DOI: 10.6503/THJCS.2016.46(4).04

Prosody Transfer in Second Language Acquisition: Tonal Alignment in the Production of English Pitch Accent by Mandarin Native Speakers*

Lu Yu-an
Department of Foreign Languages & Literatures
National Chiao Tung University

Kim Miran**
Department of English Education
Gyeongsang National University

ABSTRACT

This study examines how the tonal alignment in L1 influences the production of L2 prosody, focusing on the realization of English L*+H pitch accent (i.e., listing contour) produced by Mandarin Chinese speakers. While Mandarin uses pitch mainly to signal *lexical* contrast, English uses it to convey *discourse/pragmatic* meaning (*post-lexical*). It has been observed in the literature that the F0 of lexical tone in Mandarin is syllable bound, while English pitch accent is contextually governed by various phonetic factors, including speech rate and syllable duration. Given these functional and prosodic differences, we ask whether prosodic transfer at the level of tonal alignment occurs across languages in which the grammatical function of pitch in L1 (Mandarin) differs fundamentally from that in L2 (English). Our results show that the L1 alignment pattern (syllable-bound) remains dominant in Mandarin native speakers' production of English listing contour; additionally, the speech rate effect on tonal timing is less consistent in the Mandarin production. Implications on second language learning and pedagogy are discussed.

Key words: second language learning, tonal alignment, L1 transfer

---

* We would like to thank the three anonymous reviewers and the editors for their valuable comments and suggestions. Any remaining errors are ours.

** Author's email address: yuanlu@nctu.edu.tw; mirankim@gnu.ac.kr (corresponding author)

1. Introduction

In the last decades, research on second language (L2) learning/acquisition and loanword adaptation has drawn substantial attention in the fields of pedagogy and theoretical linguistics. The main focus, however, has mostly been on segmental and phonotactic adaptations (e.g., Weinberger 1996; Brown 1997; Dupoux et al. 1999; Brannen 2002; Y. Kang 2003; Miao 2005; Kim 2006; Lu 2009; Eckman & Iverson 2013). Suprasegmental adaptation has been relatively understudied, as Y. Kang (2010) notes. The current paper reports findings of an investigation into a case of suprasegmental adaptation at the level of tonal alignment (i.e., "temporal implementation of fundamental frequency (F0) movements with respect to the segmental string" (Prieto 2011: 1185)) in Mandarin Chinese (hereafter, Mandarin) and English, with a view to improving our understanding of the influence of L1 prosody on L2 production.

The two languages chosen as targets for this research, Mandarin and English, display fundamentally different pitch functions. Mandarin uses pitch to signal *lexical* contrast (i.e., lexical tones; Duanmu 2007), while English uses it to convey *discourse/pragmatic* meaning *post-lexically* (i.e., intonation; Pierrehumbert 2000). Second, the segmental alignment of pitch targets (e.g., F0 maxima [H], and minima [L]) has been observed to be different in the two languages. Xu (1998) reports that the F0 contours of lexical tones in Mandarin are *syllable bound* and align to the end of tone-bearing syllables. This alignment remains stable regardless of contextual differences, such as different speaking rates and segmental makeups (Xu 1998; Xu & Wang 2001). The domain of English intonation pitch contour, on the other hand, is *phrase bound*; its alignment is shown to be contextually governed by speaking rate and syllable duration (Silverman & Pierrehumbert 1990). In light of these prosodic differences, the present study investigates the possibility of prosodic transfer at the level of tonal alignment. What is the sort of transfer effect that arises when the grammatical function of pitch in a speaker's first language (L1) (i.e., Mandarin) differs fundamentally from its function in the L2 (i.e., English)? Our results show that the L1 alignment pattern (syllable-bound) remains dominant in L2 production for Mandarin speakers, and that tonal timing in the production of L2 English by Mandarin speakers is less affected by contextual differences than it is for English native speakers. The prosodic characteristics found in L2 production in our study suggest that F0 timing may be subject to prosody transfer in L2 learning; we therefore identify L1/L2 prosodic discrepancy as one of the key linguistic domains which L2 learners need to be aware of to improve L2 proficiency.

This paper is structured as follows. Section 2 reviews the literature on segmental and suprasegmental adaptation in L2 learning and loanword adaptation, and provides cross-linguistic findings regarding F0 alignment patterns. It also provides a brief summary of the phonological differences between Mandarin and English in terms of the pitch functions and F0 timing behaviors reported in the literature. Section 3 details our production experiment conducted to understand how F0 alignment differs in L2 versus L1 production, and reports our results concerning the tonal alignment of English L*+H pitch accent produced by English native speakers and Mandarin native speakers. We also discuss potential differences in F0 timing variation induced by *speech rate*, one of the phonetic factors shown to influence F0 alignment (e.g., Steele 1986). Section 4 summarizes and discusses the prosodic behaviors of L2 learners in terms of phonetic F0 alignment and speech rate. We conclude the paper by discussing implications of the findings on second language learning and raising possible directions for future studies.

2. Research Background

In this section, we combine the reviews of second language learning/acquisition and loanword adaptation due to their shared characteristics. Loanword adaptation generally occurs under two scenarios. The first one is identified as adaptation through production, in which "the borrowing may be implemented by a bilingual speaker that fills a gap in one of the languages he knows, L1, the recipient language, by taking a word from the other language he knows, L2, the donor language" (Calabrese & Wetzels 2009: 1). The second scenarios is identified as adaptation through perception, in which "the borrowing is implemented by a speaker that fills a gap in his language by taking a word from another language he knows poorly or not at all" (Calabrese & Wetzels 2009: 2). The first scenario resembles a learning situation by a proficient second language learner while the second resembles that by a beginner second language learner. More importantly, as in second language learning, L1 knowledge is heavily dependent upon (in production for the first scenario and in perception in the second scenario).

2.1 L2 Learning on the Segmental and Suprasegmental Levels

Though L2 speech characteristics are complex phenomena that are difficult to account for by one single factor, explanations of L2 speech characteristics offered in the literature generally appeal to three main phenomena: (1) L1 transfer—carry-over effects from the *structural differences* between L1 and L2 (e.g., Eckman 1977; Tarone 1987; Hancin-Bhatt 1994; Schwartz & Sprouse 1994; Gass 1996; Miao 2005), (2) *phonetic/acoustic* similarities between L1 and L2 (e.g., Y. Kang 2003; LaCharité & Paradis 2005; Boersma 2009; Broselow 2009), and (3) universal tendencies—what is considered to be common or 'unmarked' across languages (e.g., Broselow & Finer 1991; Broselow et al. 1998; Broselow 2004; Escudero & Boersma 2004).[1]

The first approach, L1 transfer, refers to native language carry-over properties found in the speech productions of second-language learners. Whether the effect is so-called "positive" or "negative", L1 transfer is widely identified in the early stages of most second-language learning. For example, LaCharité and Paradis (2005) report that French speakers tend to categorize English /b/ as French /b/, even though English /b/ is acoustically more similar to French /p/ (in the sense that both have a short-lag voice onset time (VOT)). In other words, the two-way contrast of stops in English (L2) is adapted into the French two-way contrast of stops (L1), regardless of the phonetic approximation of English /b/ and French /p/. LaCharité and Paradis (2005: 223) conclude from their findings that "loanword adaptation is overwhelmingly phonological and phonetic approximation plays a limited role…", supporting the view that structural differences between L1 and L2 have a significant impact on L2 adaptation. Another L1 transfer example comes from differences in the characterization of plosives between Mandarin and English. Both languages have a two-way contrast of stops. However, Mandarin distinguishes aspirated from unaspirated stops (e.g., /p^h/ and /p/) while English distinguishes voiced from voiceless stops (e.g., /b/ and /p/). Aspirated and unaspirated

1 For a more detailed review, see Y. Kang (2011).

voiceless stops are considered allophones of the same phoneme (/p/, in this case) in English. In a corpus study, Paradis and Tremblay (2009) show that English voiceless stops (aspirated and unaspirated) are systematically adapted into Mandarin pronunciation as aspirated stops, while English voiced stops are adapted as unaspirated stops. This, again, lends support to the legitimacy of L1 transfer effects. The two cases reviewed here are ones in which the native structure prevents the faithful mappings of foreign sounds, and thus are considered to be cases of negative transfer.

