear results, the post hoc tests (Bonferroni adjusted) further showed that H^*L was more accurately perceived than LH^* and LH for both the left ear (H^*L 71% > LH^* 48%, $p < .0001$; H^*L 71% > LH 47%, $p < .0001$) and the right ear (H^*L 61% > LH^* 52%, p = .02; H^*L 61% > LH 48%, p = .003), whereas no difference between the LH^* and LH patterns was found (left ear: p = 1; right ear: p = .944).

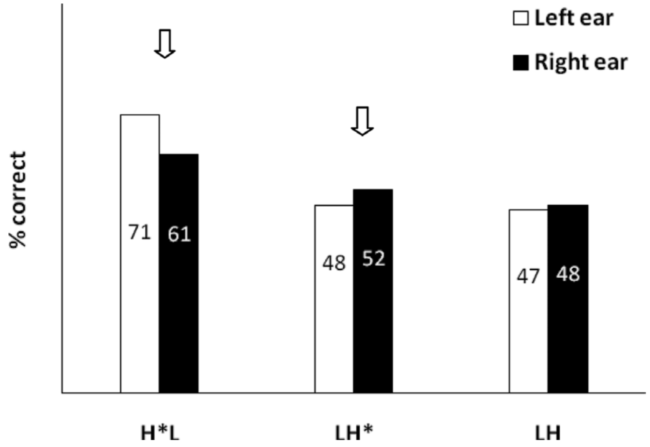


Figure 3. The percentage of correct identification in the left ear and right ear for the high–accent–low (H*L), LH*, and LH patterns across native groups. The arrow (⇩) indicates statistical significance at $p < .05$.

Syllable effect and native group. To further analyze the above-reported main effect and interactions involving syllable, sets of one-factor ANOVAs were conducted for each group using syllable as the within-subjects factor. The results showed that the syllable effect existed only in the Japanese group, F (6, 90) = 7, $p < .0001$, with *aki* (51%) being more poorly identified than the other syllables (*name*: 56%; *yuki*: 59%; *kaki*: 59%; *tama*: 61%; *hana*: 62%; particularly *take*: 65%, $p < .001$). More detailed analysis with the Japanese group revealed consistent patterns across pitch accent patterns.

Distribution of ear preference

In addition to mean perceptual accuracy, data were also examined in terms of frequency, that is, the number of listeners showing each of the three different types of ear preference: left ear advantage (LEA), REA, or no ear advantage (NEA). This was performed using Pearson's chi-square (χ^2) analysis. A three-way contingency table was created in SPSS (SPSS Inc., Chicago, IL) with ear preference as the column variable, pitch accent pattern as the row variable, and native group as the layer variable.

The results indicated a significant association between pitch accent pattern and the number of listeners showing LEA, REA, or NEA only for the English group, χ^2 (4) = 18.32, $p = .001$, but not the Japanese, χ^2 (4) = 4.85, $p = .303$, or Mandarin group, χ^2 (4) = 3.18, $p = .528$. Thus, further analysis was performed to examine only the English group's distribution of ear preference in the processing of individual pitch accent patterns. As illustrated in Figure 4, for the H*L pattern, more English listeners showed LEA (15) than those showing REA (1), χ^2 (1) = 12.25, $p < .0001$. In contrast, for the LH* pattern, the REA listeners (12) outnumbered the LEA listeners (4), χ^2 (1) = 4, $p = .046$. No English listeners

