INTERLANGUAGE PROSODY: NATIVE ENGLISH SPEAKERS' PRODUCTON OF MANDARIN *YES-NO* QUESTIONS*

Carolyn Pytlyk
University of Victoria

1. Introduction

Research in second language acquisition (SLA) agrees that a learner's native language (L1) plays a pivotal role in the perception and production of a second language (L2). In fact, Major (2001) states, "our NL [native language] has an inescapable influence on our L2" (p. 36). With regards to L2 phonology, most of the research to date has focused on the L2 phonological development at the segmental level. In contrast, research into the L2 phonological development at the prosodic level remains relatively sparse. However, recent research has recognized the need to investigate L2 prosodic development and its contributions to nonnative accents (Ueyama & Jun 1998, Vicseglia & Fodor 2006). For example, Visceglia and Fodor (2006) maintain that prosody "makes a significant contribution to the perception of a non-native accent in L2 speech" (p. 27). This research endeavours to contribute to the literature on L2 prosodic development by focussing on the L2 prosodic patterns of ***yes-no*** questions. This paper will address two important questions. First, do L1 prosodic contours influence the production of L2 prosodic contours? More specifically, do native English (NE) speakers produce Mandarin ***yes-no*** questions with a rising pitch contour as found in canonical English ***yes-no*** questions?

2. Background

2.1 Previous Literature

In the search to characterize the prosodic influence of a first language on the prosodic structure of a second language, a few studies have attempted to address the relationship between L1 and L2 prosodic patterns. While "not all L1 features directly shape L2 intonation" (Ueyama & Jun 1998, p. 644), prosodic research has found that some native-nonnative prosodic differences can be attributed to L1 interference. In particular, some prosodic interference from the L1 has been observed for Japanese and Korean learners of English (Ueyama & Jun 1998), Spanish learners of English (García Lecumberri 2001, Ramírez Verdugo 2006), Dutch learners of Greek (Mennen 2004), and French learners of Dutch and Dutch learners of French (Raiser & Hiligsmann 2007).

* This research was supported in part by the Province of British Columbia through the Ministry of Advanced Education. I gratefully acknowledge Tae-Jin Yoon and Qian Wang for their contributions.

Actes du congrès annuel de l'Association canadienne de linguistique 2008.
Proceedings of the 2008 annual conference of the Canadian Linguistic Association.

In a recent study, Visceglia and Fodor (2006) investigated the phrase- (i.e. local prosody) and sentence- (i.e. global prosody) level prosodic patterns of native English speakers learning Mandarin and native Mandarin speakers learning English. They discovered that, whereas the global prosodic contours of both English and Mandarin questions (as spoken by native speakers) are characterized by higher overall pitch than declaratives, both groups of learners failed to produce F0 distinctions between interrogatives and declaratives. In other words, L2 speakers did not positively transfer the L1 pitch distinction between declaratives and interrogatives and apply it to the L2 prosodic structure. This finding is in line with Ueyama and Jun's (1998) conclusion that not all prosodic structures are positively transferred from the L1 into the L2. Visceglia and Fodor (2006) suggested this may be attributed to attentional focus where learners may have been focused on local prosody, and therefore, paid no attention to pitch height of declaratives and questions. For example, the authors maintained that the English L2 Mandarin learners focused on accurate pronunciation of lexical tone and did not have extra attentional resources to distinguish pitch excursions. While the global prosody results indicate a lack of L1 transfer, in terms of local prosody, Visceglia and Fodor did find some evidence of negative transfer. That is, native Mandarin speakers tended to compress pitch excursions in English declaratives and interrogatives to the final syllable. Native English speakers, on the other hand, tended to use a final rise on the final syllable in Mandarin *ma* particle questions. However, as Visceglia and Fodor did not statistically evaluate their data, their results are not definitive.