Other studies have relied mainly on phonetic/acoustic similarities to explain general loanword adaptation patterns. For example, Kenstowicz and Suchato (2006) investigate adaptation patterns from English to Thai. As noted above, English has a two-way contrast of stops (e.g., /p/ and /b/) whereas Thai has a three-way contrast (e.g., /p^h/, /p/, and /b/). In a corpus study, the authors report that the aspirated allophone [p^h] in English is more likely to be adapted as /p^h/ by Thai native speakers, while the unaspirated allophone [p] is more likely to be adapted as /p/. For example, in English /sC/ clusters, the C is uniformly produced as a voiceless unaspirated stop and when voiceless stops are followed by unstressed vowels, the stops are more likely to be produced as unaspirated. These contexts are the ones in which the voiceless stops are unaspirated or weakly aspirated and Thai speakers are sensitive to these acoustic details. The phonetic/gradient nature (versus a phonological/categorical nature) suggests an effect of phonetic similarity on L2 learning.

Some cases of loanword adaptation cannot be explained by either of the above approaches. For example, Mandarin does not allow obstruents of any kind in coda position, due to strict restrictions on syllable structure phonotactics. Broselow et al. (1998) report that Mandarin learners of English tend to correctly produce words with voiceless obstruent codas, whereas words with voiced obstruent codas tend to be produced less faithfully. Voiced obstruents are more marked in coda position cross-linguistically (e.g., Charles-Luce 1985; Port & O'Dell 1985; Kawahara & Garvey 2010). The researchers draw on the notion of a typologically common or "unmarked" grammar/constraint hierarchy deeply embedded in our linguistic system to explain this contrast.

While segmental and phonotactic learning/adaptation have been extensively examined, as the discussion above shows, research on suprasegmental learning/adaptation has been less studied. Questions such as what effects and to what extents L1 transfer, phonetic proximity, universality are subject to suprasegmental learning remained to be answered. Y. Kang (2010) reviews several cases of suprasegmental adaptation in the literature and show that, when L1 and L2 share the same suprasegmental system (e.g., tone language to tone language, stress language to stress language), two major patterns of adaptation arise. First, L1 transfer may occur. An example comes from a prosodic comparison of Swedish (a final-stress language) and Finnish (an initial-stress language) (Fenyvesi & Zsigri 2006). When learning Swedish, Finnish speakers tend to transfer their L1 knowledge of initial stress to their L2 production (e.g., Swedish [mus**í**k] 'music' becomes [m**ú**siːkki] in Finnish). A similar phenomenon occurs when Huave speakers, whose native language has predictable stress (final stress on words ending in closed syllables; penultimate stress on words ending in open syllables), adapt Spanish words whose stress does not follow the native stress assignment pattern. In this case, the Huave speakers tend to alter the segmental contents of Spanish words to follow the native stress patterns (e.g., Spanish [garab**á**to] 'hook' becomes [garab**á**t] in Huave, with the final vowel deleted to form a closed syllable) (Broselow 2009). The second pattern of adaptation between suprasegmentally consistent languages capitalizes on phonetic/ acoustic similarities between L1 and L2. For example, when Gwari speakers learn Hausa, they consistently adapt the Hausa high tone to either a high or mid tone and the Hausa falling tone to either a high-falling or mid-falling tone (Maddieson 1977).

However, when L1 and L2 do not share the same suprasegmental type (e.g., tone language to stress language and vice versa), universal tendencies seem to play a greater role. For example, when Tibetan (a tone language) borrows from English (a stress language), the stress placement of the source words is completely ignored. Instead, non-initial syllables are given high tones and initial syllables are given high or low tones depending on the voicing of the onsets: H when the onsets are voiceless and aspirated (e.g., police → **pu**$^{\mathbf{H}}$li^{H}si^{H}) and L when the onsets are voiced (e.g., bottle → **po**$^{\mathbf{L}}$to^{H}ra^{H}), following the principles of tonogenesis (Hsieh & Kenstowicz 2008). This strategy is not used in native grammar and is argued to be a UG enhancement process that reinforces or replaces a voicing contrast in the onset of the syllable found in the emergence of tone cross-linguistically. Similarly, when Japanese (a pitch-accent language) borrows words from French (a stress language), the stress from the source words is again ignored.

Instead, a default pattern (extrametrical final mora and moraic trochee on the right edge) is assigned (Shinohara 2004). This finding mimics the unmarked pattern in first language acquisition (Demuth 1995, 1996). Likewise, when Taiwanese (a tone language) borrows words from Japanese, the pitch accent pattern is ignored. Instead, tone patterns are assigned based on the rhyme structure: contour tones are assigned to longer syllables (e.g., CVG/CVN) while level tones are assigned to shorter syllables (e.g., CV/CVO) (Hsieh 2006). This follows the cross-linguistic pattern in which longer rhymes (e.g., rhymes with sonorant codas, stressed syllables, phrase-final position, shorter words) are more likely to host dynamic tones while shorter rhymes (e.g., rhymes with obstruent codas, unstressed syllables) are more likely to host static tones (Zhang 2004). Although cases of L2 adaptation based on phonetic/acoustic similarities between two languages with different suprasegmental functions have been reported (e.g., Wu 2006; Ou 2010), universal tendencies seem to dominate in this type of the L2 learning/adaptation. In the current study, we examine another suprasegmental aspect—tonal alignment—between two languages with different suprasegmental functions, in order to investigate whether L1 prosodic transfer occurs in this context as well.

In the next section, we present an overview of the literature on tonal alignment.

2.2 F0 Alignment as a Phonological Contrast and Cross-linguistic Differences

Tonal alignment plays a crucial role in encoding intonational contrasts. In recent decades, a number of experiments have shown that F0 alignment patterns can show categorical distinctions within a language in much the same way that segmental contrasts can (e.g., Prieto 2011; Ramijsen 2013). These findings suggest strongly that tonal alignment differences are indeed encoded phonologically. For example, the placement of F0 targets (either H or L) within a segment string—i.e., the "timing" of F0 within the tone-bearing syllable—functions contrastively in English; Pierrehumbert (1980) has shown that early-aligned pitch accents are phonologically distinct from late-aligned pitch accents. Building on this discovery, Pierrehumbert and Steele (1989) asked English listeners to imitate pitch steps along a synthesized continuum of F0 alignments ranging from early (L+H*) to late (L*+H) pitch accents. They found that the participants did not use the full range of the continuum in their imitation performances, but instead showed a bimodal distribution of H peaks; these findings lend further support to the categorical nature of perception for tonal alignment. Other studies similarly found that timing differences among identical tonal targets are categorically perceived by native listeners of languages where timing encodes lexically or pragmatically distinct information (e.g., Fujisaki Hiroya 1983; Kohler 1987; D'Imperio 1999; Dilley 2007). Additionally, Prieto et al. (1995) show that three Romance languages (Central Catalan, Neapolitan Italian, and Pisa Italian) share a rising pitch accent (LH) but encode F0 alignment targets differently according to different prosodic boundary levels (i.e., mora, syllable, and word).

The research highlighted above shows that the alignment of F0 vis-à-vis segmental strings can help to characterize the prosodic system of a language in a significant way; such phonological properties are part of native speakers' knowledge, and may not be automatic to non-native language learners. When it comes to loanword adaptation, not only segmental contrasts, but also the phonetic manifestation of tonal alignment may therefore be significant for L2 learners. We know from the literature that the phonetic manifestations of representationally equivalent tones (i.e., H tone) may not be identical across languages, resulting in a seemingly unfaithful tonal mapping in loanword adaptation (Maddieson 1977; Yip 2002; Y. Kang 2010). For instance, the H tone of [$du^{H}bu^{H}$] 'thousand' in Hausa is adapted as an M tone in Gwari ([$du^{M}bu^{M}$]) and the HL tone pattern of [te^{H}: $bu^{L}r$] 'table' in Hausa becomes ML [$te^{M}bu^{L}l$] in Gwari, even though Gwari has a phonological H tone (examples are cited from Y. Kang 2010; originally from Maddieson 1977). Together with the fact that a single pitch contour (e.g., rising accent) may show different F0 alignments across different languages, these examples indicate that the adaptation process may involve more detailed phonetic and phonological properties than the apparent unidimensional contrasts in the languages at hand. Thus, research on tonal alignment in second language production may be able to further reveal the minor linguistic details, systematically encoded in the phonology of L1, that influence non-native learners phonetic acquisition. This paper pursues this direction of research by examining Mandarin speakers' production of non-native pitch accents (i.e., the listing contour L*+H in English) and investigating F0 alignment as a case of prosodic transfer.

The next section provides a brief overview of phonetic factors that may influence F0 alignment.

2.3 Phonetic Factors That Influence F0 Alignment

A number of studies have shown that F0 alignment can be affected by phonetic factors, such as right-edge prosodic boundaries (e.g., how many post-nuclear syllables), syllable structure, segmental duration, and speaking rate (e.g., Silverman & Pierrehumbert 1990; Pierrehumbert 2000). Steele (1986) investigates the timing of the F0 alignment for nuclear H* pitch accent in English as a function of the number of post-nuclear syllables (syllables following the H*) and speech rate. The results show that the F0 peak of the nuclear H* is much earlier, relative to the total vowel duration, when no post-nuclear syllables than when there do. Steele takes this finding as evidence for an effect of right-edge materials on tonal alignment. The same study also observes a speech rate effect, in which slower speech induces a later alignment of nuclear H*. Similarly, Silverman and Pierrehumbert (1990), in a production experiment testing English speakers' F0 peak placement, show that "when a syllable is lengthened from being spoken more slowly, the peak will occur corresponding later," suggesting a speech rate effect on English speakers' production of pitch targets (Silverman & Pierrehumbert 1990: 94).

