L2 EXPERIENCE MODULATES LEARNERS' USE OF CUES IN THE PERCEPTION OF L3 TONES

Zhen Qin, Allard Jongman

Department of Linguistics, University of Kansas, United States
qinzhenquentin2@ku.edu, ajongman@ku.edu

ABSTRACT

It is unclear what roles native language (L1) and second language (L2) play in the perception of lexical tones in a third language (L3). Listeners with different language backgrounds use different F0 cues in tonal perception. While English listeners use F0 height, Mandarin listeners rely more on F0 direction. The present study addresses whether knowledge of Mandarin as an L2 results in listeners' reliance on F0 direction in their perception of L3 (Cantonese) tones. 15 English-speaking L2 learners of Mandarin were compared to 15 monolingual English speakers and 15 native Mandarin speakers as control groups. All groups discriminated Cantonese tones either by distinguishing a contour tone from a level tone or a level tone from another level tone. The L2 learners patterned differently from the control groups by using both F0 height and direction. Hence, L2 experience as well as L1 experience was found to modulate perception of L3 tones.

Keywords: Speech perception, Third language acquisition, Chinese tones.

1. INTRODUCTION

The present study addresses the role of linguistic experience, in addition to acoustic factors, in the perception of tones in an unfamiliar language. In this study, we introduce L2 learners to an unfamiliar L3.

Linguistic experience is defined here as the influence of previously acquired languages (i.e., L1 and L2). L1 has been suggested to play a crucial role in predicting development and ultimate attainment of L2 acquisition [6, 10]. While only L1 could be the source of potential transfer in L2 acquisition, either L1 or L2 or both could be sources of transfer in the initial stage of L3 acquisition [3, 11].

In terms of acoustics, the larger the acoustic difference between two sounds, the easier their discrimination will be, regardless of which languages have been previously acquired.

The current study addresses the question of what role L1 and L2 play in L3 acquisition by testing lexical tones. Lexical tone is an informative linguistic structure for testing the influence of L1 and L2 in L3 acquisition because L2 can be shown to be the source of transfer to L3 if lexical tones are present in L2, but not L1 [11]. The present study recruited English-speaking learners with Mandarin as L2 and introduced them to Cantonese as L3. Given the greater structural similarity between L2 and L3 compared with L1 and L3 in terms of lexical tones, it is revealing to examine whether L2 experience overrides L1 influence and modulates learners' use of cues in their perception of L3 tones.

Speakers of tone languages use F0 cues to distinguish lexical meanings whereas speakers of non-tone languages typically use F0 cues to convey pragmatic meaning in intonation [8]. Previous studies found that English listeners attend to F0 height as a cue and process lexical tones as non-linguistic units because there is no lexical tone in English. In contrast, Mandarin listeners are sensitive to F0 direction as a cue and process lexical tones as linguistic units, because Mandarin is a tone language [2, 4, 5]. In addition to the contour tone contrasts found in Mandarin, Cantonese has level tone contrasts, so listeners need to attend to both F0 height and F0 direction to identify Cantonese tones. Therefore, Cantonese tones are used as stimuli to examine whether English-speaking L2 learners of Mandarin perceive Cantonese tones using F0 height like English listeners or F0 direction like Mandarin listeners.

2. METHODS

2.1. Participants

15 English-speaking learners of Mandarin were recruited as our target group. All of them were college students studying Chinese at the University of Kansas. Their average age of acquisition was 21.5 years (SD: 3.7) and they had received an average of 2.17 years (SD: 0.8) of Chinese instruction. Eleven of them self-reported their proficiency of Chinese as intermediate. Moreover, 15 English native speakers with no exposure to any tone languages (i.e., English monolinguals) and 15 Mandarin native speakers with no background in other tone languages were recruited as control groups. A questionnaire was

used to collect information about the participants' language backgrounds. They reported normal hearing and no history of speech or language disorders. The English-speaking L2 learners of Mandarin and the English monolinguals did not differ in their musical experience. All participants were paid for their participation in the experiment.

2.2. Stimuli

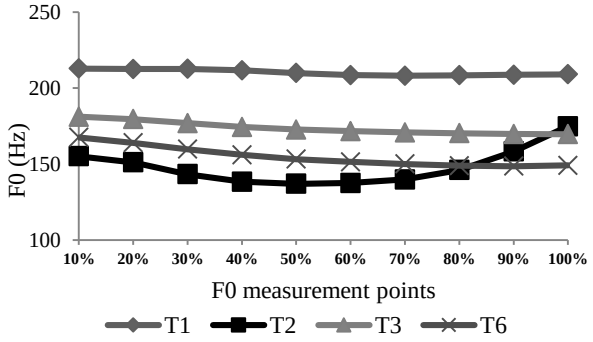
Level and contour Cantonese tones were used as stimuli, allowing for a test of both F0 height and F0 direction. All tones were carried by two syllables, [jɐu] and [se], to counterbalance the effect of voicing of initial consonants on F0 contours [7].

