ENGLISH LEARNERS' PERCEPTION AND PRODUCTION OF MANDARIN INTONATION

Shan Luo Hua Lin
University of Victoria

ABSTRACT

This study investigates pitch interplay in non-native (L2) speech. Specifically, experiments were conducted to examine how native English speakers perceive and produce Mandarin statements and unmarked questions. Results from the perceptual experiment showed English listeners did not resolve the simultaneous pitch cues from intonation and tone the same way as native Mandarin speakers. They had less difficulty identifying intonation when the pitch movement of intonation and tone was in the same direction versus when they were not. The production experiment showed that the intelligibility of the English speakers' question intonation was especially compromised due likely to their narrower pitch range than that found in native production.

Keywords: intonation, unmarked questions, Mandarin, L1 influence, second language acquisition

1. INTRODUCTION

Many studies have long recognized the importance of intonation learning in L2 acquisition [e.g., 1,2,8,9,10,11,18,19]. These studies have found that language experience influences L2 intonation at both the phonetic and phonological levels. An example of L1 phonetic influence is found in L2 pitch range [2,3,11], whereas an case of L1 phonological influence is found in L2 use of rises where native speakers would use falls, and vice versa [1,9]. As part of the ongoing efforts toward understanding L2 prosodic development, we conducted experiments to examine the intonation-tone pitch interplay in English speakers' perception and production of Mandarin sentences. Our study focuses on two types of Mandarin sentences, statement and question.

Before we present on the experiment, a brief look at the English and Mandarin intonation is in order. In English unmarked questions, pitch typically rises to a high final boundary tone (H%) such that English listeners are perceptually accustomed to hearing a rise in question modality and have a preference for such a combination [14,15,21]. In Mandarin, the whole body of a question utterance rises higher than that of a statement [7,16,17, 23,25]. Previous studies have also revealed a local pitch rise starting at or near the end of the sentence in Mandarin questions. The interaction between Mandarin intonation and tone is believed to be maximal on the sentence-final syllable [16,17,25]. The data on this well-studied topic suggest that Mandarin questions cause the value of sentence-final tone to increase when the rising intonation movement coincides with the direction of tone (i.e., a rising tone), while the tonal value tends to be neutralized when the two directions are opposing (i.e., a falling tone) [17].

Such pitch interplay raises a question of how native listeners interpret simultaneous pitch cues; for instance, whether the sentence-final tone becomes indistinguishable. One view is that intonation has little effect on native Mandarin listeners' tone perception [4]. However, a recent study [24] reported that Mandarin participants' perception was indeed compromised by the interaction of sentential intonation and lexical tone. The results of [24] found that native listeners' statement (S) identification was more accurate than their question (Q) identification, and that many Qs were misidentified as Ss. Further analysis showed that the sentence-final tone did not affect S identification but did affect Q identification. An interesting and more specific finding was that Q perception was more accurate if the sentence final tone was T4 (a falling tone), whereas it was less accurate if it was T2 (a rising tone).

It is not well understood how English listeners will perform differently from (or in the same way as) native Mandarin speakers in Mandarin intonation perception and production. A few studies indicate that L2 listeners might have different perceptual strategies from Mandarin listeners [10, 22]. One specific question to ask is how English speakers are affected by compatible/conflicting pitch environments. In view of the previous findings, our research aims to answer two questions: 1) How English speakers tease apart intonation and tone information? 2) How their production is affected by compatible/ conflicting pitch interplay?

2. THE EXPERIMENT

2.1 Stimuli

As mentioned earlier, the study focuses on statement and question. The reason for choosing the two are: a) these two types of sentences show an interesting interface between pragmatic and intonational functions without the interference of syntactical cues,

and they thus provide us with an ideal setting to examine pitch interplay; b) both Mandarin and English have the two kinds of modality, allowing us to directly compare their intonation patterns.

The stimuli used in this study are 64 Mandarin sentences, comprised of 32 minimal pairs of statements (S) and unmarked questions (Q). None of them contain syntactic markers signalling their sentence modality; that is, all sentences can be interpreted as either an S or a Q based on intonation alone. The 64 utterances were divided into eight groups, each containing eight utterances with the same number of syllables and the same grammatical structure. The eight sentences in each group contained combinations of four lexical tones and two intonations respectively. The goal was to examine the interaction between intonation and sentence-final tones. One female native Mandarin speaker recorded all the stimuli.