Tonal targets in Mandarin, on the other hand, are shown to be less affected by phonetic factors. Xu (1998) conducted a production experiment that asked Mandarin speakers to produce the four lexical tones at different speaking rates (i.e., slow, normal, fast) in syllables with different segmental makeups (CV, CVN). He found that tonal targets in Mandarin speakers' production aligned consistently with the ends of the syllables that carried the tones—regardless of speaking rates and segmental makeup—contrary to the results reported for English speakers' production.

The present paper investigates speech rate as a phonetic factor influencing F0 alignment variations in the speech of English and Mandarin native speakers. By comparing variations in the productions of these two speaker groups, we hope to gain a better understanding of how F0 alignment is affected by contextual differences in L2 production and whether tonal alignment in L2 production shows evidence of L1 prosodic transfer.

The above review showed that tonal alignment—the set of principles a language relies on to map tone targets temporally to the segmental string—is phonologically controlled, just like segmental contrast. It is therefore possible that L1/L2 differences in tonal alignment may be subject to L1 transfer interference. As mentioned earlier, Mandarin uses pitch to signal *lexical* contrast, while English uses it to convey *post-lexical* pragmatic meaning; furthermore, lexical tones in Mandarin are *syllable bound* and aligned at the end of the tone-bearing syllables. This tonal alignment of F0 contours in Mandarin remains consistent regardless of contextual differences such as speech rate and segmental makeup. Conversely, English intonation pitch contour is *phrase bound* and its alignment is contextually governed by speech rate and syllable duration (Silverman & Pierrehumbert 1990). These fundamental differences inspired a production experiment to examine how L1 tonal alignment prosody influences the production of L2 tonal alignment prosody. The salient differences between the two languages are summarized in Table 1.

Table 1: Prosodic differences between Mandarin and English

	Mandarin	English
	Lexical tone	Intonation
Pitch function	Lexical contrast	Pragmatic contrast
Prosodic domain	Syllable	Accented unit (e.g., prosodic word, phrase)
Tonal alignment	Syllable bound: F0 contours are aligned at the end of the tone-bearing syllable regardless of different speech rates or segmental makeups. (Xu 1998; Xu & Wang 2001)	Metrical structure: The alignment of intonation peaks with their syllables exhibits contextually governed variation. (Silverman & Pierrehumbert 1990)

3. Experiment

The goal of this study is to investigate how Mandarin speakers, whose native language utilizes pitch to indicate lexical differences, produce English pitch accent, the function of which is predominantly pragmatic. We are interested in whether prosodic transfer occurs in the area of tonal alignment when the grammatical function of pitch in the L1 (Mandarin) differs fundamentally from that in the L2 (English).

3.1 Method

3.1.1 Participants

A total of 18 female participants (6 English native speakers and 12 Mandarin native speakers) were recruited at Stony Brook University and received course credit or payment for their participation. The 6 participants in the English L1 group (aged 18-32) were all monolingual English speakers. The 12 participants in the Mandarin L1 group (aged 19-32) were all native speakers of Beijing Mandarin (BM) or Taiwanese Mandarin (TM), and had received up to a high school education in China or Taiwan before coming to the United States. Their exposure to an English-speaking environment was 1-2 years (6 speakers), 3-6 years (4 speakers), or 6-14 years (2 speakers) at the time of recording.

3.1.2 Experimental design and materials

The English L*+H pitch accent, typically produced during listing, was used to investigate tonal alignment in the production of the two sets of speakers. The reason of choosing L*+H as the research target is two-fold. First, Mandarin has the tone pattern that is similar to the accent type (i.e., low-rising T3+T2) and thus no explicit guide is needed in conducting the experiment. If explicit guide were to be provided, possible confound from the instruction would not be controlled for. Second, L*+H accent in English is one of the most-studied intonation. For example, in English, there is a contrast between L*+H and L+H*, both of which are associated with metrically strong syllable (i.e., stressed syllable). As shown in Pierrehumbert and Steele (1989: 182), these two are characterized by late-aligned (L*+H) and early-aligned (L+H*) pitch pattern. This distinction was also tested through a production (imitation) task to see whether these two alignment patterns are indeed categorical, and they found that native speakers could reproduce the two categories not as a gradient fashion but as a bimodal pattern. From the result of this study as well as a number of other studies (e.g., Kohler 1987; Dilley 2007), we learn the characteristic behavior of L*+H in English and thus use this type of pitch accent for a representative comparison.

The target words were two disyllabic names, *Emma* and *Anna*. These names were chosen because they should yield a one-to-one mapping of tone to syllable for the bi-tonal pitch accent L*+H: the stressed first syllable is expected to be associated with the L* tone, and the unstressed second syllable is expected to be associated with the H tone. The intervening voiced sonorant consonants in these names were chosen in order to minimize segmentally induced perturbations on fundamental frequency between the L* and the H tones. The listing intonation was elicited using a question-and-answer paradigm, shown in Table 2. This paradigm was to provide a natural speech context for the participants, other than just reading alone. In the experiment, recordings were obtained from a role play situation in which the experimenter was a wedding planner and the participant was an assistant. This staged role encouraged the participants to answer the questions in a naturally spoken listing contour during the recording.[2] A practice session was set for the participants to match the names with caricatured pictures. Once the participants were able to recognize all the names by looking at the pictures, the main recording sessions began. To elicit a normal speaking rate, the participants were asked several prepared questions, as in Table 2 with no further instructions. To elicit slow speaking rate, the experimenter would slow down in asking the questions, making sure that the participants answered with no mistakes to allow the experimenter to write down the information. When slow speaking failed, the experimenter would explicitly ask "please make sure that you recognize each of them with no mistake." To elicit fast speaking rate, the experimenter would ask the participants to respond as quickly as possible in order to confirm that the interviewee was ready to start the job.

Table 2: Question-and-answer paradigm

Question	Answer
Who are the bride's friends from high school?	***Anna**, Annie and Amelia*
Who are the bride's friends from kindergarten?	***Emma**, Lily and Olivia*
Who has confirmed to come to the wedding?	***Anna**, Marilyn and Olivia*
Who has not confirmed yet?	***Emma**, Annie and Amelia*
Who are vegetarians?	***Anna**, Norah and Olivia*
Who needs a ride from the airport?	***Emma**, Lily and Olivia*

For example, the experimenter would ask '*who are vegetarians?*' and the participants would answer '*Anna, Norah and Olivia*' based on the figurative information given to them. The target words were embedded as the first word in each list, and were each repeated three times in answering to three different questions. The target words were

[2] The experimenter (either of the authors; a staged interviewer) explained that the participant was staged as an interviewee whose role was to assist in wedding planning and thus should remember some basic information (guest lists, accommodations, etc.).

elicited at three speaking rates (normal, slow and fast), as the staged task requested.

3.1.3 Procedure

The recording was conducted in a sound-attenuated booth using a condenser microphone connected to a Marantz PMD671 recorder with a sampling rate at 44.1 kHz. Participants took part in the recording individually. Each participant was seated comfortably in the booth in front of a computer monitor. After a detailed explanation of the staged task was given to the participant, the recording began with a practice session, followed by three main recording sessions, each of which was designed to obtain spoken data at different speaking rates—slow, normal, and fast. After each of the questions listed in Table 2 was asked, the desired answers were displayed in the form of pictures on the screen, and the participants were asked to produce the names of the characters shown. Each recording session took approximately 20-30 minutes depending on the performance of the participant.

3.2 Measurements

For each participant, two words (*Emma* and *Anna*) were digitized at three speaking rates (normal, slow and fast) in three repetitions (18 participants × 2 words × 3 speaking rates × 3 repetitions = 324 tokens) using the Praat software package (Boersma 2001). Measurements were made of the durations of the whole word, the first vowel, and the second vowel, as illustrated in Figure 1.