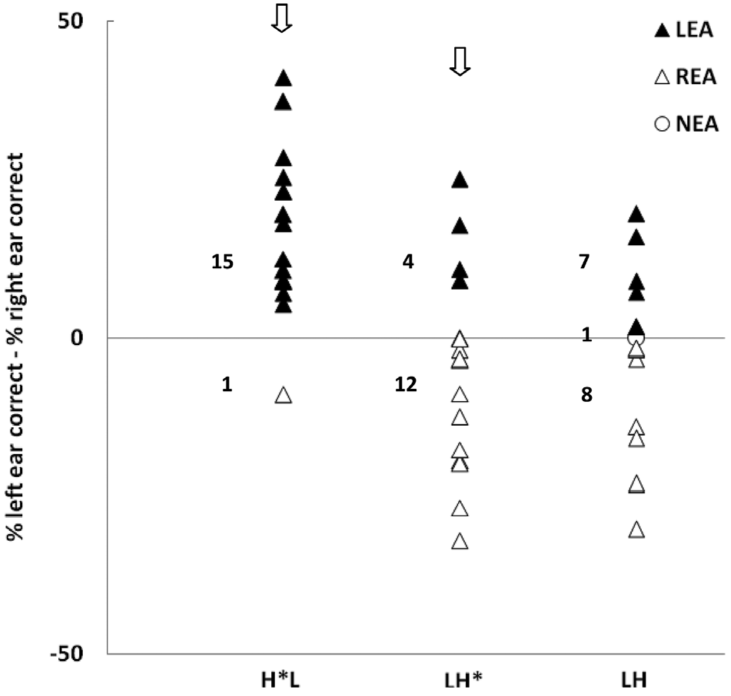


Figure 4. The number of listeners showing left ear advantage (LEA▲), right ear advantage (REA△), or no ear advantage (NEA○) for the English group ($n = 16$) in the perception of the high–accent–low (H*L), LH*, and LH patterns. The numbers indicate the occurrences of left ear advantage (LEA; above the zero-crossing line), right ear advantage (REA; below the zero-crossing line), and no ear advantage (NEA; at the zero-crossing line). The arrow (⇩) indicates statistical significance at $p < .05$.

exhibited NEA for these two pitch accent patterns. For the LH pattern, there was no significant difference in the distribution of ear preference, χ^2 (2) = 5.38, $p = .068$.

DISCUSSION

The ear preference analyses of the perceptual accuracy data indicate that all three groups tended to be less lateralized in the overall processing across three pitch accent patterns. Despite this general pattern, the results for individual pitch accent patterns varied, showing an LEA (i.e., right hemisphere dominance) for the H*L pattern, a REA (i.e., left hemisphere dominance) for the LH* pattern and NEA (i.e., no hemisphere dominance) for the LH pattern. In addition, a comparison of the pitch accent patterns across ears showed that the native and nonnative groups all identified the H*L pattern more accurately than the other two patterns. However, group difference was observed with the distribution of ear preference analysis. That is, only the English group, but not the Japanese or Mandarin group, demonstrated that the number of listeners for each hemispheric dominance pattern differed as a function of pitch accent patterns. Specifically, more English listeners showed right hemisphere dominance when processing the H*L pattern, whereas more of them showed left hemisphere dominance for the LH* pattern. In the following discussion, these results are interpreted in relation to the proposed hypotheses for the native and nonnative listeners in terms of how the acoustic properties of the target prosody interacting with linguistic functions and experience affect lateralization of Japanese pitch accent.

Native processing

Overall results. That the native Japanese listeners showed a lesser degree of lateralization in processing overall pitch accent patterns differs from previous native linguistic tone processing findings, which have shown a strong and consistent left hemisphere dominance (e.g., Gandour et al., 2002; Wang et al., 2001). This finding is consistent with Hayashi et al. (2001), claiming that the results of bilateral processing for pitch accent was presumably due to its lighter linguistic functional use compared to tone.

The results support the linguistic functional hypothesis (Gandour, Dzemidzic, et al., 2003; Van Lancker, 1980), which predicts a lesser degree of left-hemisphere dominance for those linguistic contrasts with a lower functional load. In this case, Japanese pitch accent has a lower functional load than tone, because pitch accent is less widely used (Pierrehumbert & Beckman, 1988) compared to tone, which is contrastive on every word in tone languages (Chao, 1948). Furthermore, these results may be accounted for by the acoustic hypothesis (Poeppel, 2001, 2003), in that pitch accent superimposed on disyllabic words has a larger temporal frame length compared to tone, which is typically superimposed on monosyllables. As the hypothesis posited, those linguistic domains with a longer temporal frame length tend to be processed more holistically and are thus less left lateralized, whereas those with a shorter temporal frame involve a greater degree of analytic processing which is primarily a left-hemisphere property (Bever, 1975; Gandour, Dzemidzic, et al., 2003; Poeppel, 2003).