2.2 Participants

Ten L2 Mandarin learners (F=4, M=6) with an average age of 24.8 years participated in the study. These participants had been learning Mandarin for at least one and a half year with an average learning time of four years; nine of them considered themselves as above intermediate level in their overall proficiency in Mandarin. All participants were native North American English speakers (9 Canadians, 1 American). All participants were born in English-speaking families and began their Mandarin learning after puberty. None of them had experience learning another tonal language.

2.3 Procedures

A pre-task was conducted to assess participants' proficiency levels in addition to the self-reported evaluation. Each participant was given four sentences in Mandarin. The sentences were adopted from [6] and recorded by the female native speaker who recorded all perceptual materials. Participants heard the four sentences, and recordings were made as they repeated them. Two native Mandarin speakers listened to the recordings and rated their accentedness. The average accent score over the ten participants was 3.3/10, which indicates that the participants were perceived to be slightly accented.

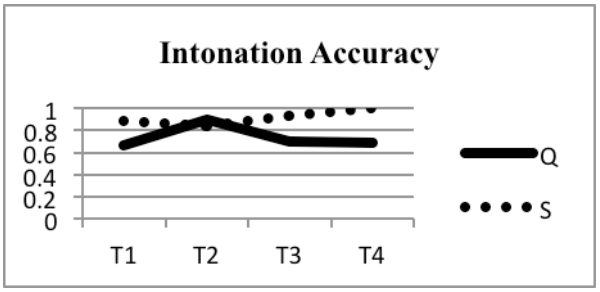
The primary experiment used E-Prime 2.0 and consisted of three tasks. Each task used the same 64 sentences in random order and each sentence appeared once in each task. In Task 1, participants listened to one sentence at a time and were asked to choose the sentence intonation (S vs. Q). In Task 2, participants were asked to identify the sentence-final tone (T1, T2, T3 or T4) when listening to the sentences again. Task 3 was a production task that involved listening to a sentence and producing it with a different intonation: if the sentence heard was a statement, participants were asked to produce it as a question, and vice versa. To reduce memory-related effects, all sentences were given on the screen in Task 3, and they were transcribed in the Romanized Mandarin alphabet *pinyin* with tone added to the tone carrying vowel symbols of all but the last syllable. Before the primary experiment, a training session was given to familiarize the participants with the procedure. The training session had the same design as was used in the primary experiment but with fewer and different testing materials (eight sentences in each training task). The whole procedure took approximately 30-45 minutes.

3. ANALYSIS AND RESULTS

3.1. Intonation Perception

In Task 1, a total of 640 responses (10 subjects * 64 sentences) were generated and subjected to analysis. The results showed that the average intonation identification accuracy was 82.5%, and that participants were more accurate at identifying statement (S) intonation (91.25%) than question (Q) intonation (73.75%). As expected, not all intonation+tone combinations were equally difficult to discriminate. The highest accuracy for S was with T4 (100%), and the highest accuracy for Q was with T2 (90%). The lowest accuracy for S was with T2 (83.75%) and that for Q was with T1 (66.25%). Note that S+T4 and Q+T2 are in compatible pitch environments, while S+T2 and Q+T4 are in conflicting environment. Figure 1 below summarizes the results of intonation perception.

Figure 1: Perception results of intonation



The effects of intonation and tone on the identification of intonation were analyzed using Two-way Repeated-Measures by subject ANOVAs. The two independent variables were sentence type

(S vs. Q) and tone (T1, T2, T3, T4). The dependent variable was average accuracy. Results showed that the main effect of sentence type was statistically significant, [F(1, 9)=20.103, p<0.002], with S easier to identify than Q. The main effect of tone was not significant [F (3, 27) =1.803, p<0.174].