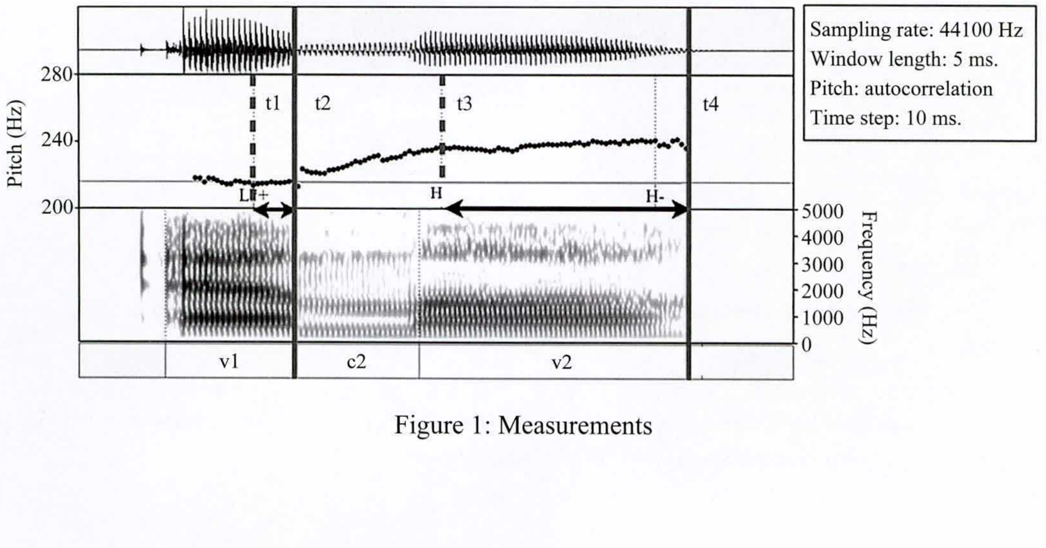

Figure 1: Measurements

The F0 was extracted for each token and the F0 minimum (corresponding to the L tone target) and F0 maximum (corresponding to the H tone target) were manually marked. The interval from each tone target to the end of the relevant tone-bearing syllable (from the L tone target to the end of the first vowel in *Emma* and *Anna*; from the H tone target to the end of the second vowel in *Emma* and *Anna*) was measured, and these values were used as indicators of the tonal alignment pattern. The shorter the intervals were, the more closely aligned the tone targets were to the end of the syllable boundary.

3.3 Results

3.3.1 Durations

Before comparing the tonal alignment patterns between the two language groups, we first needed to make sure that the durations produced by the two groups were comparable. Figure 2 shows the mean durations of the first vowel and the second vowels. The horizontal axis of Figure 2 shows the targets being measured—V1, V2, and the whole word—and the vertical axis shows the durations in milliseconds.[3] Corresponding means are given in Table 3.

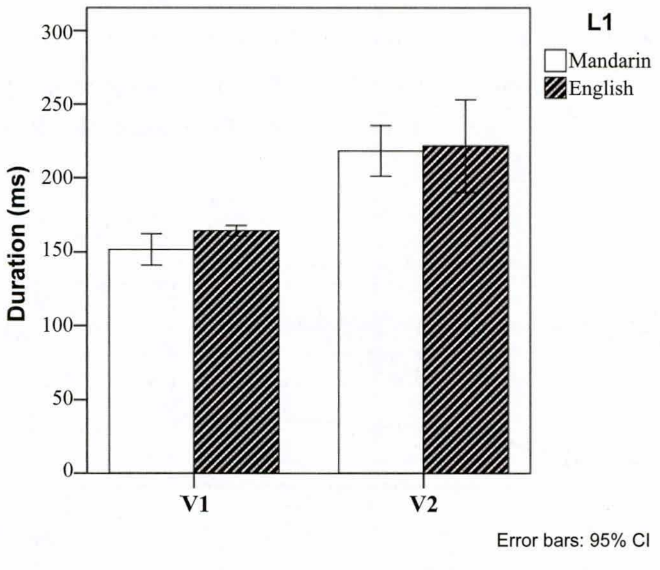

Figure 2: Mean durations in millisecond of V1 and V2 across participants

[3] The mean durations were obtained by calculating the means of V1, V2 and total durations of individual participant, and then averaged across those of all the participants.

Table 3: Mean durations in millisecond of V1, V2, and the whole word with standard deviation in parentheses

Language	V1	V2	Total
English	164.35 (21.59)	221.69 (67.77)	453.79 (89.12)
Mandarin	151.27 (21.04)	218.44 (42.48)	461.35 (62.21)

A repeated-measures analysis of variance (ANOVA) (L1 [English, Mandarin] × Duration [V1, V2]) was performed to interpret the results of the production of the experimental stimuli. The analysis showed a main effect of Duration ($F(1, 16)=46.81$, $p<.001$) but not of L1 ($F(1, 16)=1.62$, $p=0.22$). Also, no interaction was found ($F(1, 16)=0.273$, $p=0.61$). These results suggest that the durations produced by Mandarin and English native speakers were comparable. As an interesting note, we found a significant longer duration of V2 versus V1. We attribute the longer V2 duration to a phrase-final lengthening effect, given that V2 occurs at the end of the intonation phrase (e.g., Klatt 1976; Wightman et al. 1992). Although the durations were found to be comparable between the two language groups, the durations and the alignments of the tone targets (i.e. the interval between each tone target and the end of the syllable boundary) were normalized to avoid any possible influence.

3.3.2 The alignment of the tonal targets in L*+H pitch accent

We first examined the alignment pattern of the L* tone target in the first syllable of the target words *Emma* and *Anna*. Figure 3 shows the normalized intervals from the L* tone target to the end of the first syllable. The horizontal axis shows the normalized intervals (in percentages); the shorter the interval (i.e. the smaller the distance from the tone target to the end of syllable, marked as a vertical line on the right), the more closely aligned the tone target was to the syllable boundary. The vertical axis represents the different L1 groups.

We can see from Figure 3 that the alignment pattern was more variable for the English group than for the Mandarin group, and the alignment of the L* tone target of the Mandarin L1 group was closer to the end of the syllable boundary than it was for the English L1 group. One-way ANOVA showed that the difference was significant ($F(1, 17)=9.98$, $p<.01$).

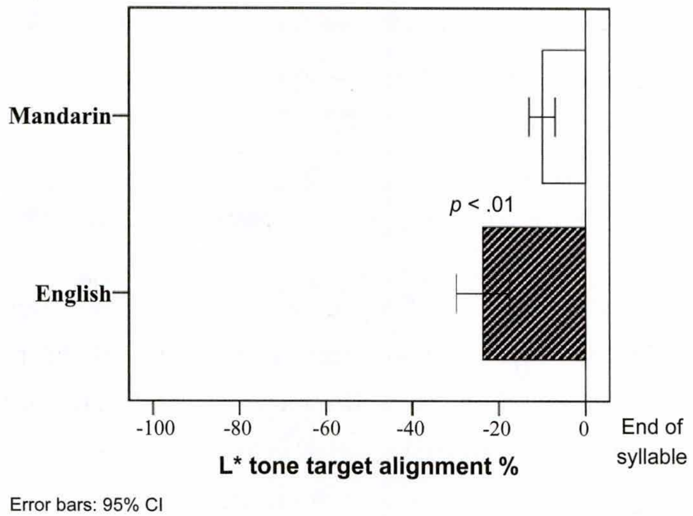

Figure 3: Alignment pattern of the L* tone target in percentage

The same pattern was observed for the H tone target associated with the second syllable, as shown in Figure 4. Again, the alignment pattern was more variable for the English L1 group than for the Mandarin L1 group, and the tone target was aligned relatively closer to the end of the syllable boundary for the Mandarin group than for the English group. Another one-way ANOVA showed that the difference was significant ($F(1, 17)=124.27$, $p<.001$).

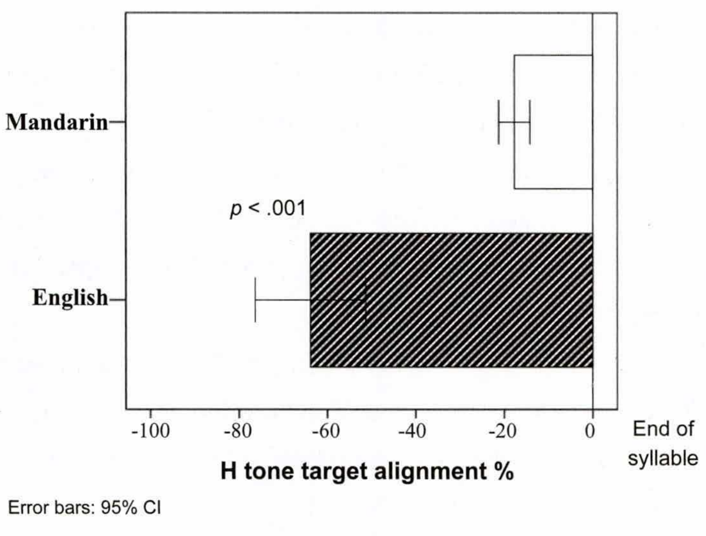

Figure 4: Alignment pattern of the H tone target in percentage

A scatterplot (with the L* tone target as a function of the H tone target) is provided in Figure 5 to summarize the results.