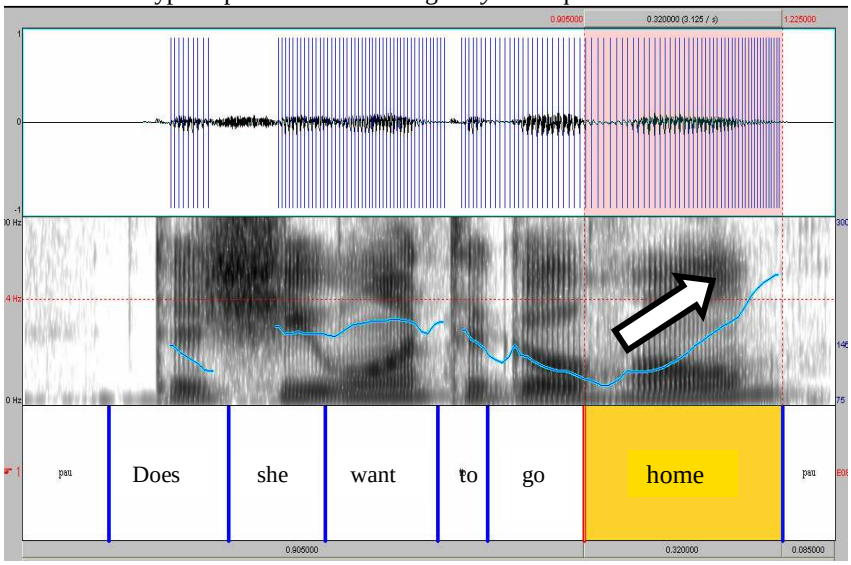
2.2 Properties of English and Mandarin prosodic structure

Since this research seeks to investigate native English speakers' productions of Mandarin *yes-no* questions in terms of prosodic structure, it is important to briefly compare the prosodic characteristics of each language.

2.2.1 English

In English, boundary tones mark utterance-level illocution. That is, speakers use different sentence-final boundary tones for different types of utterances. For example, a pitch fall from a high pitch accent on the last prominent word to an utterance-final low boundary tone is indicative of declaratives in English. Conversely, a pitch rise from a low level pitch accent to a high boundary tone is indicative of interrogatives. In other words, typically, declaratives fall to low final boundary tones while interrogatives rise to high final boundary tones (Beckman & Pierrehumbert 1986, Pierrehumbert & Hirschberg 1990, Visceglia & Fodor 2006). For illustrative purposes, Figure 1 visually demonstrates the typical final rise contour of an English *yes-no* question.

FIGURE 1 Typical pitch contour of English *yes-no* questions



2.2.2 Mandarin

Unlike English, Mandarin employs pitch lexically such that each syllable is associated with a tone which distinguishes meaning. There are four lexical tones in Mandarin. They are as follows:

- Tone 1 is a high level tone (ex. mā ‘mother’).
- Tone 2 starts at a mid-level and then rises (ex. má ‘hemp’).
- Tone 3 falls low then rises (ex. mǎ ‘horse’).
- Tone 4 starts high then falls low (ex. mà ‘to scold’).

Because pitch contours distinguish lexical meaning, Mandarin prosody must preserve lexical tone pitch variations thereby preventing illocutionary rising and/or falling final boundary contours (Visceglia & Fodor 2006). This means that the pitch contour of Mandarin *yes-no* questions ending in: a) Tone 1 will remain level, b) Tone 2 will rise, c) Tone 3 will fall then rise, and d) Tone 4 will fall. That is not to say, however, that Mandarin does not use F0 to signal utterance-level illocution. Rather, instead of employing contour changes, declaratives and questions are distinguished by the overall pitch range. In general, questions have a higher pitch range than declaratives (Shen 1990).

It is also important to note that there are two ways to construct Mandarin *yes-no* questions (Lin 2001). The first way is constructed by adding an interrogative *ma* particle to the end of a declarative sentence. For example:

(1) Tā huì chàng gē **ma**?
he/she can sing song MA-part.
'Can he/she sing a song?'

