

Effects of acoustic and linguistic experience on Japanese pitch accent processing*

XIANGHUA WU

University of California, Berkeley, U.S.A.

SAYA KAWASE

University of Western Sydney, Australia

YUE WANG

Simon Fraser University, Canada

(Received: July 13, 2015; final revision received: April 3, 2016; accepted: April 3, 2016; first published online 10 May 2016)

This study investigated the effects of L2 learning experience in relation to L1 background on hemispheric processing of Japanese pitch accent. Native Mandarin Chinese (tonal L1) and English (non-tonal L1) learners of Japanese were tested using dichotic listening. These listener groups were compared with those recruited in Wu, Tu & Wang (2012), including native Mandarin and English listeners without Japanese experience and native Japanese listeners. Results revealed an overall right-hemisphere preference across groups, suggesting acoustically oriented processing. Individual pitch accent patterns also revealed pattern-specific laterality differences, further reflecting acoustic-level processing. However, listener group differences indicated L1 effects, with the Chinese but not English listeners approximating the Japanese patterns. Furthermore, English learners but not naïve listeners exhibited a shift towards the native direction, revealing effects of L2 learning. These findings imply integrated effects of acoustic and linguistic aspects on Japanese pitch accent processing as a function of L1 and L2 experience.

Keywords: Japanese pitch accent, hemispheric processing, linguistic and acoustic experience

1. Introduction

Research has shown that hemispheric dominance of speech prosody (e.g., lexical tone, stress, intonation) can be determined by both lower-domain auditory-acoustic and higher-domain linguistic processes (Mildner, 2004; Van Lancker, 1980; Wang, Behne, Jongman & Sereno, 2004; Zhao, Shu, Zhang, Wang, Gong & Li, 2008). In addition, hemispheric processing patterns for native and non-native prosody may alter as a function of both native language (L1) and non-native language (L2) experience with linguistic pitch, as well as experience with non-speech acoustic properties of pitch (Gandour, Dziedzic, Wong, Lowe, Tong, Hsieh, Saththamnuwong & Lurito,

2003b; Hayashi, Imaizumi, Mori, Niimi, Ueno & Kiritani, 2001; Klein, Zatorre, Milner & Zhao, 2001; Tong, Gandour, Talavage, Wong, Dziedzic, Xu, Li & Lowe, 2005; Van Lancker & Fromkin, 1973; Wang et al., 2004; Wong, Parsons, Martinez & Diehl, 2004). The current study investigated hemispheric processing of Japanese disyllabic pitch accent by non-native listeners differing in their L2 experience with pitch accent (learners versus non-learners of Japanese) as well as L1 experience with lexical pitch (tonal versus non-tonal L1 listeners). The goal was to explore the extent to which non-native pitch accent processing is affected by L2 learning experience coupled with L1 background and the extent to which it is mediated through lower-level sensory-acoustic processes.

* Xianghua Wu and Yue Wang have made equal contributions to this paper. This research was conducted in the Language and Brain Lab at Simon Fraser University (SFU) and was funded by research grants to Yue Wang from the Natural Sciences and Engineering Research Council of Canada [NSERC Discovery Grants 312457-2006, 2011]. We thank Sam Gamble, Alison Kumpula, Chris Lightfoot, and Natalia Stratul from the SFU Language and Brain Lab for their assistance in stimuli development, data collection and analysis, and Jung-yueh Tu for her contributions to the first stage of this study. We also thank Drs. Alexander Francis, Kenneth de Jong, Allard Jongman, Joan Sereno, and Yuwen Lai for their valuable input at various stages of this project. Portions of this research were presented at the 164th Acoustical Society of America Meeting in Kansas City. Data from the naïve English and Chinese listeners as well as the native Japanese listeners have previously been reported in Wu, Tu, and Wang (2012).

1.1. Theoretical framework and empirical findings

Two alternative hypotheses regarding the processing of speech prosody provide the theoretical framework for the current study of pitch accent: the acoustic hypothesis postulating low-level cue-dependent processing on the basis of acoustic properties (e.g., Poeppel, 2003; Zatorre & Belin, 2001); and the functional hypothesis arguing for reliance on high-level linguistic functional load (e.g., Van Lancker, 1980). In addition to the two hypotheses, there is also an integrative view taking into account both lower-level acoustic and higher-level linguistic processing

Address for correspondence:

Xianghua Wu, Department of East Languages and Cultures, University of California, Berkeley, 3110 Dwinelle Hall, Berkeley, California 94720, U.S.A. xhwu@berkeley.edu

(Gandour, Wong, Dzemidzic, Lowe, Tong & Li, 2003a; Wong, 2002; Zatorre & Gandour, 2008).

Specifically, the acoustic account predicts that speech processing essentially involves lower-level mechanisms relying on distinctions in particular acoustic cues (Ivry & Robertson, 1998; Poeppel, 2003). Indeed, different hemispheric processing patterns have been revealed on the basis of physical acoustic differences such as temporal and spectral distinctions. For example, the processing of a voicing contrast (e.g., /ba-pa/) involving differences along a temporal span appears to be more right-hemisphere dominant than that of a rapid, localized spectral change (e.g., /ba-da/) (Ivry & Robertson, 1998). Furthermore, in the temporal domain, longer pitch patterns (~150-250 ms) are found to be more right-hemisphere lateralized than shorter pitch patterns (~20-40 ms) (Poeppel, 2003). Likewise, the processing of prosodic information across a longer temporal frame length (e.g., sentential intonation) engages a greater degree of right-hemisphere participation than more localized prosodic information (e.g., lexical tone superimposed on monosyllables) (Gandour, Wong, Lowe, Dzemidzic, Sathamnuwng, Tong & Li, 2002; Gandour, Tong, Wong, Talavage, Dzemidzic, Xu, Li & Lowe, 2004; Zhang, Shu, Zhou, Wang, & Li, 2010). Along the same lines, the processing of lexical prosody occurring over a syllable (e.g., lexical tone) involves greater right-hemisphere activities compared to that of a single segment (e.g., consonants) (Li, Gandour, Talavage, Hoffa, Lowe & Dzemidzic, 2010; Luo, Ni, Li, Li, Zhang, Zeng & Chen, 2006). In addition to temporal length, spectral information in prosody, particularly the shape of F0 contour, was also shown to affect hemispheric lateralization in the processing of pitch patterns. For instance, Wang, Wang and Chen (2013) showed that contour tones with changing F0 were more left lateralized than level tones.

Unlike the acoustically based cue-dependent hypothesis, the functional hypothesis predicts that the lateralization of speech prosody is determined by its linguistic functional load, which is measured by the extent of contrastivity between linguistic units as well as the frequency of occurrence for a given contrast (King, 1967; Surendran & Niyogi, 2006). It was suggested that processing of a prosodic entity with higher functional load is associated with greater extent of left-hemisphere dominance (Van Lancker, 1980). As such, lexical tone presumably has a higher functional load than pitch accent, because all words in a tone language are contrastive for tone, whereas only a limited number of words are contrastive for pitch accent, thus pitch accent patterns are limited in lexical selection (Pierrehumbert & Beckman, 1988; Shibata & Shibata, 1990; Tamaoka, Saito, Kiyama, Timmer, Verdonschot, Timmer & Verdonschot, 2014). Likewise, the function load of lexical stress, grammatical, and emotional intonation can be progressively low as the functional use

of these linguistic properties ranges from making lexical grammatical contrasts to serving paralinguistic functions (Cruttenden, 1997; Cutler, 1986; Hallé, Chang & Best, 2004). In line with the functional hypothesis, research has revealed increased left-hemisphere participation with the increase in functional load for different prosodic features. For instance, sentential-level emotional and grammatical intonation has been shown to employ a large extent of right-hemisphere processing (Gandour et al., 2003a; Gandour, Tong, Talavage, Wong, Dzemidzic, Xu, Li & Lowe, 2007; Shipley-Brown, Dingwall, Berlin, Yeni-Komshian & Gordon-Salant, 1988; Weintraub, Mesulam & Kramer, 1981), while lexical tone is predominantly processed in the left-hemisphere (Gandour, Wong & Hutchins, 1998; Gu, Zhang, Hu & Zhao, 2013; Van Lancker & Fromkin, 1973; Wang, Sereno & Jongman, 2001).

However, neither of the above two hypotheses alone can fully explain more complex observations of prosodic processing, particularly those involving 'intermediate-level' temporal frame length and/or functional load, such as pitch accent. In terms of acoustic temporal frame length, the disyllabic pitch accent lies intermediate to the more localized, monosyllabic lexical tones (presumably involving analytical processing in the left hemisphere) and the more global sentential-level intonation (presumably involving holistic processing in the right hemisphere). In terms of linguistic functional load, pitch accent, which is lexically contrastive in a limited number of words, is also intermediate to lexical tone (bearing higher functional use as a required lexical component and thus a left-hemisphere property) and intonation (assumed to have lower functional use and thus more right-hemisphere participation). It is therefore conceivable that patterns of pitch accent lateralization are less clear-cut.