However, considering that the interaction between sentence type and sentence-final tone was significant [F (3, 27)=6.695, p<0.004], the simple effect of sentence type was tested for each tone separately. Results showed that the effect of sentence type was significant for every tone (all p<0.001) except T2 (p<0.337). Multiple comparisons with Bonferroni corrections showed that tone had a significant effect on intonation identification for both S (p<0.017) and Q (p<0.015). This was primarily due to T2, which behaved differently from other tones in both statements and questions. In S, the only significant difference among tones was between T2 and T4 (p<0.013). In Q, the only significant difference was between T2 and T1 (p<0.024) while T2 vs. T4 was marginally significant (p<0.056). The reason that the main effect of tone was not significant is likely because the effect of T2 was opposite in S and Q: participants performed *significantly worse* at identifying intonation in S+T2 and *significantly better* at identifying intonation in Q+T2. This is expected since the pitch direction in S+T2 is conflicting but compatible in Q+T2.

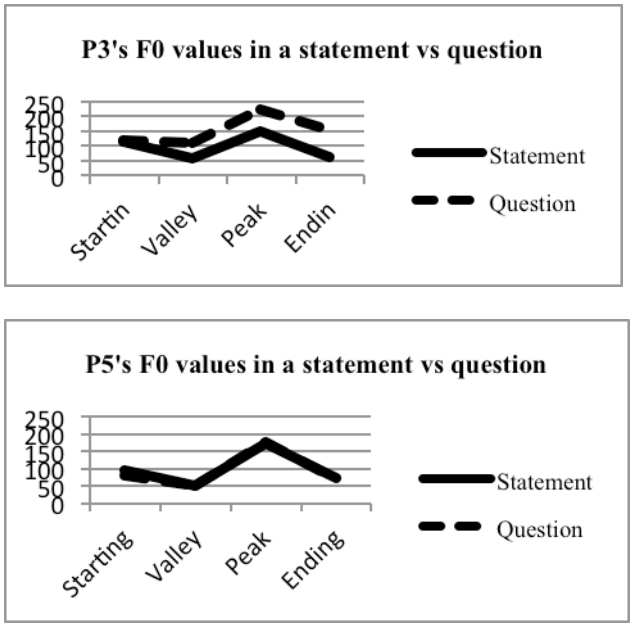
3.2. Intonation Production

Two Mandarin-speaking judges were presented with pairs of sentences (S vs. Q) produced by the participants, and were explicitly asked to identify which was the statement (S) and which was the question (Q). The first two participants' data were excluded due to slight revisions to the stimuli after their participation, which meant that eight participants' data were subject to analysis (F=4, M=4). Each participant produced a set of 64 sentences, yielding a total of 512 tokens (64 tokens * 8 participants). The results indicate that the two judges did not always agree, but the inter-rater reliability is fairly high at 79.5% (or 105/512 sentences). Correlation calculations showed that the two judges' ratings were strongly correlated [r (8)=0.951, p <0.001].

Intonation production error was calculated by counting the intended intonations that were incorrectly rated at least by one judge. The average accuracy across the eight participants was 56.25%, which is quite low considering the participants' low accent ratings and their above 80% perceptual accuracy. The results showed that only 40.6% of intended Qs were identified as such while intended Ss were identified as such 71.5% of the time. A two-tailed t-test established that the difference between Q identification and S identification was statistically significant (p<0.01), suggesting that participants performed significantly worse in producing Qs.

A closer inspection reveals that a considerable amount of inaccuracy lies in the participants' failure to produce a distinction between S and Q. This is especially to see in the data of two male participants (P5 and P10) who produced less than 20% of intended questions and yet were rated among the top 3 most fluent speakers by both judges. Their perceived speech fluency in terms of accentedness did not give them any advantage in producing the modality distinction. We measured the F0 values at four points (starting, valley, peak, and ending) from a randomly chosen sentence pair (S vs. Q) produced by P5. The F0 values of the four points were then compared to those produced by P3 (the most successful participant in intonation production task, also male) for the same pair. Figure 2 below shows that P5 did not produce much pitch variation in the two different sentence types: the F0 values at four points almost coincided in S and Q. Also, P5's pitch range was significantly smaller than P3's (50-180 Hz vs. 50-220 Hz).