However, the current results were inconsistent with Sato et al. (2007), finding a left-hemisphere specialization for native Japanese processing of pitch accent patterns H*L and LH*. This discrepancy may be explained by the different levels of task difficulty. Whereas only the H*L and LH* patterns were tested in Sato et al. (2007), the current study included all three patterns, especially the confusable pair LH* and LH, consequently increasing the task difficulty. Previous research has shown greater engagement of the right hemisphere with an increase in task difficulty (Aasland & Baum, 2003). Thus, pitch accent processing may have involved greater right-hemisphere activities compared to that in Sato et al. (2007).

Individual pitch accent patterns. Despite the lack of significant ear preference across pitch accent patterns, individual pitch accent pattern analysis did reveal different tendencies of lateralization: right hemisphere dominance for H*L, whereas left hemisphere dominance for LH* and no hemisphere dominance for LH. Given that the pitch accent triplets or pairs in the current study were selected to be of comparable frequency of use and they were all at the same level of contrastivity (i.e., the pitch accent level), it is unlikely that linguistic function played a role in these different processing patterns. On the other hand, previous findings on the acoustic processing of pitch patterns may lend some support to account for the different patterns. For example, Ivry and Lebby (1993) suggest that the left hemisphere is biased for processing (relatively) higher frequency information, whereas the right hemisphere is biased for processing lower frequency information. In addition, Walsh (1996) showed that the perception of pitch accent could not be accurately determined until the second syllable, indicating a prominent role of the ending frequency. In the case of pitch accent patterns, H*L ends with a lower frequency, whereas LH* and LH end with a higher frequency. Thus, it is speculated that when focusing on the second syllable, H*L (ending with L) was more right hemisphere-biased than LH* and LH (ending with H).

Moreover, the degree of stimulus perceptual confusion may have influenced the lateralization of individual pitch accent patterns. The across-ears results showed that LH* and LH were more poorly perceived than H*L. As previously revealed, due to their subtle acoustic distinctions in F0 maximum (Sugito, 1983; Vance, 1995), LH* and LH are difficult to distinguish and thus poorly perceived (Maniwa, 2002; Sugiyama, 2006). In terms of hemispheric processing, it has been found from tone studies that poorly perceived tones tend to show a greater degree of left hemisphere involvement (Wang et al., 2004). The current results consistently revealed this pattern, with LH* and LH involving more left-hemisphere processing than H*L.

Although further research may be necessary to test these speculations, the current results of different processing for individual pitch accent patterns suggest that future studies should not just treat a linguistic property (such as pitch accent or tone) as a single entity. Individual patterns within the same linguistic domain may involve different processing patterns due to their acoustic differences (e.g., H*L vs. LH* pitch accent patterns, or falling tone vs. rising tone).

Syllable effect. The perceptual accuracy rate across ears for each syllable indicated that the syllable *aki* was more poorly identified than the other syllables by the Japanese listeners. This may be because the familiarity ratings for the words with the *aki* syllable were relatively lower than other words (Amano & Kondo, 1999; Sekiguchi, 2006). Thus, the lower accuracy rate for the pitch accent processing of the *aki* words might result from a lower level of familiarity with *aki*. That this syllable effect was shown only in the Japanese group was conceivable, because none of these words were meaningful for the nonnative listeners and were thus devoid of any familiarity effect.

Nonnative processing

Overall results. The accuracy data revealed that both the Mandarin and English listeners showed a less-lateralized pattern when processing Japanese pitch accent across the three patterns, just as the native Japanese group did. These results are consistent with the previous findings of nonnative tone processing in that nonnative prosodic features are not processed as linguistically significant contrasts typically specialized in the left hemisphere (Gandour et al., 2002; Van Lancker & Fromkin, 1973; Wang et al., 2001, 2004).