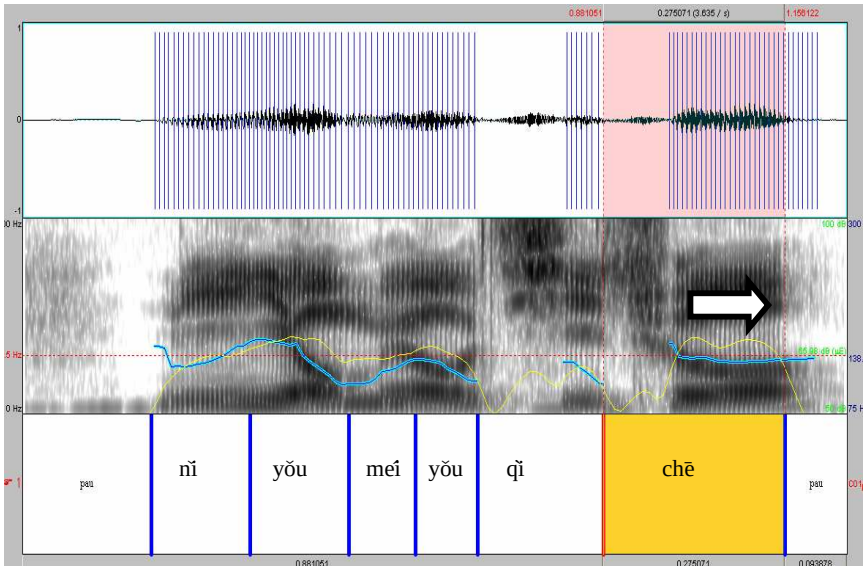
The second way to construct Mandarin *yes-no* questions is to reduplicate the verb and insert a negative morpheme bù or meí between the two verbs. For example:

(2) Tā **huì bú huì** chàng gē ?
he/she can-not-can sing song
'Can he/she sing a song?'

While the constructions are different, there are no differences in meaning between the two types of questions.

Figures 2 and 3 below demonstrate the pitch contours of both constructions. In Figure 2, the V bù V question ends with Tone 1.

FIGURE 2 Typical pitch contour of a Mandarin V bù V question ending in Tone 1[1]

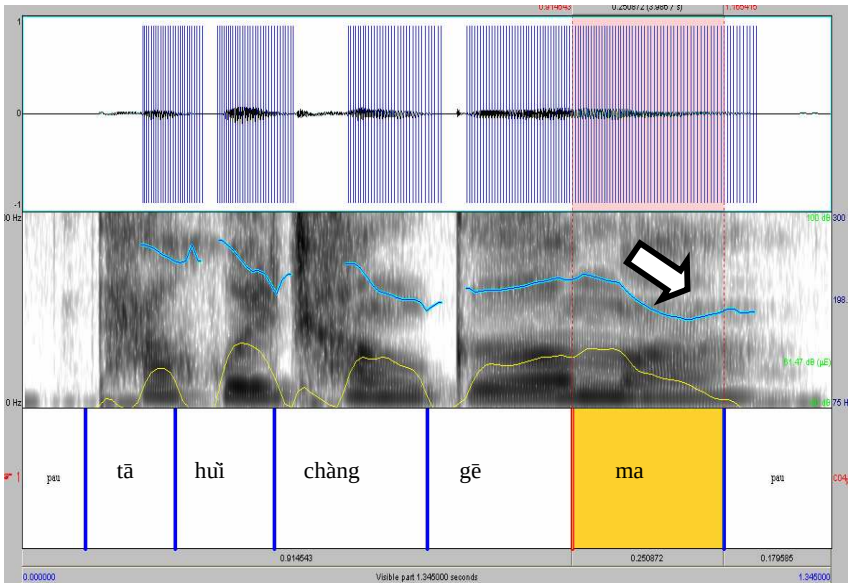


In Figure 3, the *ma* particle question also ends in a level tone. However, *ma* is not associated with any of the four lexical tones; rather *ma* is associated with the Mandarin neutral tone. According to Lin (2001), the neutral tone is underlyingly low and derives its pitch from the preceding tone such that the

[1] Translation: *Do you have a car?*

pitch falls on a neutral tone when preceded by either a Tone 1, Tone 2, or Tone 4 and the pitch rises on a neutral tone with when preceded by a Tone 3.

FIGURE 3 Typical pitch contour of *ma* particle question when *ma* follows a Tone 1[2]



2.3 Research Hypotheses

As mentioned above in Section 1, the purpose of this research is to determine whether native language prosodic contours negatively transfer into the interlanguage prosodic structure. Specifically, do native English speakers use a final rising contour that is characteristic of English *yes-no* questions when they produce Mandarin *yes-no* questions? If native prosodic structures are negatively transferred into the interlanguage grammar, two hypotheses arise with respect to native English speakers production of Mandarin *yes-no* questions.