Figure 2: F0 value comparisons of P3 and P5 in S-Q



Overall, the same pattern of a) a narrow pitch range and b) no variation between intonations was found across participants' production. Meanwhile, both P3 and the female native speaker who recorded all material for perceptual task had more than 30Hz span in pitch range in producing different intonations. Assuming pitch range is one of the

factors that lead to the intelligibility issues, we conducted one-sample t-test to examine whether there was a difference between participants' and the native speaker's pitch ranges. We found that, aside from P3, all other participants' mean F0 differences between S and Q were significantly different from the native speaker's [t (6)=-5.775, p <0.001]. This difference explains these participants' failure in making intonational distinction in production. It also establishes that native judgment is grounded in acoustic cues of pitch.

4. DISCUSSION

The perceptual findings in this study support previous research, which showed differences between native and non-native processing. First, L2 learners are affected by the conflict between intonation and tone cues differently from the native speakers (also see [22]). For instance, English speakers had particular difficulty processing sentences in which intonation and sentence-final tone moved in opposite directions; they were much more successful at identifying intonation and tone in a compatible pitch direction. That is, they found Q+T2 (both rising) and S+T4 (both falling) sentences the easiest for intonation identification, while Q+T4 (rising intonation + falling tone) and S+T2 (falling intonation + rising tone) the most two difficult. These findings contrast with L1 studies (e.g., [23, 24]), which showed that native Mandarin speakers found intonation in Q+T4 the easiest to identify.

Secondly, tone affects L2 intonation identification more significantly than L1 intonation perception. This study revealed that sentence-final tone affected L2 perception of both S and Q sentence types (see [24] for a different story that sentence-final tone only affected Q identification for Mandarin listeners). That is, English participants are more subject to pitch interplay impact than Mandarin listener. This is most likely due to language experience: only intonation is of phonological significance in English, resulting in the pitch interplay more confusing for the English speakers than for the native group. Additionally, our perceptual findings also suggest universal markedness is at work. In this study, Q is found harder to identify than S. This result is in line with findings from previous studies on L2 as well as on L1 perception [22, 24].

An interesting finding related to both universal preference and language experience is that the English participants found it difficult to indentify intonation in Q+T1 (T1: a high level tone): such combinations were mostly considered as statements. The study [15] showed English speakers' preference for having rising intonation in question modality; otherwise, they would rate those sentences not polite and unacceptable. [21] reported that English speakers could not tell whether a sentence was Q or S when a falling tone occurred sentence-finally. Our study further reveals that not only can English speakers not distinguish sentence types when a final falling pitch is involved, but they also cannot determine a question without a clear terminal-rising pitch. A flat contour, even if it is high in F0, still cues S rather than Q for English listeners. That is, they systematically rely on terminal-rising as a mechanism for Q association.

Participants' perception difficulty of Q is reflected in their production as well. The results indicate that English speakers had more difficulty producing intelligible questions (40.6%) than statements (71.5%). This again could be explained in terms of markedness: question intonation is more marked than statement intonation. Ultimately, the difficulty and the markedness hierarchy could have a physiological explanation. In [12, 13], rising pitch is associated with raising the larynx (against gravity) while falling pitch with lowering the larynx.

The observation of a smaller pitch range exhibited by English speakers in this study is also consistent with other studies concerning L2 acquisition of English intonation patterns [2,8]. The explanation could lie in the fact that L2 learners in general tend to produce a narrower and flatter pitch range, underutilizing the prosodic means available to them in the L2 [2,3, 11, 20]. This is indeed the case in this study: both native judgement and acoustic analysis showed a smaller pitch range in the non-native as opposed to the native production of the sentences. The consequence of this narrower pitch range is that the intonation patterns in the L2 production are not as recognizable by the judges as the full-pitch-range native intonation.

5. REFERENCES

[1] Backman, N. (1979). Intonation errors in second language pronunciation of eight Spanish speaking adults learning English. *Interlanguage studies bulletin*,*4*(2), 239-266.

[2] Barlow, J. S. (1998). Intonation and second language acquisition: a study of the acquisition of English intonation by speakers of other languages. PhD Dissertation, University of Hull.

[3] Chen, G.T. (1974). The pitch range of English and Chinese speakers. *Journal of Chinese Linguistics* 2, 159-171.

[4] Connell, B. A., Hogan, J. T., & Rozsypal, A. J. (1983). Experimental evidence of interaction between tone and intonation in Mandarin Chinese. *Journal of phonetics*.