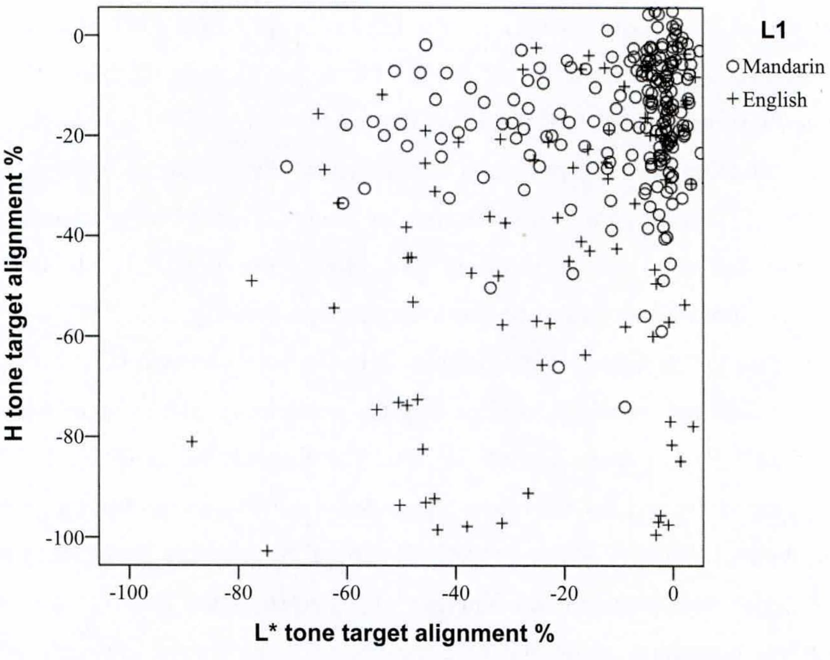

Figure 5: L* tone target as a function of the H tone target in percentage

We can see from the figure that the overall tonal alignment patterns produced by the English group (represented by crosses) displayed more variability, while those produced by the Mandarin group (represented by empty circles) clustered at the end of the tone-bearing syllable (upper-right-hand corner). In other words, significantly different tonal alignment patterns were observed in the production of English L*+H listing pitch accent by English and Mandarin native speakers: the tonal alignment pattern in English native speakers' production was more variable, whereas the pattern in Mandarin native speakers' production was consistently aligned at the boundary of the relevant tone-bearing syllable, mirroring the Mandarin L1 pattern (Xu 1998; Xu & Liu 2006).

Note that we included more Mandarin participants than English ones. One concern of the results is that the more variability of the English group versus the Mandarin group may be due to the artifact of the unbalanced sample.[4] We are not able to exclude this possible interpretation based on the current experiment. However, the findings in this study correspond to the robust findings in previous studies in which Mandarin speakers' tonal alignment is more consistent (e.g., Xu 1998; Xu & Wang 2001; Xu & Liu 2006) while that of English speakers' is more variable (e.g., Pierrehumbert & Steele 1989; Silverman & Pierrehumbert 1990; Pierrehumbert 2000; Prieto 2011). Though we cannot entirely conclude from our current results that the variability is due to different L1s, given that this is one of the first comparisons across languages on speakers' performance on L2 production and that the findings matched those found previously, we believe that our participants' production behaviors are good representatives.

There is also an important demographic factor to be taken into consideration in the analysis of the findings: would degree of English proficiency affect Mandarin speakers' tonal alignment?[5] Recall from section 3.1 that the English exposure of the Mandarin participants ranged from 1 to 14 years. Since we did not control for the participants' proficiency, we included a post-hoc variable investigating possible proficiency effect as a function of length of exposure to English. To investigate whether tonal alignment behavior differs according to degree of exposure to English, we ran another one-way ANOVA including English exposure as an independent variable (1-2 years vs. 3+ years, 6 speakers in each group). The cut-off line was set as 1-2 years vs. 3+ years according to pedagogical studies which suggest oral proficiency takes at least 3 years to develop (e.g., Hakuta et al. 2000). This follow-up test did not show an effect (L* tone alignment: $F(1, 11)=3.04$, $p=.112$; H tone alignment: $F(1, 11)=1.71$, $p=.221$). Although confirmation needs to be made in future research in which language proficiency and number of participants are better controlled for, the preliminary findings suggest that additional exposure to English did not seem to affect the alignment patterns of Mandarin speakers' L2 productions.

---

4 We thank an anonymous reviewer for the suggestion.

5 This was pointed out independently by two anonymous reviewers and the question of which proficiency level L2 learners could acquire or master L2 tonal alignment patterns is an important question to be answered in future studies.

3.3.3 Effect of speech rate

We are also interested in the extent to which contextual differences affect tonal alignment patterns in the production of L*+H listing intonation by English and Mandarin native speakers. In the literature, it has been observed that F0 alignment-to-segment in the English production of pitch accent is significantly affected by contextual differences such as speech rates (Silverman & Pierrehumbert 1990). F0 alignment in the Mandarin production of lexical tone, on the other hand, has been found to remain consistent across different speech rates (Xu 1998). If L1 transfer occurs in Mandarin speakers' production of English listing intonation, we should expect less variability across different speech rates in the Mandarin speakers' L2 tonal alignment pattern than in English speakers' production of the same pitch accent.

To compare the tonal alignment patterns between the two language groups across different speech rates, we first needed to make sure that the durations produced by the two groups varied according to speech rate and were comparable between the two groups. Figure 6 shows the total durations of the target words, *Emma* and *Anna*, produced by the two groups across different speech rates.

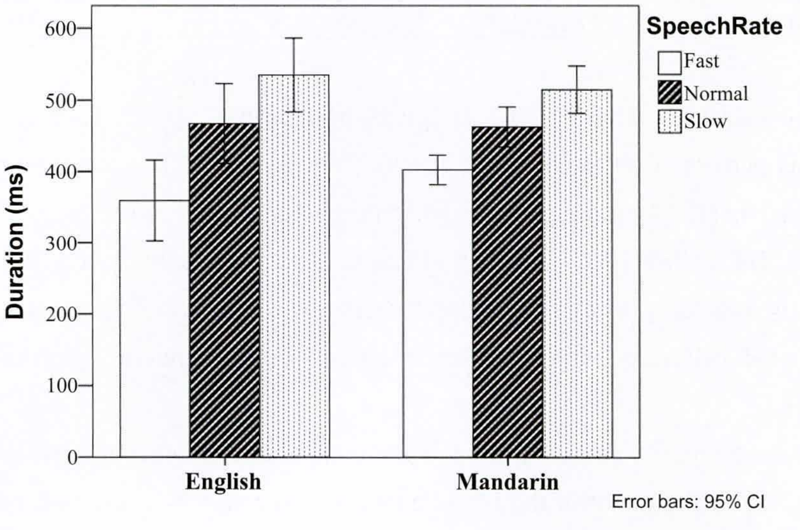

Figure 6: Mean durations in millisecond across different speech rates

A repeated-measures ANOVA (L1 [English, Mandarin] × Speech Rate [Fast, Normal, Slow]) revealed a main effect of Speech Rate ($F(2, 28)=73.257$, $p<.001$) but not of L1 ($F(1, 14)=.520$, $p=.483$). Post-hoc tests showed that all pairwise comparisons of the levels under the factor Speech Rate (i.e., Fast vs. Normal, Normal vs. Slow, Slow vs. Fast) were significantly different (all $p<.001$).[6] Overall, the results indicate differences across different speech rates, but not between the two L1 groups.

One might also wonder that the differences in duration might not be due to the different speech rate manipulation but due to a side-effect of different pitch heights (e.g., Gussenhoven & Zhou 2013; Yu et al. 2014).[7] In other words, the longer/shorter duration might be because of physical reasons (i.e., the lower the pitch, the longer the duration), and not because of different speech rates. To exclude this possibility, we extracted F0 values in Hertz at 12 points for each stimulus and calculated the pitch average from these points for each of the token to serve as dependent variable. The averaged F0 in Hz and standard deviations across speech rates are shown in Table 4.

Table 4: Mean F0 in Hz across speech rates with standard deviation in parentheses

Language	Fast	Normal	Slow
English	101.12 (8.1)	93.56 (9.26)	93.19 (9.78)
Mandarin	93.80 (8.22)	95.03 (6.57)	91.96 (9.9)

A repeated-measures ANOVA (L1 [English, Mandarin] × Speech Rate [Fast, Normal, Slow]) was performed to interpret the results from the pitch. The analysis did not show effects of L1 ($F(1, 14)=0.45$, $p=0.51$) and Speech Rate ($F(2, 28)=2.20$, $p=0.13$), nor an interaction ($F(2, 28)=1.82$, $p=0.18$). This indicates that the pitch did not differ significantly according to different speech rates and L1 group. The result indicates that the durational difference across different speech rates is not a by-product of pitch differences.

