B4

Effects of Mandarin Tones on the Production of English Boundary Tones

Min Zeng Graduate School of Education Waseda University

1．Brief introduction

Prosody is present in every spoken utterance, such that any utterance—no matter how short or in which language it is spoken— must have a certain duration, loudness, or pitch (Mennen and Leeuw, 2014). Hence prosody is undoubtedly very crucial in communication. It conveys a variety type of information. Prosody can, for example, indicate the narrative or interrogative mood of an utterance. If the sentence *They went to NEWYORK* were said with a falling pitch, it often represents a statement. Whereas if the same sequence of the words were said with a rising pitch, in combination with lengthening the final one or two syllables, it often indicates a question. Besides, prosody can be used to show the salience of important information in an utterance. If one were to say *They went to NEWYORK*, NEWYORK may be emphasized as it is contrasted with something else, such as *They went to New York, not Los Angeles*. This emphasis may achieve by using higher pitch, longer time, or louder voice for word NEWYORK in comparison to the words and syllables that are not salient. According to Trofimovitch & Baker (2006) and Rasier & Hiligsmann (2007), "Deviations in the production of L2 stress, rhythm and intonation may affect listeners' judgments more than deviations in the production of L2 vowels and consonants" – thereby showing the importance of prosodic properties for second language (L2) learning and usage

2．Literature review

Prosodic properties are notoriously difficult to acquire because prosodic transfer occurs from the native language to the target language. In English learning, for example, they are often seen as "the final hurdle, which a vast majority of speakers of English as a foreign language never manage to cross" (Banjo, 1979). Research about prosodic properties in Second Language Acquisition (SLA) has been focused on several aspects, such as the above-mentioned intonation, rhythm, and stress. Among intonation, fundamental frequency (F0) files are widely compared between native languages and target languages.

2.1 Previous studies on intonation of English and Mandarin

English is a stress language and has a system of culminating word stress, but Mandarin, a typical tone language, has no system of word stress. Rather, Mandarin has a system of lexically distinctive tones. Comparing the two languages is very interesting and meaningful. Keating and Kuo (2012) compared speaking F0 profiles of English and Mandarin. The result showed that the two languages' F0 profiles were sometimes found to differ, but these differences depended on the speech samples being compared. In studies comparing native English speakers and Mandarin EFL learners, Ploquin (2013) invited six Mandarin speakers to record English and Chinese sentences and then compared their voice with native English speakers. Results showed that the subjects produced a Mandarin tone 2 where a rising pitch accent was required and thus that Mandarin speakers relied on their lexical tones inventory to produce English prosody. In a more recent study, Ding, Hoffmann, and Hirst (2016) compared the F0 patterns of continuous speech in English, Mandarin and English produced by Mandarin EFL learners. The results

showed that Mandarin EFL learners' speech displayed a significantly larger value than native English speakers. Moreover, the same Mandarin subjects demonstrated still a larger value in their Chinese speech.

According to the above-mentioned researches, Mandarin speakers have a higher F0 level than native English speakers (Keating & Kuo, 2012). Mandarin EFL learners seem to have a wider pitch range than native English speakers, but only in the single-word level (Ding, Hoffmann & Hirst, 2016). These studies, however, analyzed data of declarative sentences or passages. There are only a few studies compared F0 profiles of Mandarin and English yes-no questions. Viger (2007) investigated differences in yes-no questions between the English and Mandarin utterance-level prosodic contours produced by native and non-native speakers. A bi-directional design permitted a double comparison of L1 English with L2 English by Mandarin speakers, and of L1 Mandarin with L2 Mandarin by English speakers. Results showed that both participant groups exhibited a considerably narrower F0 range overall in their L2 than in their L1. There was no sign of transfer of global prosodic strategies from Mandarin into English. Two questions could be raised to this study. First, the participants of this study were all females. Secondly, all participants started learning English after puberty. The English education of the 21-century in China, however, starts from the elementary level.

3.　Research questions

Bearing the problems mentioned above, this study is designed to compare the F0 profiles between native English speakers and Mandarin EFL learners who started learning English in elementary school. The research questions are as follows.

　　1.　How the pitch range of Mandarin EFL learners and native English speakers differ in producing English yes-no questions, declarative sentences and echo questions?