Indeed, empirical research has revealed complex patterns of pitch accent processing. Both behavioral (dichotic listening, Kanamura & Imaizumi, 2008) and neurophysiological (near-infrared spectroscopic, Sato, Sogabe & Mazuka, 2007) studies showed a left-hemisphere preference for Japanese pitch accent processing in stimuli with high linguistic demands (e.g., in real word contexts which request differentiation of word meaning). However, this hemispheric advantage was absent when only the acoustic information (F0 contours) of the pitch accent patterns was presented and processed. Sato et al. (2007) further indicated that the extent of left-hemisphere involvement in pitch accent processing was less than that observed in previous lexical tone studies (e.g., Gandour, Wong, Hsieh, Weinzapfel, Van Lancker & Hutchins, 2000; Wang, Sereno, Jongman & Hirsch, 2003). Moreover, in sentential contexts, Japanese pitch accent perception appeared to engage bilateral processing, as shown in electro-physiological studies (Hayashi et al., 2001; Koso & Hagiwara, 2009).

These findings suggest that hemispheric processing patterns for pitch accent are not as straightforward as those found in lexical tone processing which consistently shows a left-hemisphere preference (e.g., Gandour et al., 2000; Wang et al., 2001). Patterns of pitch accent processing appear to vary as a function of specific linguistic and acoustic contexts, where a left-hemisphere preference is only evident in short, real word contexts but not in non-linguistic or sentential contexts. These patterns indicate integration of linguistic and acoustic influences. On the one hand, the processing of pitch accent patterns which appear in the same (disyllabic) temporal frame length may or may not be left dominant depending on whether they are processed as real words or as acoustic information (F0 contours) only, demonstrating linguistic influence (cf. Kanamura & Imaizumi, 2008; Sato et al., 2007). On the other hand, pitch accent processing in linguistic (real word) contexts may or may not be left dominant depending on the temporal frame length of the contexts, i.e., (shorter) disyllabic or (longer) sentential, respectively, indicating acoustic influence (cf. Kanamura & Imaizumi, 2008; Hayashi et al., 2001). Thus results from these previous studies could not tease apart how each of these influencing factors (linguistic or acoustic) contributes to pitch accent processing patterns.

1.2. Effects of linguistic experience

One way to disentangle the extent to which different linguistic and acoustic mechanisms are involved in pitch accent processing is to involve non-native listeners with different experience with pitch accent and lexical pitch in general, since presumably listeners with no pitch accent background would rely more on acoustic processes whereas those with experience would incorporate linguistic processes.

Previous findings indicate that hemispheric lateralization of prosody can be influenced by native and non-native prosodic experience, involving interactions of sensory acoustic and functional linguistic activities (Chandrasekaran, Krishnan & Gandour, 2007; Gandour et al., 2000; Gandour et al., 2003, 2007; Wang et al., 2003; Wang et al., 2004; Wong, Perrachione & Parrish, 2007a; Wong, Skoe, Russo, Dees & Kraus, 2007b; Xu, Gandour, Talavage, Wong, Dziedzic, Tong, Li & Lowe, 2006). Specifically, studies have shown that non-native relative to native prosodic processing typically involves greater extent of cortical activities, particularly in the right hemisphere (Gandour et al., 2003, 2004; Hsieh, Gandour, Wong & Hutchins, 2001; Klein et al., 2001; Wang et al., 2001, 2004). Moreover, patterns of non-native processing appear to be affected by linguistic functional load of particular prosodic contrasts. For instance, in the processing of higher-function prosody, such as lexical tone, non-native processing tends to be less left-lateralized

than native processing (Gandour et al., 2003a; Gandour et al., 2002; Hsieh et al., 2001; Van Lancker & Fromkin, 1973; Wang et al., 2001), while in the processing of lower-function prosody, such as sentential intonation, non-native processing and native processing exhibit comparable patterns (e.g., right-hemisphere dominance or bilateral processing, Gandour et al., 2003a, 2007). Furthermore, non-native listeners' backgrounds with L1 prosodic categories do not seem to influence their hemispheric processing of L2 prosodic categories. For instance, the hemispheric lateralization in native Mandarin listeners' perception of Thai tones or in native Norwegian listeners' perception of Mandarin tones does not reflect any positive influence from their L1 backgrounds with lexical tone (for Mandarin listeners) or pitch accent (for Norwegian listeners) (Gandour et al., 2000; Wang et al., 2004).

However, despite the observed native and non-native differences, hemispheric processing of non-native prosodic patterns can be modified due to increasing knowledge of L2 (Kaan, Wayland, Bao & Barkley, 2007; Wang et al., 2003, 2004; Wong & Perrachione, 2007; Wong et al., 2007a, 2007b; Zatorre & Gandour, 2008; Wu, 2013). Compared to naïve non-native listeners, L2 learners appear to approximate native processing patterns to a greater degree. Previous laboratory training studies have revealed that after non-native learners (e.g., English) received a short training period to perceive (e.g., Cantonese, Mandarin) lexical tones, their tone processing initially involved extended cortical activation as additional resources were recruited in learning challenging tonal contrasts. However, as learners gained L2 experience, their tone processing would engage a shift towards the native direction, from greater right-hemisphere involvement to left-hemisphere dominance (Wang et al., 2003; Wong et al., 2007a; Wong, Warrier, Penhune, Roy, Sadegh, Parrish & Zatorre, 2008). Consistent with these patterns, further research revealed that training-induced modifications of hemispheric processing could also be associated with listeners' ability to acquire new language contrasts. For example, while successful tone learners exhibit increased activation in the left-hemisphere regions, learners with limited improvement show increased activation in the right auditory cortex responsible for non-linguistic pitch processing (Wong et al., 2007a).

These results consistently indicate that effective L2 tone learning involves a right-to-left shift in hemispheric dominance. However, most research has focused on lexical tone with high functional load, for which the native direction is equivalent to left-hemisphere dominance. To further investigate the integration of linguistic and acoustic aspects in prosodic processing, it would be necessary to examine the effects of L2 learning on the processing of prosodic features with lower functional load, such as pitch accent, which has not been delved into

previously. Moreover, as reviewed above, research has shown that L1 prosodic experience with Mandarin tone or Norwegian accent did not affect L2 prosodic processing of Thai or Mandarin tones, respectively (Gandour et al., 2000; Wang et al., 2004). The absence of L1 effects could be due to the linguistic functional differences between L1 and L2 prosody, where lower-function (Norwegian accent) or less complex (Mandarin tone) L1 prosodic experience may not easily affect higher-function (Mandarin tone) or more complex (Thai tone) L2 prosodic processing. Further research is needed to test if L1 experience with higher-function prosody (e.g., lexical tone) could facilitate processing of lower-function prosody (e.g., pitch accent) in an L2.

1.3. Japanese pitch accent

In Japanese, combinations of high (H) or low (L) pitch (F0) as well as an accent (*) signaling prominence form three pitch accent patterns (H*L, LH*, LH) to contrast word meaning for disyllabic words (Cruttenden, 1997; Kitahara, 2001; Sugiyama, 2006). For instance, one of the triplet stimulus sets included in the current research, *kaki*-H*L, *kaki*-LH* and *kaki*-LH, which contains the same disyllabic segment components but different pitch accent patterns, results in three words: *oyster*, *fence* and *persimmon*, respectively. The F0 contours of the three pitch accent patterns exemplified by *kaki* are shown in Figure 1.

The acoustic features characterizing pitch accent include F0, amplitude and duration, where accented syllables typically carry higher F0 and amplitude, and longer duration than unaccented syllables (Beckman, 1986; Hasegawa & Hata, 1992; Sugito, 1972): specifically, the acoustic realizations of pitch accent patterns which are salient in perception include F0 trajectories, particularly F0 direction and position of F0 peak, whereas intensity and duration are found to be secondary cues (Beckman, 1986; Hasegawa & Hata, 1992; Maniwa, 2002; Sugiyama, 2006, 2008, 2014). In terms of F0, the H*L pattern typically involves a falling F0 contour with peak F0 falling on the first syllable in a disyllabic word, while the LH* and LH patterns involve rising F0 contours with peak F0 falling on the second syllable of a disyllabic word. Furthermore, while F0 direction (pitch fall or pitch rise) primarily distinguishes the H*L and LH*/LH patterns, pitch fall has been claimed to be a more primary cue than pitch rise for recognizing accent patterns (Kindaichi, 1967; Hirano-Cook, 2011). Research has also examined the LH* and LH patterns which differ in accentedness on the second syllable (Hasegawa & Hata, 1992; Maniwa, 2002; Sugiyama, 2006, 2008, 2014; Vanditti, 2005). Overall, the F0 maximum value for LH* is higher than that for LH, and the F0 difference between the first syllable and second syllable is larger for LH* and LH when produced in a