After confirming that the durations were comparable between the two groups, let us first examine the alignment of the L* tone target (associated with the first syllable of the target words, *Emma* and *Anna*), to see if the alignment pattern is affected by changes in

6 Bonferroni was used to adjust for multiple comparisons.

7 We thank an anonymous reviewer for the possible interpretation.

duration caused by the differences in speech rate. The normalized durations from the L* tone target to the end of its tone-bearing syllable, compared across different speech rates, are shown in Figure 7.

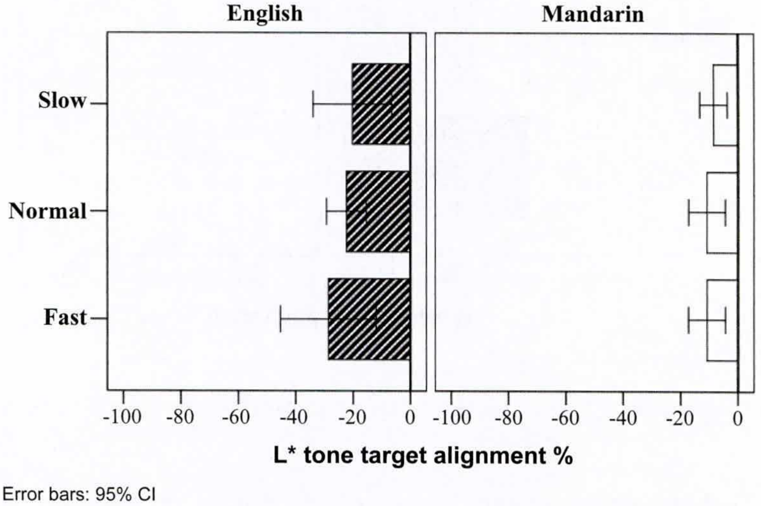

Figure 7: Alignment pattern of the L* tone target across different speech rates in percentage

Three important observations can be made based on the data in Figure 7. First, the alignment patterns produced by the English L1 group were more variable, especially in fast speech, compared to the patterns produced by the Mandarin L1 group. Second, the Mandarin speakers aligned the L* tone target more closely to the syllable boundary than the English speakers did. A repeated-measures ANOVA showed a main effect of L1 ($F(1, 14)=7.436$, $p<.05$), indicating that the alignment patterns of the two L1 groups are significantly different. Third, we can note a slight trend of earlier alignment in the English L1 group at the fast speaking rate.

Figure 8 shows the alignment patterns of the H tone target produced by the two L1 groups across different speech rates.

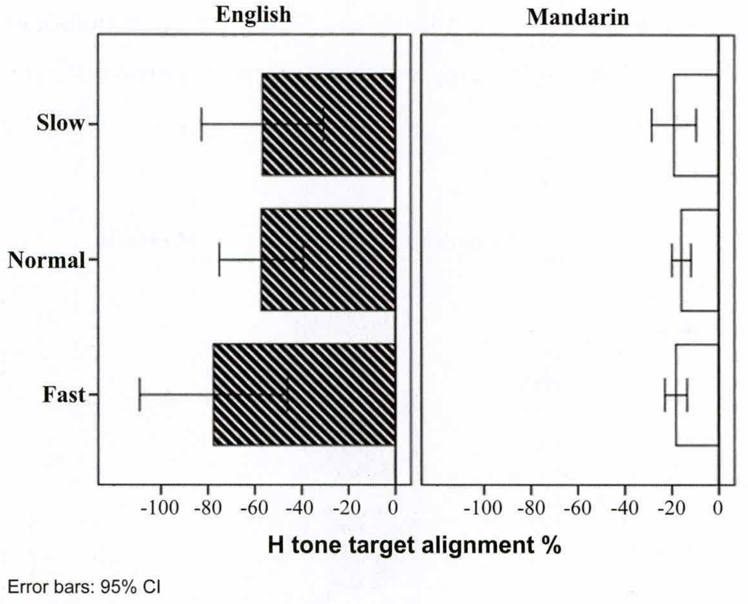

Figure 8: Alignment pattern of the H tone target across different speech rates in percentage

Figure 8 shows an even more drastic difference between Mandarin and English L1 speakers for the H tone target alignment. First, we again see that the alignment produced by the English L1 group is more variable than that produced by the Mandarin L1 group. Second, the F0 maxima produced by the Mandarin L1 group align more closely to the syllable boundary than those produced by the English L1 group. Again, a repeated-measures ANOVA showed a main effect of L1 ($F(1, 14)=187.968$, $p<.001$). Third, we can observe a progressively earlier tonal alignment in the English L1 group as the speaking rate increases. The correlation between earlier alignment and faster speaking rate found in the English L1 data supports similar findings reported in Steele (1986) and Silverman and Pierrehumbert (1990).

The final section of this paper summarizes our findings, discusses the implications of this work, and provides some possible directions for future study.

4. Discussion and Conclusion

The study reported above found that productions of English listing contours by

Mandarin and English native speakers display several significant differences: 1) In the Mandarin speakers' productions, both the L* and H tone targets align closely with the end of the respective tone-bearing syllable, and these alignment patterns remain consistent regardless of speech rate; 2) In the English speakers' productions, the alignment of the L* and H tone targets are highly variable; furthermore, we found that tonal alignment occurs progressively earlier as speaking rate increases. These findings are similar to the patterns found in previous studies on English tonal alignment (e.g., Pierrehumbert & Steele 1989; Prieto et al. 1995; Pierrehumbert 2000). Significantly, the pattern observed in the Mandarin participants' speech was similar to that observed in their native L1 pattern (Xu 1998).

The implications of this study are threefold. One possible interpretation from the findings is that prosody transfer effects may be productive even between languages with different pitch functions. Our Mandarin participants displayed tonal alignment patterns which are similar to their native *lexical tones* in their productions of English *intonation pitch contour*. This finding runs contrary to what has been reviewed in section 2.1, in which universal tendencies seem to play a greater role when L1 and L2 do not share the same suprasegmental function.

There is, however, an alternative interpretation of the results: the consistent alignment pattern produced by the Mandarin speakers might not be due to L1 transfer, but due to universal tendencies in which consistent alignment to the end of a syllable is the unmarked strategy.[8] If this interpretation is on the right track, what we found here is not a case of L1 transfer, but again a case of universal markedness, consistent with what has been reviewed in section 2.1. Unless a control group whose alignment pattern differs from that of Mandarin is included that also displays native pattern when producing English listing contour, the two interpretations cannot be teased apart.

However, there are studies looking at other aspects of prosody and found that 'foreign accent' can be detected by native speakers with L2 speech altered only in

[8] An anonymous reviewer pointed out whether the tonal alignment pattern found in this study is due to a prosodic transfer pattern or possibly attributed to a universal prosodic behavior among English learners. For example, we may conjecture that there could be an inter-language stage where L2 learners can refer to when they could not master the prosody of a target language. This universality hypothesis is an interesting question to be answered in future studies.

intonation. For example, a number of studies have examined differences in intonation between learners and native speakers of English (Russian learners in Thompson 1991; Chinese learners in Wennerstrom 1994, among others; Saudi learners in Binghadeer 2008). One common characteristic found as a foreign accent is the smaller pitch range in the production of English intonation by second language learners. O. Kang (2010) found that pitch range turns out to be the best predictor based on which English native speakers judge *accentedness* of learners' speech. That is, English native speakers rated more accented when the production is made with less pitch range.

Though these previous studies are possible support for prosody being subject to L1 transfer, nothing can be made conclusive if tonal alignment is subject to L1 transfer as well. To date, research on the suprasegmental aspects of second language acquisition has been limited, and tonal alignment has received even less attention. Prosody transfer phenomena deserve more attention if we wish to understand how L1 prosody plays a role in second language learning. One direction for future research is, thus, to examine tonal alignment patterns in other language contact scenarios where L1 and L2 do not share the same pitch function.

Second, our findings suggest that variability effects caused by phonetic factors, such as speech rate, may be subject to L1 transfer as well. In Mandarin native productions, different speech rates do not affect tonal alignment (Xu 1998). Our study has showed that the same pattern is retained in Mandarin speakers' L2 production of English. English native speakers' productions, on the other hand, are consistently affected by speaking rate. The speech rate effect was the only phonetic factor examined in the present study; many other phonetic factors are also responsible for variability in suprasegmental information. Another direction for future research will be to include other phonetic factors such as segment duration, segmental make-ups, and different prosodic boundaries, to see if the effect still holds.