1. Native English (NE) speakers' prosodic contours for Mandarin *yes-no* questions will resemble the contours produced by NE for canonical English *yes-no* questions.
2. NE speakers prosodic contours for Mandarin *yes-no* questions will not resemble the contours produced by native Mandarin (NM) speakers for canonical Mandarin *yes-no* questions.

[2] Translation: *Can s/he sing a song?*

3. Methodology

3.1 Participants

In order to test the above research hypotheses, two groups of participants were recruited at the University of Victoria. The first group was comprised of native English (NE) speakers who were intermediate learners of Mandarin. There were five participants in this group (four males and one female). The second group of participants were native Mainland Chinese Mandarin (NM) speakers who were in the ESL program at the university. There were also five participants in this group (three males and two females). Each participant received $10 for participating in this research project. All participants reported to have no known auditory or visual impairments.

3.2 Experimental Materials

To compare NE speakers' prosodic patterns of English and Mandarin *yes-no* questions, three sets of experimental materials were designed to elicit three types of *yes-no* questions. The first type of questions used the Mandarin 'V bù V' construction as in "Tā huì bú huì chàng gē?"[3] The second used the Mandarin *ma* question particle construction as in "Tā huì chàng gē ma?" It is important to note here that the tone of the final lexical word in each Mandarin question was controlled – all final words were Tone 1 (the high level tone). Finally, the third type of questions was the English translation of each Mandarin question as in "Can he sing a song?" The purpose of recording the English questions was to confirm that the NE speakers do, in fact, use a rising intonation on English *yes-no* questions and to gather a baseline against which to measure the NE production of Mandarin questions.

In order to elicit multiple experimental tokens, five different questions were constructed to reflect each of the three question types under investigation. In other words, each of the five questions was asked in: 1) an English construction (EN), 2) a Mandarin V bù V construction (BU), and 3) a Mandarin *ma* particle construction (MA). As a result, there were 15 experimental questions (5*3). The following table provides all the experimental questions.

[3] Both the V bù V and *ma* particle example questions translate as: *Can s/he sing a song?*

TABLE 1 The five questions for each of the three question constructions

ENGLISH	MANDARIN	
	V bú V	*ma* particle
1. Are you a student?	Nǐ shì bú shì xué shēng?	Nǐ shì xué shēng ma?
2. Does she want to go home?	Tā yaò bú yaò weí jiā?	Tā yaò weí jiā ma?
3. Do you have a car?	Nǐ yoǔ meí yoǔ qìchē?	Nǐ yoǔ qìchē ma?
4. Does he want to read books?	Nǐ xǐ bù xǐhuān kàn shū?	Nǐ xǐhuān kàn shu mā?
5. Can he sing a song?	Tā huì bú huì chàng gē?	Tā huì chàng gē ma?

3.3 Procedure

The experiment was divided into two blocks. In the first block, the questions were organized so that the EN questions were first, the Mandarin BU questions were second and the Mandarin MA questions were third.[4] This order was recorded three times. After the first block, the participants took a short 5 minute break. In the second block, the question sets were organized so that the EN questions were first, the Mandarin MA questions were second and the Mandarin BU questions were third. This order was recorded two times. Therefore, each participant produced a total of 75 question tokens (15*3*2). This, in turn, yielded a total of 750 tokens (10 participants*75 question tokens).

The participants were recorded in a sound treated booth in the Department of Linguistics Phonetics Lab at the University of Victoria. Each question and answer pair was presented to the participants via a computer screen (using Microsoft Office PowerPoint 2003).[5] The questions were recorded using the software Audacity (Version 1.2.4) at a sampling frequency of 44.1 kHz with an M-Audio Luna Large-Diaphragm Condenser microphone on a Windows XP workstation connected to M-Audio Firewire 4/10 Preamp Interface.

Once all the question and answer pairs were recorded and saved in a single *.wav* file for each block for each participant, the data were broken down and analysed in seven procedural steps.

1. Each target question was isolated using the software program PRAAT (Version 4.4.13).
2. A PRAAT script was used to identify and save each question token as an individual *.wav* file.
3. The consonants and vowels were segmented in the final word in each question.
4. Another PRAAT script was applied to calculate the mean F0 and the F0 standard deviation for each question token.

[4] The participants also recorded answers for each question; however, the answers were not analysed in this project.