The current results of similar native and nonnative patterns also suggest the involvement of acoustic processing for native and nonnative listeners alike, with a larger temporal domain resulting in a lesser degree of left hemisphere involvement (Bever, 1975; Gandour, Dzemidzdic, et al., 2003; Poeppel, 2003). These common patterns across the native and nonnative groups suggest that the perception of pitch accent may involve more acoustic processing than linguistic processing. Previous studies argued that in the processing of speech, listeners may integrate different levels of acoustic and linguistic cues depending on cue availability and stimulus properties (Zhang et al., 2010; Zhao et al., 2008). For example, although cortical competition may occur as a function of the competition of information from different linguistic (and nonlinguistic) domains (Zhao et al., 2008), cortical overlap may reflect common processing of acoustic properties associated with different linguistic dimensions (Zhang et al., 2010). In the current study, that the listeners with different language backgrounds showed common patterns may indicate a lesser degree of linguistic influence. Because of the low linguistic contrastivity (functional load) for pitch accent, its processing was conceivably more associated with the subtle acoustic properties, resulting in similar processing patterns for native and nonnatives.

Group-specific patterns. Despite these common patterns across groups, group-specific patterns were also evident. From the distribution of ear preference data, ear effect for individual pitch accent patterns was only found for the English group, with more right-hemisphere-biased listeners for the H*L pattern, whereas more left-hemisphere-biased ones for LH*. These patterns were consistent with the common patterns across groups from the perceptual accuracy results. That more English listeners' processing patterns were sensitive to pitch pattern difference indicates that the perception of Japanese pitch accent involved even greater degree of acoustic rather than linguistic processing for the English listeners, compared to the Japanese and Mandarin listeners. These different processing patterns may be accounted for by the influence of the nonnative listeners' prior prosodic experience, as the degree of functional load of linguistic prosody in English was lower than in Japanese or Mandarin (Van Lancker, 1980).

GENERAL DISCUSSION

The theoretical framework of this research addresses a long-deliberated issue in speech processing, regarding the extent to which speech processing is influenced by higher level linguistic-specific circuits, or reflecting lower level encoding of physical properties (Gandour et al., 2004; Poeppel, 2003; Zatorre & Gandour, 2008). The current findings from the native and nonnative groups suggest combined effects of linguistic representations and acoustic sensitivities on the processing of linguistic prosody.

In terms of linguistic function, the results support the previous finding of an experience-dependent processing of prosody, where prosody with a lower functional load involves a lesser degree of left hemisphere dominance (Gandour, Dzemidzic, et al., 2003; Van Lancker, 1980). That the native processing of Japanese pitch accent was overall less lateralized could be attributed to the low functional use of pitch accent in Japanese (compared to tone). Likewise, linguistic functional load may differentially affect nonnative processing patterns. As discussed earlier, listeners whose L1 (e.g., English) prosody has a lower functional load than that of the target prosody tend to rely more on acoustic cues in the processing of the target prosody than those from a language (e.g., Mandarin) with a greater functional use of prosody. These native and nonnative patterns consistently suggest that prosodic processing is altered by linguistic experience as indexed by the weight of functional load.

In contrast, acoustic properties such as the temporal frame length are also found to influence hemispheric processing, with a prosodic feature over a longer temporal domain being more right-hemisphere lateralized, whereas one over a shorter domain more left-hemisphere dominant (e.g., Poeppel, 2003). Following this account, the current Japanese listeners' overall less-lateralized processing could be due to the (longer) disyllabic context in which pitch accent appears (compared to the monosyllabic context for tone). These results echo the previous findings of a left-hemisphere preference for monosyllabic tone processing, whereas a right-hemisphere dominance for intonations imposed on trisyllabic sentences (Gandour et al., 2003). It is thus conceivable that pitch accent was overall less lateralized as its temporal frame length falls in between tone and intonation. Moreover, the contribution of acoustic aspects in pitch accent processing has also been revealed with the nonnatives, particularly the English group who demonstrated greater sensitivity to the acoustic differences among the pitch accent patterns.

The current results suggest that the processing of pitch accent for the native and nonnative listeners alike may have involved more acoustic than linguistic processing. First, the relatively low functional load of pitch accent in Japanese may have forced them to rely more on acoustic information, given that the linguistic cues are not prominent (Zhao et al., 2008). Moreover, the different laterality patterns for individual pitch accents as revealed by the perceptual accuracy data indicates that both native and nonnative processing may have been sensitive to subtle acoustic properties (Ivry & Lebby, 1993).

Taken as a whole, this study, along with the previous ones (e.g., Gandour et al., 2004; Gandour, Dzemidzic, et al., 2003; Gandour, Wong, et al., 2003), suggests that processing of speech prosody does not just involve a single process, a single linguistic or physical domain, or a particular hemisphere, but may rather be the integration of multiple levels and processes.