Third, this study investigated only the production of tonal alignment. The identification of a clear difference between English native productions and Mandarin non-native productions points to yet another direction for research: do discrepancies in the alignment of tone targets affect native speakers' perception of non-native productions? For example, do misaligned tone targets lead to foreign accent perception, just as misproduction of segmental information does? Perceptual studies on this topic can help us better understand the nature of tonal alignment in L2 productions and identify a possible relationship between tonal alignment acquisition and language proficiency. This direction of research also has implications for second language pedagogy. For example, if discrepancies in the alignment of tone targets *do* affect native speakers' perception of non-native productions, suggesting that the slight, non-categorical differences in alignment are detectable in perception, we can enhance L2 learners' phonological awareness by providing them perceptual training (Kraljic & Samuel 2005; Schmidt 2010). Most of the pedagogical attention has been put on segmental differences and the trainings on perception and production follow accordingly. If laboratory findings provide support of prosody transfer and of improvement from raising learners' awareness, pedagogical training can follow.

(Proofreader: Wang Si-qi, Kong Ling-an)

References

Binghadeer, Nora. "An Acoustic Analysis of Pitch Range in the Production of Native and Nonnative Speakers of English," *The Asian EFL Journal*, 10.4, 2008, pp. 96-113.

Boersma, Paul. "PRAAT, a System for Doing Phonetics by Computer," *Glot International*, 5.9/10, 2001, pp. 341-345.

______. "Cue Constraints and Their Interactions in Phonological Perception and Production," in Paul Boersma & Silke Hamann (eds.), *Phonology in Perception*. Berlin: Mouton de Gruyter, 2009, pp. 55-110.

Brannen, Kathleen. "The Role of Perception in Differential Substitution," *Canadian Journal of Linguistics*, 47.1, 2002, pp. 1-46.

Broselow, Ellen. "Unmarked Structures and Emergent Rankings in Second Language Phonology," *International Journal of Bilingualism*, 8.1, 2004, pp. 51-65. doi: 10.1177/13670069040080010401

______. "Stress Adaptation in Loanword Phonology: Perception and Learnability," in Paul Boersma & Silke Hamann (eds.), *Phonology in Perception*. Berlin: Mouton de Gruyter, 2009, pp. 191-234.

Broselow, Ellen, Chen Su-i & Wang Chilin. "The Emergence of the Unmarked in Second Language Phonology," *Studies in Second Language Acquisition*, 20.2, 1998, pp. 261-280. doi: 10.1017/S0272263198002071

Broselow, Ellen & Dan Finer. "Parameter Setting in Second Language Phonology and Syntax," *Second Language Research*, 7.1, 1991, pp. 35-59. doi: 10.1177/026765839100700102

Brown, Cynthia. "Acquisition of Segmental Structure: Consequences for Speech Perception and Second Language Acquisition," Ph.D. Dissertation, Montreal: McGill University, 1997.

Calabrese, Andrea & W. Leo Wetzels. "Loan Ponology: Issues and Controversies," in Andrea Calabrese & W. Leo Wetzels (eds.), *Loan Phonology*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2009, pp. 1-11.

Charles-Luce, Jan. "Word-final Devoicing in German and the Effects of Phonetic and Sentential Contexts," *Journal of the Acoustical Society of America*, 77.S1, 1985, p. S85. doi: 10.1121/1.2022551

D'Imperio, Mariapaola. "Tonal Structure and Pitch Targets in Italian Focus Constitutents," in John Ohala (ed.), *Proceedings of the 14th International Congress of Phonetic Sciences*. San Franscisco: University of California, 1999, pp. 1757-1760.

Demuth, Katherine. "Markedness and the Development of Prosodic Structure," in Jill N. Beckman (ed.), *Proceedings of the North Eastern Linguistic Society 25*. Amherst, MA: GLSA, University of Massachussets, 1995, pp. 13-25.

______. "The Prosodic Structure of Early Words," in James L. Morgan & Katherine Demuth (eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates, 1996, pp. 171-184.

Dilley, Laura. "The Role of F0 Alignment in Distinguishing Categories in American English Intonation," Manuscript, Bowling Green: Bowling Green State University, 2007.

Duanmu San. *The Phonology of Standard Chinese*. New York: Oxford University Press, 2007.

Dupoux, Emmanuel, Kakehi Kazuhiko, Hirose Yuki, Christophe Pallier & Jacques Mehler. "Epenthetic Vowels in Japanese: A Perceptual Illusion?," *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1999, pp. 1568-1578. doi: 10.1037/0096-1523.25.6.1568

Eckman, Fred R. "Markedness and the Contrastive Analysis Hypothesis," *Language Learning*, 27.2, 1977, pp. 315-330. doi: 10.1111/j.1467-1770.1977.tb00124.x

Eckman, Fred R. & Gregory K. Iverson. "The Role of Native Language Phonology in the Production of L2 Contrasts," *Studies in Second Language Acquisition*, 35.1, 2013, pp. 67-92. doi: 10.1017/S027226311200068X

Escudero, Paola & Paul Boersma. "Bridging the Gap Between L2 Speech Perception Research and Phonological Theory," *Studies in Second Language Acquisition*, 26.4, 2004, pp. 551-585. doi: 10.1017/S0272263104040021

Fenyvesi, Anna & Gyula Zsigri. "The Role of Perception in Loanword Adaptation: The Fate of Initial Unstressed Syllables in American Finnish and American Hungarian," *SKY Journal of Linguistics*, 19, 2006, pp. 131-146.

Gass, Susan M. "Second Language Acquisition and Linguistic Theory: The Role of Language Transfer," in Suzanne Flynn & Wayne O'Neil (eds.), *Linguistic Theory in Second Language Acquisition*. Dordrecht/Boston/London: Kluwer Academic Publisher, 1996, pp. 384-403.

Gussenhoven, Carlos & Zhou Wenchui. "Revisiting Pitch Slope and Height Effects on Perceived Duration," Paper presented at the 'INTERSPEECH 2013: 14th Annual

Conference of the International Speech Communication Association', 25-29 August 2013, Lyon, France: INRIA Rennes Bretagne Atlantique and Europa Organisation.

Hakuta Kenji, Yuko Goto Butler & Daria Witt. *How Long Does It Take English Learners to Attain Proficiency?* Stanford, CA: The University of California Linguistic Minority Research Institute, 2000.

Hancin-Bhatt, Barbara Jean. "Phonological Transfer in Second Language Perception and Production," Ph.D. Dissertation, Urbana: University of Illinois Urbana-Champaign, 1994.

Hiroya Fujisaki. "Dynamic Characteristics of Voice Fundamental Frequency in Speech and Singing," in Peter F. MacNeilage (ed.), *The Production of Speech*. New York: Springer, 1983, pp. 39-55. doi: 10.1007/978-1-4613-8202-7_3

Hsieh Feng-fan. "High Infidelity: The Non-mapping of Japanese Accent onto Taiwanese Tone," *MIT Working Papers in Linguistics*, 52, 2006, pp. 1-27.

Hsieh Feng-fan & Michael J. Kenstowicz. "Phonetic Knowledge in Tonal Adaptation: Mandarn and English Loanwords in Lhasa Tibetan," *Journal of East Asian Linguistics*, 17, 2008, pp. 279-297. doi: 10.1007/s10831-008-9027-7

Kang Okim. "Relative Salience of Suprasegmental Features on Judgments of L2 Comprehensibility and Accentedness," *System*, 38.2, 2010, pp. 301-315. doi: 10. 1016/j.system.2010.01.005

Kang Yoonjung. "Perceptual Similarity in Loanword Adaptation: English Postvocalic Word-final Stops in Korean," *Phonology*, 20, 2003, pp. 219-273.

______. "Tutorial Overview: Suprasegmental Adaptation in Loanwords," *Lingua*, 120.9, 2010, pp. 2295-2310. doi: 10.1016/j.lingua.2010.02.015

______. "Loanword Phonology," in Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology vol. 4: Phonological Interfaces*. West Sussex, United Kingdom: Blackwell Publishing Ltd., 2011, pp. 2258-2282.

Kawahara, Shigeto & Kelly Garvey. "Testing the P-map Hypothesis: Coda Devoicing," Manuscript, New Brunswick, NJ: Rutgers University, 2010.