　　2.　Do bilingual (Mandarin and English) and trilingual (Mandarin, English, and Japanese) speakers exhibit different pitch range in producing English yes-no questions, declarative sentences and echo questions?

4.　Method

4.1 Overview

This study was designed to investigate the differences between the English and Mandarin utterance level F0 profiles produced by native and non-native speakers. The materials were controlled for sentence length and illocutionary type. Pitch range was the particular focus of the study. Of interest was whether Mandarin bilingual and trilingual learners would show a statistically significant difference in producing three illocutionary types of English sentences (yes-no questions, declarative sentences, and echo questions).

4.2 Hypothesis and Predictions

Based on the previous research (Keating & Kuo 2012; Ploquin 2013), which demonstrated a higher pitch level in Mandarin L1 than English L1, it is predicted that Mandarin EFL learners will exhibit higher pitch level than English native speakers in producing three illocutionary types of English sentences.

4.3 Participants

The participants in this study were all adult Mandarin EFL learners, with advanced proficiency level (TOEFL scored no less than 80 or IELTS scored no less than 6.5). A questionnaire was administered before the voice

recording, which checked their language background, including age of beginning language instruction, amount of in-country exposure, and experiences in English speaking countries. All participants had begun English learning before puberty, had acquired most of their language skills in formal classroom settings, and were university students or Japanese language school students. There were two groups of Mandarin EFL learners. Members of the bilingual group were all advanced English learners, and members of the trilingual group were all advanced English and Japanese learners.

The members of native English group were American. All of them were born and raised in America and have been in non-English speaking countries for less than 6 months. There were twenty Mandarin EFL learners (ten bilingual and ten trilingual speakers) and ten native English speakers.

4.4 Materials

Experimental materials were designed to elicit F0 values from three sentence types: yes-no question, declarative sentence, and echo question. All sentences consisted of nine syllables in yes-no questions and eight syllables in declarative sentences and echo questions. The English experimental sentences ended with (a) final, (b) penultimate, and (c) antepenultimate nuclear-stressed syllables. There were twelve experimental items and six fatigue items. The twelve experimental stimuli at the end of each sentence were "wall", "line", "fan", "pin", "onion", "winner", "lawyer", "money", "enemy", "animal", "foreigner", and "millionaire".

4.5 Procedures

Participants were seated before the same computer in a quiet room to do voice recording. They were shown with the same powerpoint which gave them the experimental materials. They were instructed to speak the materials as natural as possible as if they were talking to a friend. They were also instructed to produce each sentence after they feel ready and comfortable to speak. Each experimental sentence got three repetitions.

Each set of materials included twelve experimental items, with three practice items at the beginning, and a fatigue item after two sets of experimental items. Each experimental item consisted of a three-sentence mini-dialogue containing a sequence of (1) a yes-no question, (2) an affirmative answer to that question, and (3) an echo question to confirm that answer. To minimize the possibility that the production of one sentence set would be influenced by the semantic content of the preceding set, a filler sentence (The next sentence set is number X) was read aloud after each experimental set.

The experimental procedures were carried out in the following order:

1.   Participants read a description of our experiment and sighed a consent form, which explicitly granted permission to record their voice (administered in the participant's L1).