CONCLUDING REMARKS

The current study supports the prevalent hypothesis on the combined linguistic and acoustic effects on prosodic processing (Zatorre & Gandour, 2008). For further research to examine the interactions of functional load and acoustic properties in prosodic processing, a comparison of nonnative processing of stress and pitch accent could be informative. The disyllabic pitch accent and stress may have the same temporal domain, but different levels of functional load. Thus, comparing the native and nonnative processing of pitch accent and stress, along with the previous findings on nonnative tone and intonation processing, can shed light on our understanding of the interaction of linguistic and acoustic aspects in hemispheric processing of prosodic features. Moreover, different processing patterns of individual pitch accent patterns imply that future studies of hemispheric processing of a linguistic property should involve a more detailed examination of different patterns of the property, for example, to separately analyze rising versus falling tones for tone processing, or H*L versus LH* for pitch accent patterns. Because individual patterns of prosody may not be processed in the same way, the acoustic processing involving individual linguistic patterns needs to be taken into account. Furthermore, this research with naive nonnative listeners of pitch accent provides a foundation for further studies on the role of linguistic experience on the processing of pitch accent with learners of Japanese or bilinguals examining the extent to which the processing patterns differ as a function of proficiency in Japanese.

Thus, to converge evidence in unraveling the neural processing of linguistic prosody, future studies should take into account different acoustic and linguistic domains, different levels of linguistic properties, as well as listeners with diversified language proficiency levels and backgrounds.

ACKNOWLEDGMENTS

All three authors made equal contributions to this study. This research was supported by a Discovery Research grant (312457–2006) from the Natural Sciences and Engineering Research Council of Canada (NSERC). We acknowledge Amalia Dray for her contribution to the pilot study of this research, and Angela Cooper, Andrea Lalic, Tia Rad, Yasuko Sakurai, and Rebecca Simms from Simon Fraser University for their assistance. We thank Allard Jongman, Dawn Behne, Kenneth de Jong, Kenneth Hugdahl, Nancy Hedberg, and Yukiko Sugiyama for their feedback and discussions. Special thanks to Lihui Zhao at Harvard University for assistance in statistical analysis. Thanks also to an Associate Editor and the two anonymous reviewers for their comments and suggestions. Portions of this research were published in the Proceedings of the 2008 Canadian Acoustical Association Annual Conference and presented at the 2008 Canadian Acoustical Association Annual Conference in Vancouver, the 156th Acoustical Society of America Meeting in Miami, and the 2th International Conference on East Asian Linguistics in Vancouver.

NOTE

1. Note that “dominance” refers to a greater degree of involvement in one hemisphere than the other one. It does not exclude involvement of the other hemisphere (cf. Wang et al., 2004).

REFERENCES

Aasland, W., & Baum, S. (2003). Temporal parameters as cues to phrasal boundaries: A comparison of processing by left-hemisphere-damaged and right-hemisphere-damaged individuals. *Brain and Language, 87*, 385–399.

Amano, S., & Kondo, T. (1999). *Nihongo-no goitokusei* [Lexical properties of Japanese]. Tokyo: Sanseido.

Arciuli, J., & Slowiaczek, L. M. (2007). The where and when of linguistic word-level prosody. *Neuropsychologia, 45*, 2638–2642.

Baum, S. (2002). Word recognition in individuals with left and right hemisphere damage: The role of lexical stress. *Applied Psycholinguistics, 23*, 233–246.

Baum, S., & Pell, M. (1999). The neural bases of speech prosody: Insights from lesion studies and neuroimaging. *Aphasiology, 13*, 581–608.

Berker, F., & Reinvang, I. (2007). Mismatch negativity elicited by tones and speech sounds: Changed topographical distribution in aphasia. *Brain and Language, 100*, 69–78.

Bever, T. G. (1975). Cerebral asymmetries in humans are due to the differentiation of two incompatible processes: Holistic and analytic. *Annals of the New York Academy of Sciences, 163*, 251–262.

Bryden, M. P., & Murray, J. E. (1985). Toward a model of dichotic listening performance. *Brain and Cognition, 4*, 241–257.