Kenstowicz, Michael & Atiwong Suchato. "Issues in Loanword Adaptations: A Case Study from Thai," *Lingua*, 116.7, 2006, pp. 921-949. doi: 10.1016/j.lingua.2005.05.006

Kim Miran. "Multiple Cue Interpretation in Non-native Phoneme Perception: Korean Listeners' Perception of English /s/ in Varied Prosodic Contexts," *Journal of the Acoustical Society of America*, 120, 2006, pp. 3172-3173. doi: 10.1121/1.4787934

Klatt, Dennis H. "Linguistic Uses of Segmental Duration in English: Acoustic and Perceptual Evidence," *Journal of the Acoustical Society of America*, 59.5, 1976, pp. 1208-1221. doi: 10.1121/1.380986

Kohler, Klaus. "Categorical Pitch Perception," in Thomas V. Gamkrelidze (ed.), *Proceedings of the 11th International Congress of Phonetic Sciences*. Tallinn: Academy of Sciences of the Estonian SSR, 1987, pp. 331-333.

Kraljic, Tanya & Arthur G. Samuel. "Perceptual Learning for Speech: Is There a Return to Normal?," *Cognitive Psychology*, 51.2, 2005, pp. 141-178. doi: 10.1016/j.cogpsych.2005.05.001

LaCharité, Darlene & Carole Paradis. "Category Preservation and Proximity versus Phonetic Approximation in Loanword Adaptation," *Linguistic Inquiry*, 36.2, 2005, pp. 223-258. doi: 10.1162/0024389053710666

Lu Yu-an. "Adaptation of English Interdental Fricatives by Speakers of Taiwan Mandarin," in Linguistic Society of Korea (ed.), *Current Issues in Unity and Diversity of Languages*. Seoul: LSK, 2009, pp. 1151-1165.

Maddieson, Ian. "Tone Loans: A Question Concerning Tone Spacing and a Method of Answering It," *UCLA Working Papers in Phonetics*, 36, 1977, pp. 49-83.

Miao Ruiqin. "Loanword Adaptation in Mandarin Chinese: Perceptual, Phonological and Sociolinguistic Factors," Ph.D. Dissertation, Stony Brook: Stony Brook University, 2005.

Ou Shu-chen. "Taiwanese EFL Learner's Perception of English Word Stress," *Concentric: Studies in Linguistics*, 36.1, 2010, pp. 1-23.

Paradis, Carole & Antoine Tremblay. "Nondistinctive Features in Loanwords Adaptation: The Unimportance of English Aspiration in Mandarin Chinese Phoneme Categorization," in Andrea Calabrese & W. Leo Wetzels (eds.), *Loan Phonology*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2009, pp. 211-224. doi: 10.1075/cilt.307.09par

Pierrehumbert, Janet B. *The Phonology and Phonetics of English Intonation*. Cambridge: Massachusetts Institute of Technology, 1980.

______. "Tonal Elements and Their Alignment," in Merle Horne (ed.), *Prosody: Theory and Experiment*. Berlin: Kluwer Academic Publishers, 2000, pp. 11-36.

Pierrehumbert, Janet B. & Shirley A. Steele. "Categories of Tonal Alignment in English," *Phonetica*, 46, 1989, pp. 181-196. doi: 10.1159/000261842

Port, Robert & Michael O'Dell. "Neutralization of Syllable-final Voicing in German," *Journal of Phonetics*, 13.4, 1985, pp. 455-471.

Prieto, Pilar. "Tonal Alignment," in Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology*. West Sussex, United Kingdom: Blackwell Publishing Ltd., 2011, pp. 1185-1203.

Prieto, Pilar, Mariapaola D'Imperio & Barbara Gili Fivela. "Pitch Accent Alignment in Romance: Primary and Secondary Associations with Metrical Structure," *Langauge and Speech*, 48.4, 1995, pp. 359-396. doi: 10.1177/00238309050480040301

Ramijsen, Bert. "Tonal Alignment is Contrastive in Falling Contours in Dinka," *Language*, 89.2, 2013, pp. 297-327. doi: 10.1353/lan.2013.0023

Schmidt, Richard. "Attention, Awareness, and Individual Differences in Language Learning," Paper presented at 'CLaSIC', 2-4 December 2010, Singapore: National University of Singapore.

Schwartz, Bonnie D. & Rex A. Sprouse. "Word Order and Nominative Case in Non-native Language Acquisition: A Longitudinal Study of (L1 Turkish) German Inter-langauge," in Teun Hoekstra & Bonnie D. Schwartz (eds.), *Language Acquisition Studies in Generative Grammar*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1994, pp. 317-368. doi: 10.1075/lald.8.14sch

Shinohara Shigeko. "Emergence of Universal Grammar in Foreign Word Adaptations," in René Kager, Joe Pater & Wim Zonneveld (eds.), *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press, 2004, pp. 292-320.

Silverman, Kim E. A. & Janet B. Pierrehumbert. "The Timing of Prenuclear High Accents in English," in John Kingston & Mary E. Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press, 1990, pp. 72-106.

Steele, Shirley A. "Nuclear Accent F0 Peak Location: Effects of Rate, Vowel, and Number of Following Syllables," *Journal of the Acoustical Society of America*, 80.S1, 1986, p. S51. doi: 10.1121/1.2023842

Tarone, Elaine E. "Some Influences on the Syllable Structure of Interlangauge Phonology," *International Review of Applied Linguistics in Langauge Teaching*, 18, 1987, pp. 139-153.

Thompson, Irene. "Foreign Accents Revisited: The English Pronunciation of Russian Immigrants," *Language and Learning*, 41.2, 1991, pp. 177-204. doi: 10.1111/j.1467-1770.1991.tb00683.x

Weinberger, Steven H. "Minimal Segments in Second Language Phonology," in Allan James & Jonathan Leather (eds.), *Second Language Speech: Structure and Process*. Berlin: Mouton de Gruyter, 1996, pp. 263-312. doi: 10.1515/9783110882933.263

Wennerstrom, Ann. "Intonaltional Meaning in English Discourse: A Study of Non-native Speakers," *Applied Linguistics*, 15.4, 1994, pp. 399-420. doi: 10.1093/applin/15.4.399

Wightman, Colin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf & Patti J. Price. "Segmental Durations in the Vicinity of Prosodic Phrase Boundaries," *Journal of the Acoustical Society of America*, 91.3, 1992, pp. 1707-1717. doi: 10.1121/1.402450

Wu Hsiao-hung. "Stress to Tone: A Study of Tone Loans in Mandarin Chinese," *MIT Working Papers in Linguistics*, 52, 2006, pp. 227-253.

Xu Yi. "Consistency of Tone-syllable Alignment Across Different Syllable Structures and Speaking Rates," *Phonetica*, 55, 1998, pp. 179-203. doi: 10.1159/000028432

Xu Yi & Liu Fang. "Tonal Alignment, Syllable Structure and Coarticulation: Toward an Integrated Model," *Italian Journal of Linguistics*, 18.1, 2006, pp. 125-159.

Xu Yi & Q. Emily Wang. "Pitch Targets and Their Realization: Evidence from Mandarin Chinese," *Speech Communication*, 33, 2001, pp. 319-337. doi: 10.1016/S0167-6393(00)00063-7

Yip, Moira. "Perceptual Influence in Cantonese Loanword Phonology," *The Journal of the Phonetical Society of Japan*, 6.1, 2002, pp. 4-21.

Yu, Alan C. L., Lee Hyunjung & Jackson Lee. "Variability in Perceived Duration: Pitch Dynamics and Vowel Quality," Paper presented at the '4th International Symposium on Tonal Aspect of Langauges (TAL-2014)', 13-16 May 2014, Nijmegan, The Netherlands: Radboud University.

Zhang Jie. "The Role of Contrast-specific and Language-specific Phonetics in Contour Tone Distribution," in Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically Based Phonology*. Cambridge: Cambridge University Press, 2004, pp. 157-190. doi: 10.1017/CBO9780511486401.006

第二語言學習中之母語韻律轉移現象
——以中文使用者產出英文語調為例

盧郁安
國立交通大學外國語文學系
yuanlu@nctu.edu.tw

金美蘭
國立慶尚大學英語教育系
mirankim@gnu.ac.kr

摘 要

本文探討母語韻律中之音調對整對第二語言韻律產出之影響，研究對象為中文使用者如何產出英文中 L*H 之語調。中文為聲調語言，不同音調可區別字義之不同，而英文為語調語言，不同音調可區別不同言談及語用之功能。前人研究顯示，中文裡聲調中基頻的對整以音節為單位，且一致地對整於音節尾，而英文的語調是以韻律片語為單位，且對整常因語速或音節長度等語音現象而產生不同。基於此二語言韻律及音調使用上之不同，本文探討母語轉移現象是否會發生。研究結果顯示，中文使用者在產出英文中 L*H 語調時，使用與母語一致的音調對整方式，且其對整不受語速之影響。相對的，英文使用者產出 L*H 語調時，對整上較為彈性，且受語速快慢而有對整前後之不同。

關鍵詞：第二語言學習，音調對整，母語轉移現象

（收稿日期：2016. 2. 17；修正稿日期：2016. 5. 12；通過刊登日期：2016. 6. 7）