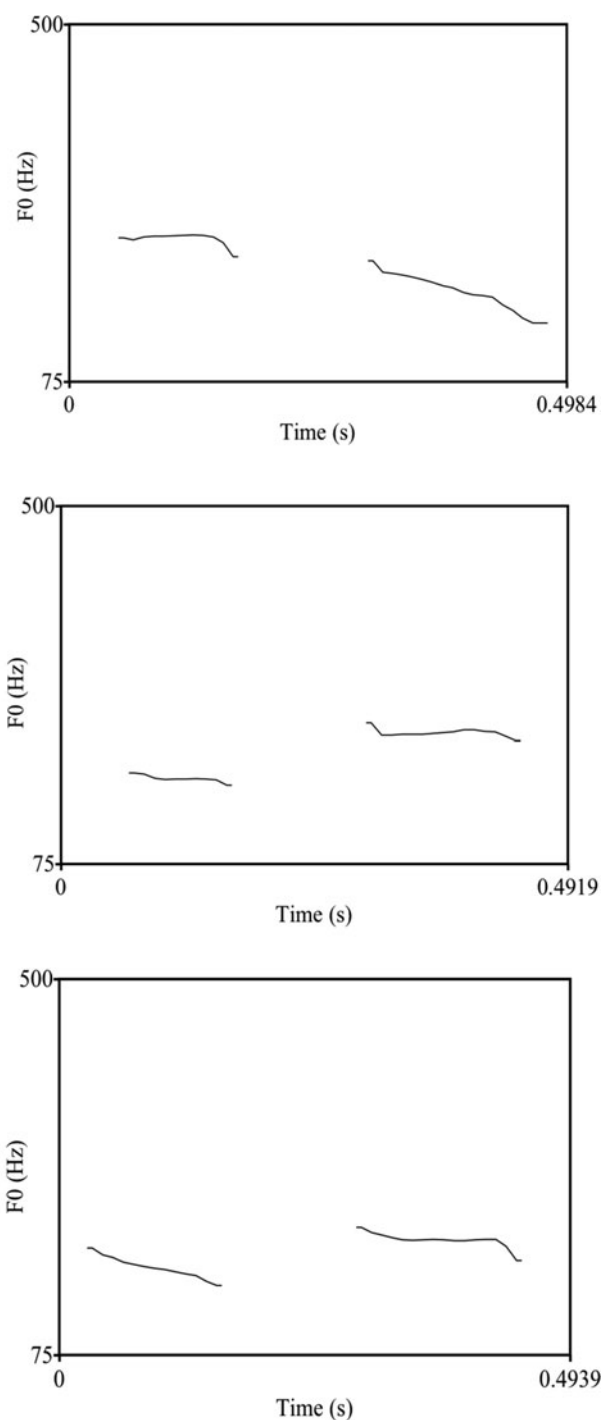


Figure 1. Fundamental frequency (F0) contours of the three pitch accent patterns: high-accent-low (H*L), low-high-accent (LH*), and LH (low-high), from top to bottom, exemplified by the disyllable *kaki*. Each of these example disyllabic words was excised from a phrasal context, a particle *が* (-ga) that follows it.

sentential context, while these differences are neutralized when produced in isolation (Sugito, 1983; Sugiyama, 2006).

1.4. The current study

As reviewed above, previous findings indicate effects of linguistic and acoustic influences on pitch accent processing. However, research has not identified how each of these two factors contributes to the observed processing patterns. Examining non-native listeners with different linguistic experience with lexical pitch allows disentangling the different mechanisms underlying pitch accent processing, as listeners with less pitch accent experience would presumably rely more on acoustic processes whereas those with more experience would incorporate linguistic processes to a greater extent.

The current study explores the interactive effects of acoustic and linguistic experience on the hemispheric processing of Japanese pitch accent patterns. This is a follow-up study based on our previous research (Wu, Tu & Wang, 2012). Using dichotic listening, Wu et al. (2012) examined hemispheric processing of Japanese pitch accent by native Japanese listeners as well as native English and Mandarin Chinese listeners with no Japanese or pitch accent background. The results revealed that Japanese pitch accent processing is less lateralized compared to the left-hemisphere dominant lexical tone processing found in previous research. Detailed analyses of individual pitch accents across groups indicated a reliance on acoustic cues, showing a right hemisphere preference for processing the H*L pattern, a left hemisphere preference for LH*, and no hemisphere dominance for LH. However, since Wu et al. (2012) tested naïve non-native listeners without Japanese background, one would expect them to process Japanese accent with a greater degree of acoustic-level processing than linguistic-level processing. One subsequent question is the extent to which non-native listeners will encode linguistic information as they gain experience with pitch accent.

The current study thus follows up on Wu et al. (2012) to explore the influence of linguistic experience, with the focus on the effects of L2 experience and its relation to L1 backgrounds, on the processing of Japanese pitch accent patterns by non-native learners of Japanese whose L1 was either tonal (Mandarin Chinese) or non-tonal (English). The recruitment of these two groups is unique in terms of learners' L1 backgrounds relative to their L2 experience, as pitch accent (the target L2 lexical prosodic property) has a lower linguistic functional load than lexical tone (the primary L1 lexical prosodic experience for Chinese learners) while a higher functional load than stress or focus (the English learners' L1 prosodic experience). Including the three groups of listeners from Wu et al. (2012), the five language groups in this study are: native Mandarin Chinese learners of Japanese (CL), native English learners of Japanese (EL), native Mandarin Chinese listeners with no Japanese learning experience (i.e., naïve Chinese

listeners or 'Chinese non-learners', CNL), native English listeners with no Japanese learning experience (i.e., naïve English listeners or 'English non-learners', ENL), and native Japanese listeners (NJ). These five groups represent listeners with a gradation of lexical pitch experience in pitch accent processing: from native experience with pitch accent (NJ), to L2 pitch accent experience with and without L1 lexical tone experience (CL and EL, respectively), and to lack of pitch accent experience but with and without L1 tone experience (CNL and ENL, respectively). Thus, the inclusion of these listeners allows us to investigate the effects of L2 learning in relation to L1, particularly to examine potential interactions of L1 experience (non-tonal versus tonal L1) and L2 experience (naïve listener versus learner) in pitch accent processing. In a broad theoretical context, this research will contribute to the understanding of the extent to which pitch accent processing involves lower-domain acoustic mechanisms and the extent to which it is influenced by higher-domain linguistic experience (cf. Zatorre & Gandour, 2008).

In terms of the effects of L2 experience, Chinese and English learners are expected to approximate the native Japanese patterns to a greater extent and/or engage a greater extent of linguistic-level processing, as compared to naïve Chinese and English listeners. Moreover, we predict interactive effects of L1 and L2 experience. First, the Chinese learners' pitch accent processing patterns would more likely reflect the influence of their L1 experience with lexical tone (bearing higher functional load than pitch accent), whereas the English learners' L1 experience with post-lexical prosody (bearing lower functional load than pitch accent) would influence their pitch accent perception to a lesser degree. In terms of lateralization patterns, a higher degree of left-hemisphere processes may be observed for Chinese listeners compared to English learners due to the Chinese learners' L1 tonal experience. Furthermore, a comparison between English learner and naïve Chinese listener patterns would disentangle the relative contribution of L2 learning experience and L1 experience with lexical prosody. Finally, any patterns that are common across all groups would indicate acoustic-level processing. Overall, listeners with (L2) pitch accent or (L1) lexical pitch background would presumably incorporate greater linguistic processes, whereas those with less experience with either L1 or L2 lexical prosody would rely more on acoustic processes.

2. Methods

2.1. Participants

Fourteen native Mandarin Chinese learners of Japanese (CL, 9 males) and 14 native English learners of Japanese (EL, 9 males) participated in the dichotic listening study.

Table 1. Participant group characteristics and language background information

Group	Native Japanese (NJ)	Chinese learner (CL)	English learner (EL)	Chinese non-learner (CNL)	English non-learner (ENL)
Number of participants recruited*	17	16	16	22	22
Number of participants included in dichotic listening test	16	14	14	16	16
Included in Wu et al. (2012)	yes	no	no	yes	yes
Japanese experience	native control	3-12 months as L2	3-12 months as L2	none	none
L1 background	pitch accent	tonal	non-tonal	tonal	non-tonal

* Among all the recruited participants, a few in each group did not reach the criterion for inclusion in the dichotic listening test (see Section 2.3 for detailed information).

Additionally, 16 native Mandarin Chinese listeners with no Japanese learning experience (CNL, 5 males), 16 native English listeners with no Japanese learning experience (ENL, 7 males), and 16 native Japanese listeners (NJ, 6 males) from Wu et al. (2012) were included. All five groups of participants were recruited from the undergraduate and graduate student population at Simon Fraser University, and were matched in age (mean: 23 years). The Chinese and English learners had three months to one year of experience learning Japanese as an L2 at a Canadian university, and had no other pitch accent language experience apart from Japanese. As reported in Wu et al. (2012), the Chinese and English naïve listeners had no pitch accent language experience. No Japanese or English listeners had any tone-language experience. All the participants reported normal hearing and speech ability, and all were right-handed based on the Edinburgh Handedness Inventory (Oldfield, 1971) that they were required to complete prior to testing. None of the listeners had any formal musical training (cf. Wong et al., 2007b). The listeners were paid for their participation in this research. Table 1 displays the characteristics of the five groups of participants included in the current study.

2.2. Stimuli

The stimuli in the current study were the same as those used in Wu et al. (2012), that is, twenty-one Japanese disyllabic words, consisting of seven minimal triplets: 3 pitch accent patterns (H*L, LH*, LH) x 7 syllables (*aki, hana, kaki, nami, take, tama, yuki*). All the words are commonly used in Japanese. Eighteen of the 21 words were adapted from Sugiyama (2006), selected due to their relatively high familiarity ratings based on a computerized dictionary (Amano & Kondo, 2000). Similarly, the remaining three words were rated as high frequency by an online Japanese dictionary, *Denshi Jisho* (<http://www.jisho.org/>).