[5] DeFrancis, J.F. (1963). *Beginning Chinese*. New Haven: Yale University Press.
[6] Guo, J., & Shi, F. (2011). Experimental Study of Declarative and Interrogative Intonation in English and Chinese. *Contemporary Foreign Languages Studies*, 5-11.
[7] Ho, A. T. (1976). The acoustic variation of Mandarin tones. *Phonetica*, *33*(5), 353-367.
[8] Jilka, M. (2007). Different manifestations and perceptions of foreign accent in intonation. In J. Trouvain and U. Gut (Eds.), *Non-native Prosody: Phonetic Description and Teaching Practice* (pp. 77-96). New York: Mouton de Gruyter
[9] Lepetit, D. (1989). Cross-Linguistic Influence in Intonation: French/Japanese and French/English. *Language Learning*, *39*(3), 397-413.
[10] Liang, J., & van Heuven, V. J. (2007). Chinese tone and intonation perceived by L1 and L2 listeners, In C. Gussenhoven & T. Riad (Eds.), *Tones and Tunes, Volume 2: Experimental studies in word and sentence prosody.* Vol. 12-2 (pp. 27-61). Berlin/NewYork: Mouton de Gruyter.
[11] Mennen, I. (2007). Phonological and phonetic influences in non-native intonation. In J. Trouvain and U.Gut (Eds.), *Non-native Prosody:Phonetic Description and Teaching Practice* (pp. 53-76). New York: Mouton de Gruyter.
[12] Moisik, S., Lin, H., & Esling, J. (2013). Larynx height and constriction in Mandarin tone. In Gang Peng & Feng Shi (eds). Eastward Flows the Great River: Essays on Contemporary Linguistics and Language Studies. Festschrift in honor of Professor William S-Y. Wang on his 80th Birthday. Hong Kong: City University of Hong Kong Press. p. 187-205.
[13] Moisik, S., Lin, H., & Esling, J. (2014). A study of laryngeal gestures in Mandarin citation tones using simultaneous laryngoscopy and laryngeal ultrasound (SLLUS). Journal of the International Phonetic Association, 44 (01), pp. 21 – 58.
[14] Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P.Cohen, J.Morgan, M.Pollack (Eds.), *Intentions in communication* (pp. 271-311). MIT Press, Cambridge, MA.
[15] Scherer, K. R., Ladd, D. R., & Silverman, K. E. (1984).Vocal cues to speaker affect: Testing two models. *The Journal of the Acoustical Society of America*, *76*(5), 1346-1356.
[16] Shen.-N.S. (1986). Tone, stress, and intonation in Mandarin Chinese. PhD dissertation, University of California, Berkeley.
[17] Shen.-N.S. (1990). *The prosody of Mandarin Chinese*. Berkeley, CA: University of California Press.
[18] Ueyama, M. (2000). *Prosodic transfer: An acoustic study of L2 English vs. L2 Japanese*. PhD dissertation, University of California Los Angeles.
[19] Vaissière, J. (2005). Perception of Intonation, In D.B.Pisoni and R.E.Remez (Eds.), *The Handbook of Speech Perception* (pp. 236-263). MA: Blackwell, Malden.
[20] Visceglia, T. & Fodor, J.D. (2006). Fundamental frequency in Mandarin and English: Comparing first- and second-language speakers. In C. Lleoa (Ed.), *Interfaces in Multilingualism: Acquisition and Representation* (pp.27-59). Amsterdam: John Benjamins.
[21] Wells, J.C. (2006). *English Intonation*. Cambridge: Cambridge University Press
[22] Yang, C., & Chan, M. K. (2010). The perception of Mandarin Chinese tones and intonation by American learners. *Journal of Chinese Language Teachers Association*, *45*(1), 7-36.
[23] Yuan, J.H. (2004). Intonation in Mandarin Chinese: Acoustics, Perception, and Computational Modeling. PhD Dissertation. Cornell University, Ithaca.
[24] Yuan, J.H. (2011). Perception of intonation in Mandarin Chinese. *The Journal of the Acoustical Society of America*, 130 (6), 4063-4069.
[25] Zeng, X. L., Martin, P., & Boulakia, G. (2004). Tones and intonation in declarative and interrogative sentences in Mandarin. In *International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages*.