[5] The Mandarin questions were presented in the Pinyin orthography rather than Chinese characters.

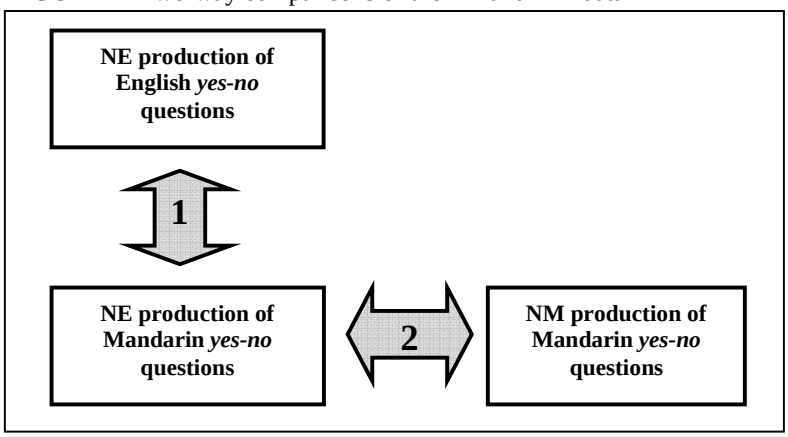
5. The beginning and end F0 values for each final vowel were gathered to calculate the raw F0 differences.
6. The normalized F0 differences for each final vowel using the final vowel beginning F0 and the end F0 as well as the overall mean F0 and F0 standard deviation for each question token was calculated (see Figure 5).
7. Statistical analyses on the raw F0 and normalized F0 differences were conducted using the software SPSS.

The motivations for investigating the final vowel as well as the calculations of both the raw and normalized F0 differences are discussed in the following section.

3.4 Data analysis

In order to test the two research hypotheses articulated in Section 2, the data collected for this research were analysed and evaluated in two ways. First, the contours were compared in terms of NE production of English and Mandarin questions. Second, the contours of NE productions of Mandarin questions were compared with the contours of native Mandarin speakers' productions of the same questions. For clarity, Figure 4 below provides a visual representation of the two comparisons where the numbered arrows correspond with the hypotheses outlined in Section 2.3.

FIGURE 4 Two-way comparisons of the NE and NM data



For this research, the domain of investigation was the final vowel in each target question. As consonants, especially voiceless consonants, can have unrepresentative (or non-existent) pitch contours, the goal here was to gather data from uninterrupted pitch movements (i.e. the vowels). It was assumed that if NE speakers rise from the final pitch accent to the final boundary tone, a rising pitch contour would still be observed on the final vowel even if the final

vowel was not associated with the final pitch accent of the intonational phrase. The data were analysed by measuring the pitch differences between the beginning and the end final vowel. To calculate these pitch differences, four measurements were taken: 1) the mean F0 values and 2) the standard deviations for each question token as a whole, 3) the beginning F0, and 4) the end F0 for the final vowel in each question. The raw F0 differences were then calculated by subtracting the beginning F0 from the end F0. However, as the pitch values of individuals is variable, it was necessary to normalize the individual pitch variation. Figure 5 provides the equation used to normalize the F0 differences.

FIGURE 5 Equation used to normalize speaker pitch variation

$$\text{Normalized F0 difference} = \frac{(\text{endF0} - \text{meanF0})}{\text{sdF0}} - \frac{(\text{beginF0} - \text{meanF0})}{\text{sdF0}}$$

4. Results

Recall that there were two hypotheses at the outset of this paper. The first hypothesis predicted that the pitch contours produced by native English (NE) speakers on Mandarin *yes-no* questions would closely resemble canonical English *yes-no* questions. In other words, NE speakers would produce Mandarin *yes-no* questions with a similar rising intonation that is indicative of English *yes-no* questions. The second hypothesis predicted that the NE speaker pitch contours on Mandarin *yes-no* questions would differ significantly from the pitch contours produced by native Mandarin speakers. The results indicated that while the second hypothesis was supported, the first hypothesis was only partially supported.

Figure 6 presents the raw F0 differences and Figure 7 presents the mean normalized F0 differences for the NE and NM speakers for the English questions (EN), the Mandarin V bù V questions (BU), and the Mandarin *ma* particle questions (MA). The columns denote the F0 differences between the onset and the offset of the final vowels for each question type. The columns above zero indicate a rising contour while the columns below zero indicate a falling contour on the final vowel. The error bars represent +/- one standard error. These two figures show very similar patterns for both the raw and the normalized F0 differences data.