Each word was recorded four times by a female linguistically-trained native speaker of Tokyo Japanese

(aged 32) 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 followed by a monosyllabic particle, including *が* (-ga), *を* (-wo), *に* (-ni), *と* (-to), and *の* (-no), 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 x 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, e.g., [*hana-H*L, hana-LH**], [*hana-H*L, hana-LH*], or [*hana-LH*, hana-LH*]. 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 to the same RMS intensity (70 dB) using Sound Forge 6.0 (Sonic Foundry, Inc.). 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 444–581 ms (533 ms on average).

2.3. Procedure

The procedure and setting of the current study were the same as those used in Wu et al. (2012). The experiment was created using E-prime 1.0 (Psychology Software Tools, Inc.) and conducted in a sound-attenuated booth in the Language and Brain Lab at Simon Fraser University.

Prior to the main dichotic listening test, three preliminary tasks were administered: pitch accent familiarization, binaural pitch accent identification, and word recognition (which was additionally designed for the learner groups in the current study). These tasks were to familiarize all participants with the three pitch accent patterns and ensure that they could distinguish

these patterns when presented binaurally. Only those participants whose binaural pitch accent identification accuracy for the twenty-one target words was higher than 60% (well above the chance level, 33%) could proceed with the dichotic listening test. By this criterion, the numbers of participants excluded from the total recruited in each group were: one out of 17 for NJ, six out of 22 for CNL, six out of 22 for ENL, two out of 16 for CL, and two out of 16 for EL (cf. Table 1). As mentioned above, the word recognition task was designed to familiarize the two learner groups with the meanings of the target words to ensure that all learners were equally familiar (or unfamiliar) with all the target words such that they would not be biased by (un)familiarity with the meanings of the words when performing the dichotic listening task. Furthermore, in order to determine whether the variance in the length of Japanese learning experience (three months to one year) among learners affected the homogeneity within each learner group in terms of their pitch accent perception ability, correlation analyses were performed between length of learning and correct binaural identification of pitch accent for each of the learner groups. No significant correlation was found for the Chinese learners [$r(14) = -.18, p = .53$] or the English learners [$r(14) = -.52, p = .06$], indicating that the learners within each learner group were comparable in binaural pitch accent identification accuracy.

The dichotic listening test procedures were modeled after similar previous studies (e.g., Wang et al., 2001, 2004). A total of 168 trials were presented, including four repetitions of the 42 dichotic pairs in four blocks. Each pair was auditorily presented to the participants with one word in the left ear and the other in the right ear simultaneously. The task was forced-choice identification, for which participants were asked to identify both stimuli by indicating what they heard in their left ear and what they heard in their right ear. They provided their responses by clicking the pitch accent labels (H*L, LH*, LH) prompt on a computer screen. To avoid response order bias within participants, right- and left-ear responses were counterbalanced after two blocks. Additionally, two versions of the test were created 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 (LR) and then reverse the order (RL), while for the other half of the participants, the order was RL followed by LR. Furthermore, to eliminate channel effects, the participants were requested to reverse the headphone channels across blocks. The dichotic test for each participant lasted approximately 30 minutes.

2.4. Perception data analysis

Perceptual accuracy for the left and the right ear was calculated, following previously established analysis

procedures (e.g., Kimura, 1961; Shipley-Brown et al., 1988; Wang et al., 2001, 2004). Hemispheric asymmetry patterns were determined based on the difference in correct response between the right ear and the left ear (% right-ear correct minus % left-ear correct). A positive value indicates right-ear advantage (REA) for left-hemisphere dominance, a negative value indicates left-ear advantage (LEA) for right-hemisphere dominance, and “0” indicates no ear advantage for balanced bilateral processing. Based on this criterion, the numbers of participants in each group showing left- versus right-hemisphere dominance were compared to determine the distribution of hemispheric lateralization patterns.

2.5. Acoustic analysis

The acoustic features of the different pitch accent patterns of the disyllabic triplets used in the current dichotic test were analyzed using Praat (Boersma & Weenink, 2013). The acoustic features included here have been shown to be relevant perceptual cues, as reviewed above (e.g., Hasegawa & Hata, 1992; Maniwa, 2002; Sugiyama, 2006, 2008, 2014). First, to track F0 trajectories and determine contour direction for different pitch accent patterns (primarily between H*L and LH*/LH), F0 values at 0%, 25%, 50%, 75% and 100% of the first and second syllables of the disyllabic words, as well as the average F0 values for each syllable, were measured. These F0 values included all the voiced segments in a syllable. Then, to compare LH* and LH, “F0 rise” values were obtained by subtracting the F0 minimum of the first syllable from the F0 maximum of the second syllable (Sugiyama, 2006, 2008). In addition, secondary features including vowel mean intensity and syllable duration were measured for the first and second syllables.

3. Results

As stated previously, in order to examine the effects of L2 learning experience in relation to L1 backgrounds, the five listener groups were directly compared in a single set of analyses, combining the data acquired from the current two learner groups (CL, EL) and the Wu et al. (2012) data from the two non-learner groups (CNL, ENL) as well as the native Japanese group (NJ).

3.1. Perceptual accuracy

Overall results

Percent correct identification was analyzed through three-factor repeated measures analyses of variance (ANOVAs) using Ear (left, right) and Pitch accent pattern (H*L, LH*, LH) as the within-subjects factors, and Group (NJ, CL, EL, CNL, ENL) as the between-subjects factor. The analyses showed a significant main effect of Ear [$F(1, 71) = 7.1$,

Table 2. Percent correct identification (%) of the three Japanese pitch accent patterns (H*L, LH* and LH) in the left and right ear by native Japanese, Chinese learner, English learner, naïve Chinese, and naïve English listener groups (NJ, CL, EL, CNL and ENL, respectively). Standard deviation (SD) values are provided in parentheses.

Group	% Correct identification (SD)					
	H*L		LH*		LH	
	Left ear	Right ear	Left ear	Right ear	Left ear	Right ear
NJ	74 (24)	69 (25)	47 (19)	51 (17)	57 (17)	56 (20)
CL	74 (20)	69 (25)	54 (17)	48 (17)	44 (19)	51 (20)
EL	70 (20)	54 (19)	58 (14)	57 (22)	40 (20)	46 (16)
CNL	71 (14)	61 (24)	46 (15)	49 (18)	42 (19)	44 (18)
ENL	69 (16)	52 (19)	50 (16)	55 (13)	40 (12)	44 (14)

$p = .01, \eta_p^2 = .09$], with the percent correct identification of the stimuli presented in the left ear (55.4%) being higher than that in the right ear (53.7%), indicating an overall left-ear advantage. A significant main effect for Pitch accent pattern was also observed [$F(2, 142) = 40.6, p < .0001, \eta_p^2 = .364$]. Bonferroni-adjusted post hoc analyses further revealed that the H*L pattern (65.7%) was identified more accurately than the LH* pattern (51.5%, $p < .0001$) and the LH pattern (46.4%, $p < .0001$), but there was no difference between the LH* and LH patterns ($p = .16$). No main effect of Group was observed [$F(4, 72) = .9, p = .5, \eta_p^2 = .05$].

The analyses also revealed significant interactions between Ear and Pitch accent pattern [$F(2, 142) = 14.9, p < .0001, \eta_p^2 = .174$], and between Group and Pitch accent pattern [$F(8, 142) = 2.2, p = .03, \eta_p^2 = .112$], but not between Ear and Group [$F(4, 71) = .2, p = .93, \eta_p^2 = .011$], or Ear, Pitch accent, and Group [$F(8, 142) = 1.68, p = .107, \eta_p^2 = .09$]. Further analyses were performed on the basis of these significant interactions.

Table 2 displays the average correct identification of the three pitch accent patterns in the left and right ears by listeners from the five groups.

Ear and pitch accent pattern

Based on the significant Ear and Pitch accent pattern interaction, further one-factor repeated measures ANOVA analyses were conducted for the ear effect on the perception of individual pitch accent patterns. As shown in Figure 2, the results revealed better performance in the left ear (70%) than the right ear (61%) for the H*L pattern [$F(1, 75) = 28.3, p < .0001, \eta_p^2 = .274$], consistent with the overall results of a left-ear advantage. However, for LH, the right ear (48%) excelled the left ear (45%) [$F(1, 75) = 4.8, p = .03, \eta_p^2 = .06$], and for LH*, no significant difference was found between the left (51%) and right (52%) ears [$F(1, 75) = .09, p = .35, \eta_p^2 = .01$]. Moreover, one-factor repeated measures ANOVAs for each ear showed significant effects of pitch accent pattern

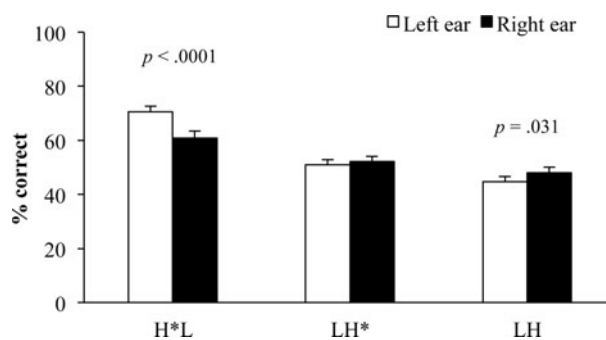


Figure 2. Mean percent correct identification in the left and right ear for the H*L, LH*, and LH patterns across groups. The p values are provided for significant differences between the left and right ears. Error bars indicate standard errors.

for both the left ear [$F(2, 150) = 60.4, p < .0001, \eta_p^2 = .442$] and the right ear [$F(2, 150) = 11.3, p < .0001, \eta_p^2 = .13$]. Bonferroni-adjusted post hoc tests further showed that, consistent with the across-ear results, the H*L pattern was significantly more accurately perceived than both the LH* and LH patterns for both the left ear and the right ear ($ps < .001$), whereas no difference between the LH* and LH patterns was found ($ps > .5$).

