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,

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 nance. It has been found that nonnative listeners, such as English listeners of

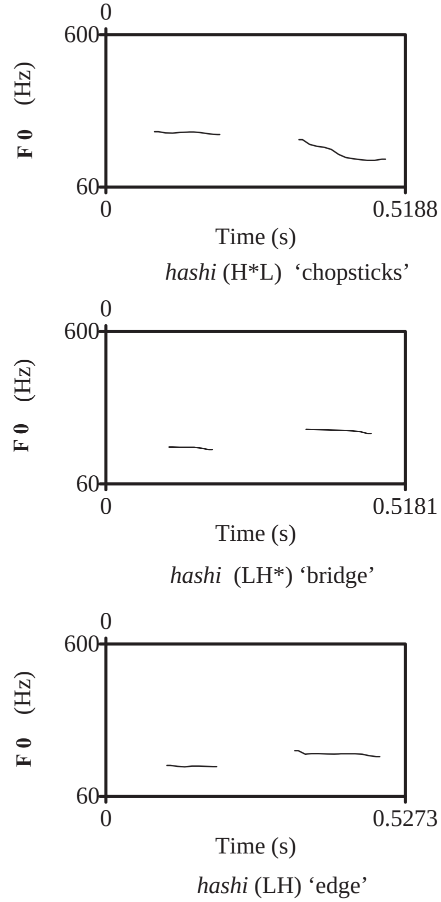

Figure 1. The fundamental frequency (F0) contours of the three pitch accent patterns high–accent–low (H*L), LH*, and LH, exemplified by the syllable *hashi*.

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, 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 (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

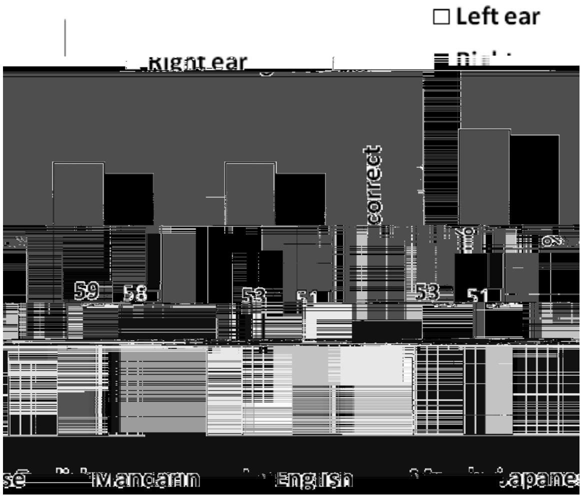

Figure 2. The percentage of correct identification in the left ear and right ear for Japanese, Mandarin. and English listeners.

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.0001;

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

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 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, HꜜL ends with a lower frequency, whereas LHꜜ and LH end with a higher frequency. Thus, it is speculated that when focusing on the second syllable, HꜜL (ending with L) was more right hemisphere-biased than LHꜜ 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 LHꜜ and LH were more poorly perceived than HꜜL. As previously revealed, due to their subtle acoustic distinctions in F0 maximum (Sugito, 1983; Vance, 1995), LHꜜ 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 LHꜜ and LH involving more left-hemisphere processing than HꜜL.

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., HꜜL vs. LHꜜ 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,

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*

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.

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.