2.   Participants filled out a language background questionnaire (administered in the participant's L1).

3.   Participants read the experimental instructions (administered in English).

4.   Participants spoke the experimental sentences aloud.

5.   Praat was used to elicit pitch values of the experimental stimuli (including Nuclear Pitch Accent, NPA is the last pitch accent in a phrase; Phrase Accent, PA is an additional tone after the NPA; and Boundary Tone, BT is a rise or fall in pitch at the end of the intonational phrases or sentences), according to Tone and Break Indices (ToBI) system (Ladefoged & Johnson 2014, Silverman et al. 1992,)
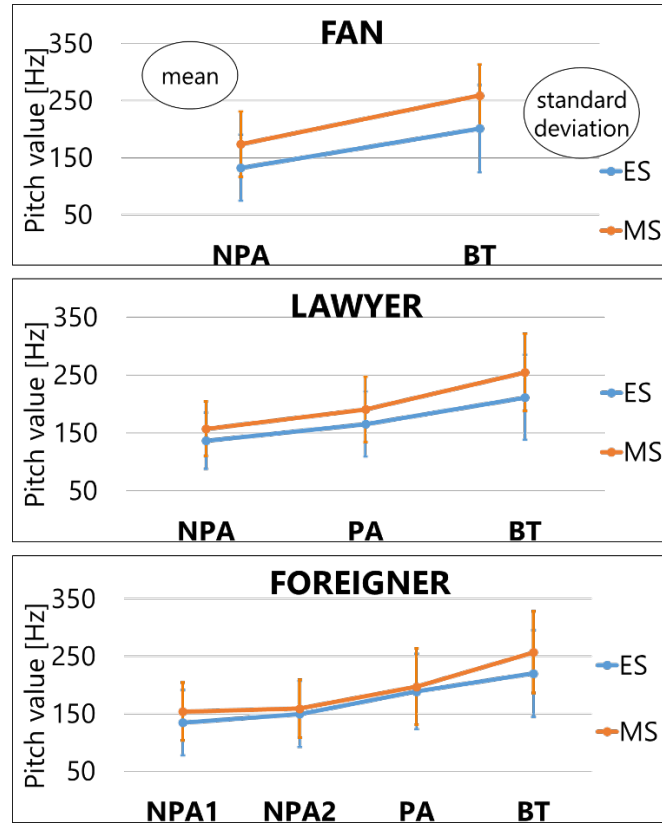
Figure 1: Pitch range from NPA to BT of the ES and MS groups in English yes-no questions

5. Results

This section focuses on differences in the production of F0 profiles between Mandarin EFL learners and American English speakers. MANOVA (Multivariate Analysis of Variance) was performed to compare the pitch values of the two groups. The data of yes-no questions including eleven native English speakers (6 males and 5 females) and twelve Mandarin EFL learners (6 males and 6 females, bilingual and trilingual speakers were treated as the same group) have been analyzed. The MANOVA results of words FAN, LAWYER and FOREIGNER are used as examples.

Figure 1 shows the pitch range from NPA to BT of the English speakers and Mandarin speakers in English yes-no questions. English speakers and Mandarin speakers show a clear pitch rising in English yes-no questions. Mandarin speakers exhibit a higher pitch level than English speakers.

In testing whether the general result is statistically significant, MANOVA was performed three times to check each word respectively. Table 1 shows the results of MANOVA which compares pitch values (FAN, LAWYER, and FOREIGNER, respectively) between the English speakers and Mandarin speakers. The independent variables were Nationality and Gender. The dependent variables were NPA and BT in FAN; NPA, PA, and BT in LAWYER; and NPA1, NPA2, PA, and BT in FOREIGNER.

The Box's Test of Equality of Covariance Matrices checks the assumption of homogeneity of covariance across the two groups using $p < .05$ as a criterion. P-values (.027, .011, .001) < (.05) indicates that there are significant differences between the covariance matrices. The following is the MANOVA using the Pillai's Trace

	Box's Test	Multivariate Test - Pillai's Trace						
FAN	Sig.	Effect	Value	F	df	Error df	Sig.	Partial Eta Squared
		Intercept	.972	1104.54	2	64	**.000**	.972
	.027	**Nationality**	.323	15.289	2	64	**.000**	**.323**
		Gender	.666	63.892	2	64	**.000**	.666
		Nationality * Gender	.010	.327	2	64	**.722**	.010

	Box's Test	Multivariate Test - Pillai's Trace						
LAWYER	Sig.	Effect	Value	F	df	Error df	Sig.	Partial Eta Squared
		Intercept	.979	951.624	3	62	**.000**	.979
	.011	**Nationality**	.197	5.059	3	62	**.003**	**.197**
		Gender	.753	63.003	3	62	**.000**	.753
		Nationality * Gender	.025	.536	3	62	**.660**	.025

	Box's Test	Multivariate Test - Pillai's Trace						
FOREIGNER	Sig.	Effect	Value	F	df	Error df	Sig.	Partial Eta Squared
		Intercept	.980	761.065	4	62	**.000**	.980
	.001	**Nationality**	.189	3.619	4	62	**.010**	**.189**
		Gender	.786	56.988	4	62	**.000**	.786
		Nationality * Gender	.101	1.742	4	62	**.152**	.101