Group and pitch accent pattern

Based on the significant Group x Pitch accent pattern interaction, further one-factor repeated ANOVAs were performed for separate groups, resulting in a significant effect of pitch accent pattern for each of the five groups, as shown in Figure 3: NJ [$F(2, 30) = 11.2, p < .0001, \eta_p^2 = .43$]; CL [$F(2, 26) = 10.5, p < .0001, \eta_p^2 = .43$]; EL [$F(2, 26) = 6.3, p = .006, \eta_p^2 = .325$]; CNL [$F(2, 30) = 9.8, p = .001, \eta_p^2 = .397$]; and ENL [$F(2, 30) = 11.2, p < .0001, \eta_p^2 = .428$]. However, Bonferroni-adjusted post hoc analyses further revealed different pitch accent perception patterns for different groups. While for the NJ, CL, and CNL listeners, perception was more accurate for H*L than LH*

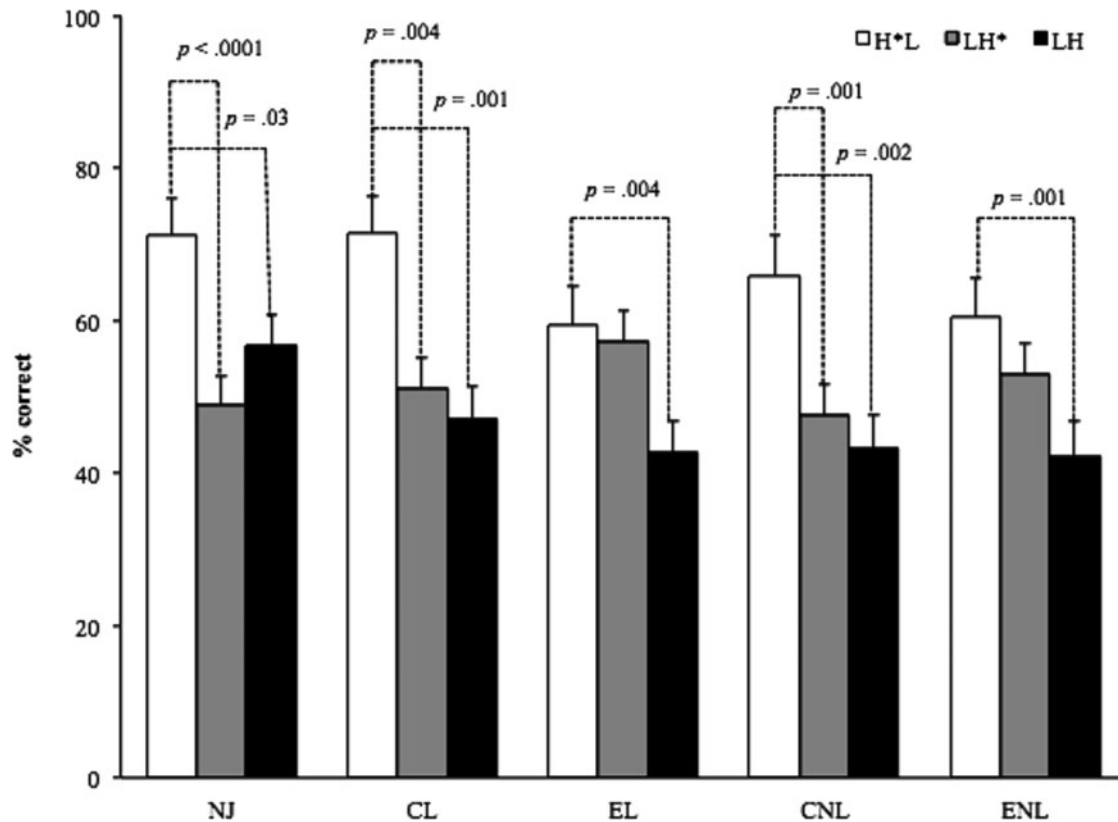


Figure 3. Mean percent correct identification for the H*L, LH*, and LH patterns across ears by listeners of the five groups (NJ: Native Japanese; CL: Chinese learners of Japanese; EL: English learners of Japanese; CNL: Chinese naïve listeners of Japanese; ENL: English naïve listeners of Japanese). The p values are provided for significant differences between pitch accent patterns. Error bars indicate standard error.

($ps < .004$) and LH ($ps < .03$), the EL and ENL groups perceived H*L more accurately than LH only ($ps < .004$) but not LH* ($ps = .09$). No difference was found between LH* and LH for any of the groups. Moreover, one-factor ANOVAs examining group effect on the identification of each pitch accent pattern showed no significant group difference: H*L [$F(4, 71) = 1.3, p = .28$], LH* [$F(4, 71) = .9, p = .48$], and LH [$F(4, 71) = 2.13, p = .09$].

3.2. Distribution of ear preference

Pearson's chi-square (χ^2) analyses were performed to examine the distribution of ear dominance patterns in term of the number of listeners showing left-ear advantage (LEA) versus right-ear advantage (REA)¹ for each group and pitch accent pattern, as displayed in Table 3.

The results revealed significant differences between the number of listeners showing LEA or REA for different groups and pitch accent patterns. For H*L, significantly more listeners in the EL [$\chi^2(1) = 7.14, p = .008$] and ENL

Table 3. The number of listeners for each group (NJ, CL, EL, CNL, ENL) showing left-ear advantage (LEA) and right-ear advantage (REA) in the processing of H*L, LH* and LH patterns. Bold numbers in shaded cells indicate statistically significant differences between the distribution of LEA and REA ($p < .05$).

	H*L LEA-REA	LH* LEA-REA	LH LEA-REA
NJ	9-7	6-10	7-7
CL	9-5	9-5	3-10
EL	12-2	8-6	6-7
CNL	9-6	5-10	8-8
ENL	15-1	4-12	7-8

[$\chi^2(1) = 12.25, p < .0001$] groups showed LEA than those showing REA. For LH*, a larger number of listeners in the ENL group showed REA than that showing LEA [$\chi^2(1) = 4, p = .046$]. No differences were observed for the other groups and pitch accent patterns.

¹ Listeners showing no ear advantage were excluded in order to compare differences in ear advantage patterns.

In addition, in order to examine if the Chinese and English learners' lateralization patterns (REA or LEA) were affected by the length of their Japanese learning experience (which varied four months to a year across learners), correlation analyses were performed between learners' lateralization pattern (difference between right and left ear correct identification) and length of learning. No significant correlation was found either for the Chinese learners [$r(14) = .29, p = .32$] or for the English learners [$r(14) = .06, p = .85$].

3.3 Acoustic results

To compare the acoustic features of the three pitch accent patterns, two-factor repeated measures ANOVAs were performed, with pitch accent pattern (H*L, LH* and LH) and syllable position (1st syllable and 2nd syllable) as independent variables. The dependent variables include average F0, mean intensity, and duration of the first or second syllable, respectively. Data were analyzed across the seven target triplets.

For average F0, a significant main effect was observed for pitch accent pattern [$F(2, 12) = 15, p = .001, \eta_p^2 = .71$] but not for syllable position [$F(1, 6) = 2.2, p = .19, \eta_p^2 = .27$]. However, there was a significant interaction of pitch accent pattern and syllable position [$F(2, 12) = 156, p < .001, \eta_p^2 = .96$]. To further explore this interaction, one-factor repeated ANOVAs were performed to test the effects of pitch accent pattern for each syllable, where significant main effects were observed for both the 1st syllable [$F(2, 12) = 96.6, p < .0010, \eta_p^2 = .94$] and the 2nd syllable [$F(2, 12) = 12.4, p = .001, \eta_p^2 = .67$]. Bonferroni-adjusted post hoc analyses indicated that, as expected, the average F0 of the 1st syllable in H*L (245 Hz) is higher than that in LH* (179 Hz) and LH (170 Hz) ($p < .0001$), whereas that of the latter two did not differ ($p = .13$). The average F0 for the 2nd syllable in H*L (189 Hz) is lower than that in LH* (223 Hz) ($p = .03$) but not in LH (201 Hz) ($p = .09$), and the average F0 of LH is lower than that of LH* as well ($p = .03$). Additionally, one-factor ANOVAs on syllable position for each pitch accent pattern revealed the expected differences: for H*L, higher F0 in the 1st than in the 2nd syllable [$F(1, 6) = 456, p < .0001, \eta_p^2 = .99$]; and for both LH* [$F(1, 6) = 29, p = .02, \eta_p^2 = .83$] and LH [$F(1, 6) = 51.7, p < .0001, \eta_p^2 = .9$], lower F0 in the 1st than the 2nd syllable.