Chao, Y. R. (1948). *Mandarin primer: An intensive course in spoken Chinese*. Cambridge, MA: Harvard University Press.

Chernigovskaya, T. V., Svetozarova, N. D., Tokareva, T. I., Tret'yakov, D. A., Ozerskii, P. V., & Strel'nikov, K. N. (2000). Specialization of cerebral hemispheres in the perception of Russian intonations. *Human Physiology, 26*, 142–147 (Translated from *Fiziologiya Cheloveka, 26*, 24–29).

Cruttenden, A. (1997). *Intonation*. Cambridge: Cambridge University Press.

Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech, 29*, 201–220.

Cutler, A., & Clifton, C. (1999). Comprehending spoken language: A blueprint of the listener. In C. Brown & P. Hagoort (Eds.), *The neurocognition of language*. New York: Oxford University Press.

Dwyer, J., Blumstein, S. E., & Ryalls, J. (1982). The role of duration and rapid temporal processing on the lateral perception of consonants and vowels. *Brain and Language, 17*, 272–286.

Gandour, J., Wong, D., Dzemidzic, M, Lowe, M., Tong, Y., & Li, X. (2003). A cross-linguistic fMRI study of perception of intonation and emotion in Chinese. *Human Brain Mapping, 18*, 149–157.

Gandour, J., Dzemidzic, M., Wong, D., Lowe, M., Tong, Y., Hsieh, L, et al. (2003). Temporal integration of speech prosody is shaped by language experience: An fMRI study. *Brain and Language, 84*, 318–336.

Gandour, J., Tong, Y., Wong, D., Talavage, T., Dzemidzic, M., Xu, Y, et al. (2004). Hemispheric roles in the perception of speech prosody. *NeuroImage, 23*, 344–357.

Gandour, J., Wong, D., Lowe, M., Dzemidzic, M., Satthamnuwing, N., Tong, Y., et al. (2002). A cross-linguistic fMRI study of spectral and temporal cues underlying phonological processing. *Journal of Cognitive Neuroscience, 14*, 1076–1087.

Gazzaniga, M. S. (Ed.). (1984). *Handbook of cognitive neuroscience*. New York: Plenum Press.

Grimshaw, G., Kwasny, K. M., Covell, E., & Johnson, R. A. (2003). The dynamic nature of language lateralization: Effects of lexical and prosodic factors. *Neuropsychologia, 41*, 1008–1019.

Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics, 32*, 395–421.

Hayashi, R., Imaizumi, S., Mori, K., Niimi, S., Ueno, S., & Kiritani, S. (2001). Elicitation of N400m in sentence comprehension due to lexical prosody incongruity. *NeuroReport, 12*, 1753–1756.

Ivry, R., & Lebby, P. (1993). Hemispheric differences in auditory perception are similar to those found in visual perception. *Psychological Science, 4*, 41–45.

Kaan, E., Wayland, R., Bao, M., & Barkley, C. M. (2007). Effects of native language and training on lexical tone perception: An event-related potential study. *Brain Research, 1148*, 113–122.

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology, 15*, 166–171.

Kimura, D. (1967). Functional asymmetry of the brain in dichotic listening. *Cortex, 3*, 163–178.

King, R. (1967). Functional load and sound change. *Language, 43*, 831–852.

Kitahara, M. (2001). *Category structure and function of pitch accent in Tokyo Japanese*. PhD dissertation, Indiana University.

Klein, D., Zatorre, R. J., Milner, B., & Zhao, V. (2001). A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *NeuroImage, 13*, 646–653.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Maniwa, K. (2002). Acoustic and perceptual evidence of complete neutralization of word-final tonal specification in Japanese. *Kansas Working Papers in Linguistics, 26*, 93–112.

Millay, K., Roeser, R. J., & Godfrey, J. J. (1977). Reliability of performance for dichotic listening using two response modes. *Journal of Speech and Hearing Research, 20*, 510–518.

Mitchell, R. L. C., & Crow, T. J. (2005). Right hemisphere language functions and schizophrenia: The forgotten hemisphere? *Brain, 128*, 963–978.

Moen, I. (1993). Functional lateralization of the perception of Norwegian word tones—Evidence from a dichotic listening experiment. *Brain and Language, 44*, 400–413.