FIGURE 6 The mean raw F0 differences for the NE and NM speakers for the three question types

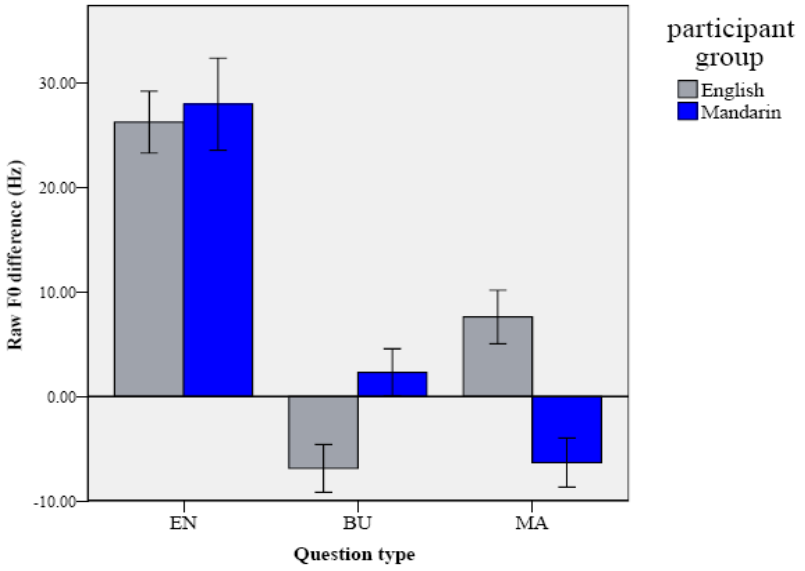
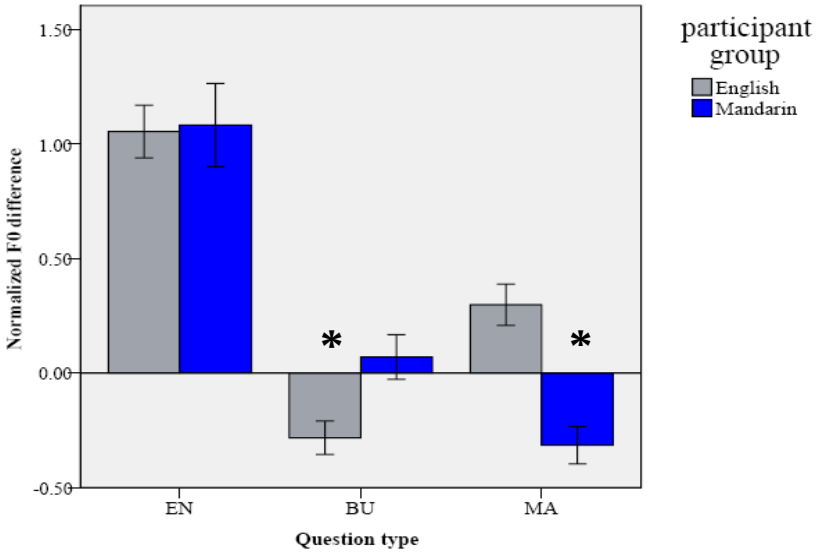


FIGURE 7 The mean normalized F0 differences of the NE and NM speakers for the three questions types



From these figures, three observations are readily apparent. First, NE and NM speakers employ as very similar degree of rising on the final vowel in EN questions. However, since this comparison was not a focus in this paper, it will not be discussed further. Second, NM speakers appear to rise slightly on BU questions that end with a high level tone. In contrast, NE speakers use a falling contour on BU questions. Third, NM speakers fall on *ma* questions while the NE speakers rise.

The next question then is: how do these data statistically compare? The statistical analyses reported below were conducted on the normalized F0 differences data. To test the first hypothesis, an analysis of variance (ANOVA) for the NE speaker data showed that the effect of question type was significant, $F(2, 70) = 50.28$, $p < 0.001$. Post hoc Tukey analyses indicated that NE speakers raised the F0 on the final vowel in EN questions significantly more than they did for BU questions ($p < 0.001$), and for MA questions ($p < 0.001$). The post hoc Tukey analyses also revealed that NE speaker F0 differences for BU questions were also significantly different from the F0 differences of their MA questions ($p < 0.001$).