For intensity, the 2-factor ANOVAs showed no significant main effects of either pitch accent pattern [$F(2, 12) = 3.8, p = .054, \eta_p^2 = .39$] or syllable position [$F(1, 6) = 3.2, p = .13, \eta_p^2 = .35$], but a significant interaction of the two was observed [$F(2, 12) = 18.7, p < .0001, \eta_p^2 = .76$]. Subsequent one-factor repeated measures ANOVAs were conducted on pitch accent pattern for each syllable. The analysis revealed significant effects for the 1st syllable [$F(2, 12) = 17.6, p < .0001, \eta_p^2 = .75$], with Bonferroni

adjusted post hoc analyses showing higher intensity for H*L (78 dB) than LH* (74 dB) ($p = .01$) and LH (74 dB) ($p = .02$), but no difference between the latter two ($p = 1$). No significant effect was found for the 2nd syllable (H*L: 77 dB, LH*: 76 dB, LH: 77 dB, [$F(2, 12) = .45, p = .64, \eta_p^2 = .07$]). Additionally, one-factor repeated measures ANOVAs on syllable position for each pitch accent pattern showed that, for H*L, the 1st syllable had a higher intensity than the 2nd syllable [$F(1, 6) = 8.6, p = .03, \eta_p^2 = .59$]; but for both LH* [$F(1, 6) = 9.4, p = .02, \eta_p^2 = .61$] and LH [$F(1, 6) = 8.3, p = .02, \eta_p^2 = .58$], the 1st syllable had a lower intensity than the 2nd syllable.

For duration, no significant difference was revealed from the 2-factor ANOVAs for pitch accent pattern [$F(2, 12) = 2.6, p = .12, \eta_p^2 = .3$] or the interaction of pitch accent pattern and syllable position [$F(2, 12) = 3.2, p = .08, \eta_p^2 = .35$]. However, a significant main effect was found for syllable position [$F(1, 6) = 50, p < .0001, \eta_p^2 = .89$], with the 1st syllable (164ms) being shorter than the 2nd syllable (253ms) across pitch accent patterns.

Moreover, LH* and LH were compared for F0 rise, maximum F0 and minimum F0 using paired-samples *t*-tests. The results indicate greater F0 rise for LH* (64 Hz) than LH (53 Hz) [$t(6) = 3.4, p = .015$] and higher maximum F0 for LH* (232 Hz) than LH (213 Hz) [$t(6) = 4.1, p = .006$], but no difference in minimum F0 was found between LH* (167 Hz) and LH (160 Hz) [$t(6) = 1.5, p = .18$].

The acoustic measurements are summarized in Table 4. Additionally, Appendix A displays the F0 values at 0%, 25%, 50%, 75% and 100% of the first and second syllables to indicate F0 trajectories of the three pitch accent patterns.

Overall, the acoustic features of the current stimuli are consistent with those claimed previously, characterizing the H*L pattern as having a falling F0 contour and LH*/LH having rising F0 contour shapes (Hirano-Cook, 2011; Kindaichi, 1967; Kitahara, 2001; Sugiyama, 2006). Moreover, differences were found between LH* and LH, where F0 rise and maximum F0 are higher for LH* than LH. Given that the current stimuli were recorded in phrasal contexts, these results agree with the previous findings of such differences when LH* and LH words were produced in sentential contexts (Sugito, 1983; Sugiyama, 2006).

4. Discussion

The results revealed an overall left-ear advantage across the five groups of listeners, suggesting a right-hemisphere preference in the processing of Japanese pitch accent. Detailed analyses with individual pitch accent patterns showed different hemispheric processing patterns: right-hemisphere preference for H*L, left-hemisphere preference for LH, and balanced bilateral

Table 4. Measurements of relevant acoustic features of the three pitch accent patterns, including average F0, intensity and duration for the first and second syllables. For LH* and LH, measurements include minimum F0 of the first syllable and maximum F0 of the second syllable, as well as F0 rise (difference between maximum F0 and minimum F0). SDs are shown in parentheses.

	H*L		LH*		LH	
	1 st syllable	2 nd syllable	1 st syllable	2 nd syllable	1 st syllable	2 nd syllable
Average F0	245	189	179	223	170	201
[Hz]	(16)	(16)	(8)	(21)	(4)	(10)
F0 minimum	-	-	167	-	160	-
[Hz]			(10)		(11)	
F0 maximum	-	-	-	231	-	213
[Hz]				(8)		(9)
F0 rise	-	-	64	-	53	-
[Hz]			(15)		(12)	
Vowel intensity	78	77	74	76	74	77
[dB]	(3)	(3)	(3)	(5)	(3)	(5)
Syllable duration	166	239	168	259	157	261
[ms]	(31)	(26)	(25)	(45)	(20)	(49)

processing for LH*. In terms of group difference, while the Chinese listeners (CNL, CL) followed the native Japanese (NJ) patterns with more accurate perception of H*L than with both LH* and LH, the English listeners' (ENL, EL) perception of H*L was only better than LH, but not LH*. Furthermore, for H*L, more EL and ENL listeners showed LEA than those showing REA; but for LH*, a larger number of ENL listeners showed REA than that showing LEA.

4.1. Cue-dependent processing

The overall results show a right-hemisphere preference in Japanese pitch accent processing, not only for the naïve listeners (CNL, ENL) with no experience with pitch accent, but also for the learners (CL, EL) as well as native Japanese listeners (NJ) to whom pitch accent is a linguistically meaningful entity. The lack of a left-hemisphere dominance exhibited by the Japanese natives in the current study (as well as in Wu et al., 2012) is consistent with the previous claim that pitch accent, due to its lower functional load, engages a lesser degree of linguistic processing as compared to lexical tone (Sato et al., 2007; Van Lancker, 1980). For the learner groups, experience with learning pitch accent as a linguistic entity did not lead to any processing advantage in the putative language-relevant left hemisphere. Particularly, the Chinese learners' L1 experience with (higher-function) lexical tone processing in the left hemisphere did not appear to be recruited in (lower-function) pitch accent processing in an L2. These patterns are in line with our prediction that pitch accent processing may largely rely on lower-level acoustic information since the overall

processing patterns were similar for native and non-native listeners alike and were not influenced by either L1 tonal experience or L2 learning experience.

These results support the cue-dependent account that the lateralization of prosodic information may be determined by acoustic properties such as its temporal frame length, where the right hemisphere is more dominant for processing global prosodic information while the left hemisphere is more dominant for localized information (Gandour et al., 2002, 2004; Poeppel, 2003; Zhang et al., 2010). Compared to lexical tone which is typically superimposed on a single syllable, pitch accent minimally spans over two syllables, which perception cannot be accurately determined until the second syllable (Walsh, 1996) and requires a phrasal context (Maniwa, 2002). It is thus conceivable that the processing of such less localized information involves less left-hemisphere participation (as compared to lexical tone) for native and non-native listeners alike. Additionally, previous studies revealed that pitch accent processing is dominant in the left hemisphere only in high linguistically demanding contexts (Kanamura & Imaizumi, 2008; Sato et al., 2007). However, the current dichotic listening task required listeners to indicate pitch accent patterns rather than identifying the meanings of the presented words. This task did not necessarily involve retrieval of higher-level (lexical semantic) linguistic information since the acoustic input alone was sufficient to complete the task. This may account for the right-ward preference exhibited by not only the naïve listeners but also learners as well as native listeners.

The results of individual pitch accent patterns reveal that the overall right-hemisphere preference only applies

to the H*L pattern, whereas the LH and LH* patterns exhibit a left-ward or balanced bilateral processing, respectively, across participant groups. Upon closer inspection, it appears that the right-hemisphere advantage for the H*L pattern consistently occurs for each of the five participant groups both in terms of the accuracy scores (cf. Table 2) and of the number of participants showing this preference pattern (cf. Table 3). In contrast, the individual-group data (Tables 2 and 3) indicate that the across-group patterns for LH and LH* do not consistently apply to all the groups. For example, while naïve Chinese and English listeners as well as Japanese natives show a left-hemisphere advantage for LH*, Chinese and English learners exhibit a right-ward preference. This inconsistency may account for the lack of dominance pattern for LH* in the overall results. Moreover, it may also explain why the current across-group result of a bilateral pattern for LH* does not agree with the Wu et al. (2012) pattern of a left-ward preference, since Wu et al. (2012) only included the naïve listener and Japanese groups. Similar cross-group inconsistencies are also observed for the LH pattern in the current study as well as Wu et al. (2012).

Thus, these individual pitch accent pattern results indicate that listeners across groups may be able to more reliably pick up the acoustic cues to H*L using right-hemisphere mechanisms, since H*L is more acoustically salient for perception as compared to LH* and LH. Indeed, previous research on pitch accent has indicated that the primary cues for pitch accent pattern identification is present in the pitch fall rather than the pitch rise (Hirano-Cook, 2011; Kindaichi, 1967). Consistent with these findings, the current acoustic measurements show a very clear F0 falling contour for H*L which significantly distinguishes it from both LH* and LH with rising F0 contours, whereas the minor F0-rise differences (9 Hz, cf. Table 4) between LH* and LH may not be easily distinguishable (Sugiyama, 2006; Maniwa, 2012). The acoustic distinctiveness of H*L is also reflected in perception, where the overall accuracy for H*L is higher than that for LH* and LH (cf. Figure 2). These results indicate acoustic-level processing of the pitch accent patterns that are acoustically salient, as similar processing mechanisms are involved across listener groups regardless of L1 or L2 backgrounds.