Moen, I., & Sundet, K. (1996). Production and perception of word tones (pitch accents) in patients with left and right hemisphere damage. *Brain and Language, 53*, 267–281.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97–113.

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

Pihan, H., Tabert, M., Assuras, S., & Borod, J. (2008). Unattended emotional intonations modulate linguistic prosody processing. *Brain and Language, 105*, 141–147.

Poeppel, D. (2001). Pure word deafness and the bilateral processing of the speech code. *Cognitive Science, 25*, 679–693.

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as “asymmetric sampling in time.” *Speech Communication, 41*, 245–255.

Sadock, J. M., & Zwicky, A. M. (1985). Speech acts distinctions in syntax. In T. Shopen (Ed.), *Language typology and syntactic description*. Cambridge: Cambridge University Press.

Sato, Y., Sogabe, Y., & Mazuka, R. (2007). Brain responses in the processing of lexical pitch-accent by Japanese speakers. *NeuroReport, 18*, 2001–2004.

Sekiguchi, T. (2006). Effects of lexical prosody and word familiarity on lexical access of spoken Japanese words. *Journal of Psycholinguistic Research, 35*, 369–384.

Shah, A. P., & Baum, S. R. (2006). Perception of lexical stress by brain-damaged individuals: Effects on lexical–semantic activation. *Applied Psycholinguistics, 27*, 143–156.

Shipley-Brown, F., Dingwall, W. O., Berlin, C. I., Yeni-Komshian, G., & Gordon-Salant, S. (1988). Hemispheric processing of affective and linguistic intonation contours in normal subjects. *Brain and Language, 33*, 16–26.

Sugito, M. (1983). Nihongo no akusento to intoneeshon—Tokyo hogen no “hana” to “hana” no soui [Japanese accent and intonation—The difference between “hana ‘flower’” and “hana ‘nose’” in the Tokyo dialect]. In M. Sugito (Ed.), *“hana” to “hana” Nihongo no kenkyu 5*. Osaka: Izumishoin. [Reprinted from Kotoba to onsei, “kotoba” Series 18]

Sugiyama, Y. (2006). Japanese pitch accent: Examination of final-accented and unaccented minimal pairs. *Toronto Working Papers in Linguistics, 26*, 73–88.

Surendran, D., & Niyogi, P. (2006). Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. In O. Nedergaard Thomsen (Ed.), *Competing models of language change: Evolution and beyond*. Amsterdam: John Benjamins.

Van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *Papers in Linguistics, 13*, 201–277.

Van Lancker, D., & Fromkin, V. (1973). Hemispheric specialization for pitch and tone: Evidence from Thai. *Journal of Phonetics, 1*, 101–109.

Vance, T. J. (1995). Final accent vs. no accent: Utterance-final neutralization in Tokyo Japanese. *Journal of Phonetics, 23*, 487–499.

Walsh, D. L. (1996). Limiting-domains in lexical access: Processing of lexical prosody. In M. Dickey & S. Tunstall (Eds.), *University of Massachusetts occasional papers in linguistics 19: Linguistics in the laboratory*. Amherst: GLSA.

Wang, Y., Behne, D., Jongman, A., & Sereno, J. A. (2004). The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Psycholinguistics, 25*, 449–466.

Wang, Y., Sereno, J. A., & Jongman, A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language, 78*, 332–348.

Wang, Y., Sereno, J. A., Jongman, A., & Hirsch, J. (2003). fMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of Cognitive Neuroscience, 15*, 1–9.

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10*, 420–422.

Zatorre, R. J., & Gandour, J. T. (2008). Neural specializations for speech and pitch: Moving beyond the dichotomies. *Philosophical Transactions of the Royal Society, 363*, 1087–1104.

Zatorre, R. J., & Samson, S. (1991). Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain, 114*, 2403–2417.

Zhang, L., Shu, H., Zhou, F., Wang, X., & Li, P. (2010). Common and distinct neural substrates for the perception of speech rhythm and intonation. *Human Brain Mapping, 7*, 1106–1116.

Zhao, J., Shu, H., Zhang, L., Wang, X., Gong, Q., & Li, P. (2008). Cortical competition during language discrimination. *NeuroImage, 43*, 624–633.