To test the second hypothesis, another ANOVA was conducted to determine the statistical relationship between the NM speakers' production of the three questions types. The analysis showed that the effect of question type was significant, $F(2, 66) = 30.993$, $p < 0.001$, which indicates that there was a significant difference between the F0 differences of at least two of the question types. Post hoc Tukey analyses revealed that the NM speakers' F0 differences between the EN questions and the BU questions ($p < 0.001$) and the EN questions and the MA questions ($p < 0.001$) were significant. However, the F0 differences between the NM's BU and MA questions was not significant ($p = 0.124$). In addition, two independent samples t-tests were conducted. The first test showed that the F0 differences between the NE falling contour and the NM slight rising contour on BU questions was significant, $t(42) = -2.946$, $p < 0.001$. The second t-test indicated that the F0 differences between the NE rising contour and NM falling contour on MA questions was also significant, $t(46) = 5.04$, $p < 0.001$. These findings are discussed at length in terms of the research hypotheses in the following section.

5. Discussion

With these results in hand, it is now possible to turn to a discussion of the research hypotheses and answer the question of whether L1 prosodic contours negatively transfer over into the interlanguage prosodic system. The first hypothesis, which predicted that NE speakers' pitch contours of Mandarin *yes-no* questions would be the same as English *yes-no* questions, was only partially supported. The results indicated that NE speakers' Mandarin BU and MA question pitch contours differ significantly from the canonical rising contour of English *yes-no* questions in terms of the degree of F0 differences. However, in terms of directionality, there is some similarity between English questions and Mandarin questions, but only with the *ma* particle construction. The NE speakers did tend to rise on the final vowel in the MA questions, although not to

the same degree as the final vowels in EN questions. In contrast, the NE speakers did not rise on the final syllable of BU questions; rather, the NE speakers tended use a falling contour for this type of question. In short, the first hypothesis was only partially supported as the NE speakers only employed a rising contour for one type of Mandarin *yes-no* questions, the MA questions.

The second hypothesis, which predicted that NE speakers would not produce the same pitch contours as NM speakers for Mandarin *yes-no* questions, was supported. In other words, while the NM pitch contour differences for BU and MA questions was not significant, the NE speakers' intonation patterns did differ significantly from their NM counterparts for both types of Mandarin *yes-no* questions in terms of degree and direction of F0 difference. In fact, the NE speakers pitch contours for Mandarin *yes-no* questions was the exact opposite from the native speakers. While the native Mandarin speakers used a slight rising contour on BU questions ending in Tone 1, native English speakers used a falling contour on the same questions. Similarly, while the native Mandarin speakers used a falling contour on MA questions, the native English speakers used a rising contour on these questions. In sum, low intermediate native English learners of Mandarin demonstrated non-native-like prosodic contours for both types of Mandarin *yes-no* questions.

Although "not all L1 features directly shape L2 intonation" (Ueyama & Jun 1998, p. 644), the current data suggest that the rising contour indicative of English *yes-no* questions was, to a certain extent, negatively transferred into NE speakers' production of Mandarin *yes-no* questions. This finding corroborates, in part, Visceglia and Fodor's (2006) finding that "L1 English speakers imported a final-syllable rise on the final syllable in Mandarin questions" (p. 52). However, since Visceglia and Fodor did not investigate BU construction questions, it would be imprudent to generalize a rising contour tendency for both types of Mandarin questions. Indeed, as shown in this research, the negative transfer of L1 question prosody appears to be restricted to Mandarin MA questions. That is, NE speakers only employed and L1-like final syllable rise on MA questions while they employed a final syllable fall on BU questions.

It is not clear at this point why NE speakers treat the two types of Mandarin questions differently. The fact that NE speakers tend to rise on MA questions suggests that these constructions are perceived as questions. However, why do the NE learners fall on BU questions? One possibility is that native English learners of Mandarin consider BU questions to be reminiscent of declarative sentences. Perhaps, BU construction questions are not syntactically salient enough to trigger question intonation. Anecdotal evidence suggests that this may indeed be the case. In some language learning situations, when constructing BU questions, language learners have been observed to append a *ma* question particle to the end of a BU construction questions. That is, instead of correctly creating a BU question such as the example in (3), some learners will add the additional question marker to their questions which results in an incorrect question construction like the one given in (4).