4.2. Role of linguistic experience

The lack of a left-hemisphere dominance for pitch accent may also be interpreted in terms of linguistic functional load. As reviewed previously, unlike lexical tone, pitch accent only applies to a small number of words to signal lexical contrasts (Pierrehumbert & Beckman, 1988; Shibata & Shibata, 1990; Tamaoka et al., 2014), and thus bears lower functional load as compared to tone. The current results support the linguistic functional

hypothesis, which predicts different degrees of left-hemisphere participation in the processing of prosodic entities as a function of their functional use, ranging from a left-hemisphere dominance for lexical tone to right-hemisphere dominance for emotional prosody (Gandour et al., 2003; Van Lancker, 1980). This finding is consistent with the previous claim that the lack of left-hemisphere dominance for pitch accent was presumably due to its lighter linguistic functional use compared to tone (Hayashi et al., 2001). The similar right-ward processing across groups in the current study, including the Chinese and English learners, further suggests that experience with processing and learning low-function prosodic contrasts may not necessarily lead to left-hemisphere dominance, as is also shown by previous intonation studies (Gandour et al., 2003a, 2007). In such cases, acoustic-level processing is dominant for listeners across L1 and L2 backgrounds.

Despite the primary findings of acoustic-level processing across groups, listener group differences in the perception of individual pitch accent patterns also provide some evidence for the effects of linguistic experience. These patterns may reflect differences in acoustic cue-weighting as a function of linguistic experience, where listeners' sensitivity to the acoustical variations of pitch accent is triggered by the linguistic functions of their L1 prosodic patterns. As the results show, the Chinese listeners as well as the Japanese listeners perceived the H*L pattern (with a falling F0 contour) better than both LH* and LH (both with a rising F0 contour), while their perception of the latter two did not differ. As discussed above, the native Japanese patterns are consistent with the previous findings that H*L is more perceptually salient than LH* and LH, while the latter two are not easily distinguishable (Hirano-Cook, 2011; Maniwa, 2002; Sugiyama, 2006). Similarly, in Mandarin Chinese, the falling tone contour has also shown to be perceptually more salient than the rising tone contours, particularly for challenging listening tasks such as dichotic listening (Wang et al., 2001) or in noisy listening backgrounds (Liu, Azimi, Bhandary & Hu, 2014). This may account for the Chinese listeners' more accurate perception of the falling H*L than the rising LH* and LH pitch accents. That the perception of both LH* and LH was poor may also have been due to the confusion between these two patterns, since fine-tuned F0 differences between the two rising contours do not involve tonal category distinctions (Wu, Munro & Wang, 2014). This is in line with previous findings that tonal L1 listeners exhibit decreased sensitivity to within-category F0 variations (Gandour, 1983; Huang & Johnson, 2010; Lee, Lekich & Zhang, 2014). In contrast, the English listeners' perception of H*L was better than that of LH, but not than LH*. The acoustic analysis showed that the F0 of the second syllable is only different between H*L and LH*, but not between H*L and LH. It is conceivable that the falling (H*L) and rising (LH*) contours are familiar to the

English listeners from their experience with F0 patterns in declarative and interrogative intonation, respectively (cf. Braun & Johnson, 2011; Broselow, Hurtig, & Ringen, 1987), allowing them to perceive the two equally well. However, the lack of difference in 2nd-syllable F0 between LH and H*L reduces the acoustic distinctiveness of LH, presumably leading to its poorer perception. Thus, the English listeners' perception of different pitch accent patterns reflects an integration of linguistic- and acoustic-level influence.

Moreover, the group differences in the distribution of left- and right-hemisphere advantage further reflect the influence of both L1 and L2 learning experience. First, as Table 3 shows, only the naïve English group demonstrated both a greater distribution of right-hemisphere preference for H*L and a greater distribution of left-hemisphere preference for LH*. These results indicate that naïve English listeners' perception may have relied on the left hemisphere (as well as the right hemisphere) to a greater extent as compared to the other groups, presumably due to the fact that pitch accent distinctions were most challenging for the naïve English listeners who were least experienced among all the listeners and whose processing thus required more brain resources. Previous research indeed suggests that perceptual difficulty in prosody may increase activation in the non-dominant as well as dominant hemisphere (Aasland & Baum, 2003). Particularly, lexical tone studies with inexperienced listeners reveal that poorly perceived tones tend to show a greater degree of left-hemisphere involvement (Wang et al., 2004). In contrast, it is noteworthy that the English learners of Japanese exhibited an approximation to the native direction. Following the native Japanese patterns (and in contrast to the naïve English group), the English learners' ear-preference distribution was not biased in favor of the left-hemisphere for LH*, indicating a reduced degree of reliance on the left hemisphere. This result is consistent with previous findings of lexical tone processing, where experienced bilingual listeners relative to naïve listeners demonstrated more native-like lateralization patterns (Wang et al., 2004).

5. Concluding remarks

Taken together, the current results indicate that the processing of Japanese pitch accent is characterized primarily by its acoustic features, although also affected by listeners' linguistic experience both with their L1 and L2 learning. With regard to acoustic aspects, the results indicate consistency in processing patterns for listeners from diverse language backgrounds, suggesting a primary reliance on acoustic properties, probably in company with the relatively low linguistic functional load of pitch accent. Moreover, results of individual pitch accent patterns reveal different processing patterns due to their acoustic

differences, with the most acoustically salient pattern (H*L) relying primarily on sensory-acoustic processing and the less acoustically distinctive patterns (LH*, LH) additionally engaging linguistic influences. In light of the influence of L1 and L2 experience, the results show a gradation of hemispheric processing changes among the listener groups towards the native Japanese patterns as a function of linguistic experience: from the native Chinese listeners who, due to their L1 experience with lexical tone, followed the native patterns, to the English learners whose L2 Japanese-learning experience geared them towards the native direction, and to the naïve English listeners whose processing was most resource-demanding due to their lack of either L1 tonal or L2 Japanese experience.

The current findings of pitch accent perception involving acoustic processing interacting with both L1 and L2 experience thus provide new data to complement the previous pitch accent studies focusing on native processing and L1 effects (Kanamura & Imaizumi, 2008; Wu et al., 2012) as well as the previous studies on lexical tone which carries a heavier linguistic functional load than pitch accent (e.g., Wang et al., 2001, 2004). Together, the data from different listener backgrounds (natives, non-natives, learners) and different stimuli functional types (tone, pitch accent) implicate that hemispheric processing of lexical prosody involves multi-level and dynamic processes, differing as a function of acoustic realizations, as well as levels of linguistic functional load, experience, and learning.

Lastly, it should also be noted that some of the current findings would warrant further research. For instance, the current results indicate that learners' processing may become more efficient (involving reduced participation of the non-dominant hemisphere) as they gain more experience with L2 pitch accents. However, previous neuroimaging research on lexical tones has also indicated that such learning-induced efficiency may also involve reduced activity in particular cortical areas (Wang et al., 2003; Wong et al., 2007a; Wong et al., 2008). Although dichotic listening has been a commonly used technique that provides straightforward patterns to indicate hemispheric lateralization of speech prosody, the limitations of this technique also prevent further investigations of regional cortical changes in prosodic learning. Furthermore, it has also been shown that the patterns exhibited from dichotic listening testing may not only be triggered by hemispheric lateralization patterns, but may also be affected by structural and functional involvement of the corpus callosum as information is transferred from one hemisphere to other (Westerhausen & Hugdahl, 2008). Thus, further research on non-native pitch accent processing and learning may utilize neuroimaging and electrophysiological techniques to acquire new learner data to complement the current behavioral findings.

Appendix A. Average F0 values at 0%, 25%, 50%, 75% and 100% positions of the first and second syllables of the target words with the H*L, LH* and LH pitch accent patterns. SDs are shown in parentheses.

	1 st syllable					2 nd syllable				
	0%	25%	50%	75%	100%	0%	25%	50%	75%	100%
H*L	233 (22)	238 (22)	250 (18)	256 (15)	248 (18)	236 (19)	213 (15)	184 (8)	156 (7)	155 (58)
LH*	187 (13)	183 (10)	179 (9)	179 (9)	168 (10)	210 (35)	218 (23)	226 (15)	231 (5)	233 (36)
LH	181 (13)	171 (7)	170 (6)	168 (4)	162 (12)	199 (24)	204 (14)	207 (9)	207 (6)	191 (7)