Table 1: Results of **MANOVA** to compare pitch values (FAN, LAWYER, and FOREIGNER, respectively) between the ES and MS groups

test. Using an alpha level of .05, we see that these tests are significant, Pillai's = .323, $F(2, 64) = 15.29$, $p < .05$, multivariate $\eta^2 = .32$ for FAN; Pillai's = .197, $F(3, 62) = 5.06$, $p < .05$, multivariate $\eta^2 = .20$ for LAWYER; Pillai's = .19, $F(4, 62) = 3.61$, $p < .05$, multivariate $\eta^2 = .19$ for FOREIGNER. These significant Fs indicate that there are significant differences among the English speakers and Mandarin speakers on a linear combination of the two, three and four dependent variables, respectively. Although the sample sizes were relatively small and the data were not in normal distribution, since Pillai's Trace test is a test statistic that is very robust and not highly linked to assumptions about the normality of the distribution of the data, we would like to say that the results were statistically significant.

6. Discussion and further study

According to the results of the statistical analysis, Mandarin speakers have a higher pitch level than English speakers in producing L2 English yes-no questions. Mandarin's high pitch level influences Mandarin speakers in producing English intonation (Ding, Hoffmann & Hirst, 2016; Eady, 1982; Ploquin, 2013).

The following steps of the study will be analyzing the remaining data for larger sample sizes to see if there were a change in the results of the Box's Test. Also, performing MANOVA for declarative sentences and echo questions to see if the results were consistent. Furthermore, dividing the Mandarin speakers into two groups - bilingual and trilingual group to see if the two groups had significant differences in pitch range at boundary tones.

7. References

Banjo, A. (1979). Beyond intelligibility in Nigerian English. In E. Ubahakwe (Ed.), *Varieties and functions of English in Nigeria* (pp. 7–13). Ibadan, Nigeria: African Universities Press .

Braun, B., Galts, T., & Kabak, B. (2014). Lexical encoding of l2 tones: The role of l1 stress, pitch accent and intonation. *Second Language Research*, *30*(3), 323–350. https://doi.org/10.1177/0267658313510926

Ding, H., Hoffmann, R., & Hirst, D. (2016). *Prosodic transfer: A comparison study of f0 patterns in l2 english by chinese speakers*. 756–760. https://doi.org/10.21437/SpeechProsody.2016-155

Kainada, E., & Lengeris, A. (2015). Native language influences on the production of second-language prosody. *Journal of the International Phonetic Association*, *45*(3), 269–287. https://doi.org/10.1017/S0025100315000158

Keating, P., & Kuo, G. (2012). Comparison of speaking fundamental frequency in english and mandarin. *Acoustical Society of America*, *132*(2), 1050–1060.

Ladefoged, P., & Johnson, K. (2014). *A course in phonetics*. Nelson Education.

Li, A., & Post, B. (2018). L2 acquisition of prosodic properties of speech rhythm: Evidence from l1 mandarin and german learners of english. *Studies in Second Language Acquisition*, *36*, 223–255. https://doi.org/10.1017/S0272263113000752

Mennen, I., & de Leeuw, E. (2014). Beyond segments. *Studies in Second Language Acquisition*, *36*(02), 183–194. https://doi.org/10.1017/S0272263114000138

Nguyễn, T. A. T., Ingram, C. L. J., & Pensalfini, J. R. (2008). Prosodic transfer in vietnamese acquisition of english contrastive stress patterns. *Journal of Phonetics*, *36*(1), 158–190. https://doi.org/10.1016/j.wocn.2007.09.001

Ortega-llebaria, M., & Colantoni, L. (2014). L2 english intonation: Relations between form-meaning associations, access to meaning, and l1 transfer. *Studies in Second Language Acquisition*, *36*, 331–353. https://doi.org/10.1017/S0272263114000011

Ploquin, M. (2013). Prosodic Transfer: From chinese lexical tone to english pitch accent. *Advances in Language and Literary Studies*, *4*(1), 68–77. https://doi.org/10.7575/aiac.alls.v.4n.1p.68

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., ... & Hirschberg, J. (1992). ToBI: A standard for labeling English prosody. *In Second international conference on spoken language processing*.

Viger, T. L. (2007). *Fundamental frequency in Mandarin Chinese and English: Implications for second-language speakers* (Doctoral dissertation, City University of New York).

伍艳红, 陶建华, & 路继伦. (2006). 汉语疑问语调的韵律分析. In *第七届中国语音学学术会议论文集*.