(3) Tā **shì bú shì** xué shēng?
he/she is-not-is student
'Is s/he a student?'

(4) *Tā **shì bú shì** xué shēng **ma**?
he/she is-not-is student MA-part.
'Is s/he a student?'

So, in essence, the question in (4) is a twice marked question. The syntactic saliency of Mandarin *yes-no* questions is an intriguing question that requires further research.

Another interesting question is: if L1 question contours are negatively transferred into the interlanguage, why do the NE speakers not employ the L1 contours to the same degree in their interlanguage. Could this be due to the tone marking inherent in the Pinyin orthography? It is possible that the visual tone marks in the Pinyin system limit the effects of negative prosodic transfer. Therefore, it is logical to ask: do the tone marks constrain the prosodic patterns of the language learners? In other words, would the F0 differences be higher for Mandarin *yes-no* questions when learners read in characters as opposed to when they read in Pinyin? Again, this is research that needs to be investigated in the future.

6. Conclusion

This research suggests that L1 prosodic contours negatively transfer into the interlanguage grammar and influence the production of the target language. The data indicated that native English learners of Mandarin do not use the same prosodic contours for Mandarin *yes-no* questions that native speakers do. Specifically, native English learners of Mandarin tend to employ a final syllable rise on Mandarin MA particle questions. Native Mandarin speakers, on the other hand, tend to use a falling contour on MA questions when the *ma* particle is preceded by a Tone 1. This disparity between the L1 and L2 speakers suggests that L1 prosodic contours do, in fact, influence the production of an L2. That is not to say, however, that BU questions do not also exhibit negative transfer. While NE speakers do not exhibit the same prosodic contours for BU questions, it is still possible that this type of question also demonstrates the effects of negative L1 transfer as the prosodic contour does not reflect the same pattern as native Mandarin speaker production. Instead, NE speakers may negatively transfer the falling contour pattern of English declarative sentences due to the lack of syntactic saliency of Mandarin BU construction questions.

References

Archibald, John. 1998. *Second Language Phonology.* Amsterdam: John Benjamins Publishing Company.

Beckman, Mary E, and Janet Pierrehumbert. 1986. Intonational Structure in Japanese and English. *Phonology Yearbook*, 3, 255-309.

Cruttenden, Alan. 1997. *Intonation: 2nd Edition*. Cambridge: Cambridge University Press.

García Lecumberri, M.L. 2001. Native language influences in learners' assessment of English focus. *International Journal of English Studies*, 1(1), 53-71.

Lin, Hua. 2001. *A Grammar of Mandarin Chinese.* Munich: LINCOM Europa.

Major, Roy. 2001. *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology.* Mahwah: Lawrence Erlbaum Associates, Inc.

Mennen, Ineke. 2004. Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics*, 32, 543-563.

Pierrehumbert, Janet, and Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In *Inventions in Communication, eds.* P. Cohen, J. Morgan, and M. Pollack, 271-311. Cambridge, MA: MIT Press.

Raiser, Laurent, and Philippe Hiligsmann. 2007. Prosodic transfer from L1 to L2. Theoretical and methodological issues. *Nouveaux cahiers de linguistique française*, 28, 41-66.

Ramírez Verdugo, Dolores. 2006. Prosodic realization of focus in the discourse of Spanish learners and English native speakers. *Estudios Ingleses de la Universidad Complutense*, 14, 9-32.

Shen. Xiao-nan Susan. 1990. *The Prosody of Mandarin Chinese.* Berkeley, CA: University of California Press.

Ueyama, Motoko, and Sun-Ah Jun. 1998. Focus realization in Japanese English and Korean English intonation. *Japanese/Korean Linguistics 7*, Stanford, CSLI, pp.629-645.

Visceglia, Tanya, and Janet D. Fodor. 2006. Fundamental frequency in Mandarin and English: Comparing first- and second-language speakers. In *Interfaces in Multilingualism: Acquisition and Representation,* ed. Conxita Lleao, 27-59. Amsterdam: John Benjamins.