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.
- Beckman, M. E. (1986). *Stress and Non-stress Accent*. Dordrecht, Holland: Foris.
- Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.51, <http://www.praat.org/>.
- Braun, B., & Johnson, E. K. (2011). Question or tone? How language experience and linguistic function guide pitch processing. *Journal of Phonetics* 39, 585–594.
- Broselow, E., Hurtig, R. R., & Ringen, C. (1987). The perception of second language prosody, in: Georgette Ioup and Steven H. Weinberger, eds., *Inter-language Phonology: The Acquisition of Second Language Sound System*, Cambridge: Newbury House Publishers, 350–361.
- Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007). Experience-dependent neural plasticity is sensitive to shape of pitch contours. *Neuroreport*, 18, 1963–1967.
- 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.
- Gandour, J. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics*, 11, 149–175.
- Gandour, J., Wong, D., & Hutchins, G. (1998). Pitch processing in the human brain is influenced by language experience. *Neuroreport*, 9, 2115.
- Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D., & Hutchins, G. D. (2000). A cross-linguistic PET study of tone perception. *Journal of Cognitive Neuroscience*, 12, 207–222.
- Gandour, J., Wong, D., Lowe, M., Dzemidzic, M., Sathammuwong, N., Tong, Y., & Li, X. (2002). A cross-linguistic fMRI study of spectral and temporal cues underlying phonological processing. *Journal of Cognitive Neuroscience*, 14, 1076–1087.
- Gandour, J., Wong, D., Dzemidzic, M., Lowe, M., Tong, Y., & Li, X. (2003a). 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., Sathammuwong, N., & Lurito, J. (2003b). 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., Li, X., & Lowe, M. (2004). Hemispheric roles in the perception of speech prosody. *Neuroimage*, 23, 344–357.
- Gandour, J., Tong, Y., Talavage, T., Wong, D., Dzemidzic, M., Xu, Y., Li, X., & Lowe, M. (2007). Neural basis of first and second language processing of sentence-level linguistic prosody. *Human Brain Mapping*, 28(2), 94–108.
- Gu, F., Zhang, C., Hu, A., & Zhao, G. (2013). Left hemisphere lateralization for lexical and acoustic pitch processing in Cantonese speakers as revealed by mismatch negativity. *Neuroimage*, 83, 637–645.
- 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.
- Hasegawa, Y., & Hata, K. (1992). Fundamental frequency as an acoustic cue to accent perception. *Language and Speech*, 35, 87–98.
- 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.
- Hirano-Cook, E. (2011). Japanese pitch accent acquisition by learners of Japanese: Effects of training on Japanese accent instruction, perception, and production. Unpublished Ph.D. Dissertation, University of Kansas.
- Hsieh, L., Gandour, J., Wong, D., & Hutchins, G. D. (2001). Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience. *Brain and Language*, 76(3), 227–252.
- Huang, T., & Johnson, K. (2010). Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners. *Phonetica*, 67, 243–267.
- Ivry, R. B., & Robertson, L. C. (1998). *The two sides of perception*. Cambridge, MA: The MIT Press.
- 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.

- Kanamura, R., & Imaizumi, S. (2008). Linguistic versus non-linguistic processing of speech prosody in dichotic listening. *Acoustics 08 Paris*, 8733–8737.
- Kindaichi, H. (1967). *Nihongo on'in no kenkyu* [Study of Japanese Phonology]. Tokyo: Tokyodo Shuppan.
- Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, 15(3), 166.
- King, R. D. (1967). Functional load and sound change. *Language*, 831–852.
- Kitahara, M. (2001). Category structure and function of pitch accent in Tokyo Japanese. Unpublished Ph.D. 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.
- Koso, A., & Hagiwara, H. (2009). Event-related potential evidence of processing lexical pitch accent in auditory Japanese sentences. *Neuroreport*, 20, 1270–1274.
- Lee, C. Y., Lekich, A., & Zhang, Y. (2014). Perception of pitch height in lexical and musical tones by English-speaking musicians and nonmusicians. *Journal of the Acoustical Society of America*, 135, 1607–1615.
- Lehiste, I. (1970). *Suprasegmentals*, Cambridge, MA: MIT Press.
- Li, X., Gandour, J., Talavage, T., Wong, D., Hoffa, A., Lowe, M., & Dziedzic, M. (2010). Hemispheric asymmetries in phonological processing of tones vs. segmental units. *Neuroreport*, 21(10), 690–694.
- Liu, C., Azimi, B., Bhandary, M., & Hu, Y. (2014). Contribution of low-frequency harmonics to Mandarin Chinese tone identification in quiet and six-talker babble background. *Journal of the Acoustical Society of America*, 135 (1), 428–438.
- Luo, H., Ni, J.-T., Li, Z.-H., Li, X.-O., Zhang, D.-R., Zeng, F.-G., & Chen, L. (2006). Opposite patterns of hemisphere dominance for early auditory processing of lexical tones and consonants. *Proceedings of the National Academy of Sciences of the United States of America*, 103 (51), 19558–19563.
- 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.
- Mildner, V. (2004). Hemispheric asymmetry for linguistic prosody: A study of stress perception in Croatian. *Brain and Cognition*, 55, 358–361.
- 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, MA: MIT Press.
- Poeppl, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication*, 41(1), 245–255.
- Sato, Y., Sogabe, Y., & Mazuka, R. (2007). Brain responses in the processing of lexical pitch-accent by Japanese speakers. *Neuroreport*, 18, 2001–2004.
- Shibata, T., & Shibata, R. (1990). Akusento wa doo'ongo o donoteido benbetsu shiuruno ka: Nihongo, eigo, chuugokugo no baai (How much can accent distinguish homophones cases of Japanese, English and Chinese). *Keiryoo Kokugo-gaku (Mathematical Linguistics)*, 17, 317–327.
- 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.
- 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*. John Benjamins.
- Sugito, M. (1972). Ososagari-koo: Dootai-sokutei ni yoru nihongo akusento no kenkyu [Delayed pitch fall: An acoustic study]. Shoin Joshi Daigaku Ronshuu 10. Reprinted in M. Tokugawa (Ed.), *Akusento [Accent]* (pp. 201–229). Tokyo: Yuuseidoo, 1980.
- 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]. *Kotoba to onsei ('Sotoba' Series 18)*, 23–37.
- Sugiyama, Y. (2006). Japanese pitch accent: Examination of final-accented and unaccented minimal pairs. *Toronto Working Papers in Linguistics*, 26, 73–88.
- Sugiyama, Y. (2008). The nature of Japanese pitch accent: An experimental study. Unpublished Ph.D. Dissertation. The State University of New York at Buffalo.
- Sugiyama, Y. (2014). Perceiving pitch accent in the absence of F0. *Proceedings of the annual meetings of the Berkeley Linguistics Society*, 482–493.
- Tamaoka, K., Saito, N., Kiyama, S., Timmer, K., Verdonschot, G., Timmer, K., & Verdonschot, R. (2014). Is pitch accent necessary for comprehension by native Japanese speakers? – An ERP investigation. *Journal of Neurolinguistics*, 27, 31–40.
- Tong, Y., Gandour, J., Talavage, T., Wong, D., Dziedzic, M., Xu, Y., Li, X., & Lowe, M. (2005). Neural circuitry underlying sentence-level linguistic prosody. *Neuroimage*, 28, 417–428.
- 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.
- 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, X., Wang, Y., & Chen, L. (2013). Hemispheric lateralization for early auditory processing of lexical tones: Dependence on pitch level and pitch contour. *Neuropsychologia*, 51, 2238–2244.

- Wang, Y., Behne, D. M., Jongman, A., & Sereno, J. A. (2004). The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Psycholinguistics*, 25(3), 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(7), 1019–1027.
- Weintraub, S., Mesulam, M.-M., & Kramer, L. (1981). Disturbances in prosody: A right-hemisphere contribution to language. *Archives of Neurology*, 38, 742–745.
- Westerhausen, R., & Hugdahl, K. (2008). The corpus callosum in dichotic listening studies of hemispheric asymmetry: a review of clinical and experimental evidence. *Neuroscience and Biobehavioral Reviews*, 32, 1044–1054.
- Wong, P. (2002). Hemispheric specialization of linguistic pitch patterns. *Brain Research Bulletin*, 59(2), 83–95.
- Wong, P., Parsons, L. M., Martinez, M., & Diehl, R. L. (2004). The role of the insular cortex in pitch pattern perception: the effect of linguistic contexts. *The Journal of Neuroscience*, 24(41), 9153–9160.
- Wong, P., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, 28, 565–585.
- Wong, P., Perrachione, T. K., & Parrish, T. B. (2007a). Neural characteristics of successful and less successful speech and word learning in adults. *Human Brain Mapping*, 28, 995–1006.
- Wong, P., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007b). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, 10, 420–422.
- Wong, P., Warrier, C. M., Penhune, V. B., Roy, A. K., Sadehh, A., Parrish, T. B., & Zatorre, R. J. (2008). Volume of left Heschl's Gyrus and linguistic pitch learning. *Cerebral Cortex*, 18, 828–836.
- Wu, X. (2013). A Cross-Language Investigation of Phonetic and Phonological Processing of Lexical Tone. Ph.D. dissertation, Simon Fraser University.
- Wu, X., Munro, J. M., & Wang, Y. (2014). Tone assimilation by Mandarin and Thai listeners with and without L2 experience. *Journal of Phonetics*, 46, 86–100.
- Wu, X., Tu, J. Y., & Wang, Y. (2012). Native and nonnative processing of Japanese pitch accent. *Applied Psycholinguistics*, 33, 623–641.
- Xu, Y., Gandour, J., Talavage, T., Wong, D., Dziedzic, M., Tong, Y., Li, X., & Lowe, M. (2006). Activation of the left planum temporale in pitch processing is shaped by language experience. *Human Brain Mapping*, 27, 2, 173–183.
- Zatorre, R. J., & Gandour, J. (2008). Neural specializations for speech and pitch: Moving beyond the dichotomies. *Philosophical Transactions of the Royal Society*, 363, 1087–1104.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 11, 946–953.
- 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(3), 624–633.