One female native speaker of Hong Kong Cantonese produced all the stimuli. The recording was carried out in a quiet room. The speaker read a randomized list of words embedded in a carrier phrase *ngo5duk6_ zi6* "I read the word___" three times at a normal speech rate. The carrier sentence was selected to avoid the final lengthening effect on the tone stimuli. Tone sandhi does not apply in this sentence. The recording was done directly on disk using a sampling rate of 44.1 kHz. Three repetitions of each target tone carried by each syllable were excised, which resulted in 24 tokens (4 tones × 2 syllables × 3 repetitions) in total. Intensity of all the stimuli was normalized at 70 dB. Duration of the stimuli was normalized within each syllable by using the average values of [jɐu] (345 ms) and [se] (488 ms), respectively.

Four tones (T1, T2, T3, T6) were chosen from the six Cantonese tones as our target stimuli. The other two tones (T4 and T5) were excluded because they are often confused with other tones even by native listeners [9]. Ten temporally equidistant F0 measurements were taken for each stimulus using ProsodyPro [13] and averaged across tokens, as shown in Figure 1.

In order to test listeners' sensitivity to F0 height and F0 direction, the four tones (T1, T2, T3, T6) were then paired into four tonal contrasts; two pairings were level-level (T1-T6, T3-T6) and the other two were contour-level (T2-T1, T2-T6). The level-level tone pairs in which two tones in each pair differ in F0 height were used to test listeners' sensitivity to F0 height. While T1-T6 has a larger acoustic difference (55 Hz) in F0 height, T3-T6 has a smaller acoustic difference (19 Hz). Hence, T1-T6 and T3-T6 are acoustically easy and hard level-level tone pairs, respectively. The contour-level tone pairs in which two tones in each pair differ primarily in F0 direction were used to assess listeners' use of F0 direction. While T2-T1 has a larger acoustic difference (62 Hz), T2-T6 has a smaller acoustic difference (7 Hz). Hence, T2-T1 and T2-T6 are acoustically easy and hard contour-level tone pairs, respectively.

Figure 1: F0 measurements of our stimuli collapsed across tokens and syllables.



2.3. Procedures

An AX forced-choice tone discrimination task was conducted for the three groups of listeners. Two types of tone pairs were used as stimuli, AA pairs (pairs with the same tone) as fillers and AB pairs (pairs with different tones) as target stimuli. In total, 144 AB pairs (4 pairs × 2 orders × 2 syllables × 3 tokens× 3 tokens) and 144 AA pairs (4 pairs × 2 syllables × 3 tokens × 3 tokens × 2 repetitions) were created, respectively, to counterbalance the number of targets and fillers. The presentation order of each AB pair was also counter-balanced. The stimuli were presented in random order to the participants.

The stimuli were presented to participants via stereo headphones with the volume adjusted to a comfortable level. The participants were told that they would hear pairs of sounds from a new language that they were not familiar with. Their task was to judge whether the two tones in each pair were "the same" or "different" as quickly and accurately as possible by pressing one of two buttons, labelled "same" or "different", respectively, on the keyboard.

The experiment was conducted using Paradigm software. The inter-stimulus interval (ISI) was 500 ms to induce a phonetic mode of processing. Invalid responses were excluded from analysis. The instructions were given orally as well as displayed visually on a computer screen before the test. No feedback was given. Before the actual experiment, the participants familiarized themselves with 12 practice trials randomly selected from the test stimuli.

2.4. Data analysis

Logit mixed-effects models were conducted on the participants' accuracy. The models were fitted in R, using the lmer( ) function from the lme4 package for mixed-effects models [1].

A first model analyzed the accuracy of each of the three groups in perceiving contour-level vs. level-level tone pairs, respectively, with Tone pair (contour-level vs. level-level), Group (L2 Learners vs. English monolinguals and Mandarin listeners), and the interaction between the two as fixed variables. In this model, the contour-level tone pairs and the L2 learners group were used as baselines. A second model analyzed only the two monolingual groups' accuracies, with Group (English vs. Mandarin) as a fixed variable. The baseline for Group was English.

A third model analyzed the accuracy of each of the three groups in perceiving the two different contour-level tone pairs, with Tone pair (T2-T1 vs. T2-T6), Group (L2 Learners vs. English monolinguals and Mandarin listeners), and the interaction between the two as fixed variables. In this model, T2-T1 and the L2 learners group were used as baselines. A fourth model analyzed only the two monolingual groups' accuracies, with Group (English vs. Mandarin) as a fixed variable. The baseline for Group was English. A fifth model analyzed the accuracy of each of the three groups in perceiving the two different level-level tone pairs, with Tone pair (T1-T6 vs. T3-T6), Group (L2 Learners vs. English monolingual and Mandarin monolingual), and the interaction between the two as fixed variables. In this model, T1-T3 and the L2 learners group were used as baselines. A sixth model analyzed only the two monolingual groups' accuracies, with Group (English vs. Mandarin) as a fixed variable. The baseline for Group was English.

In all the models, participant and item were crossed random variables.

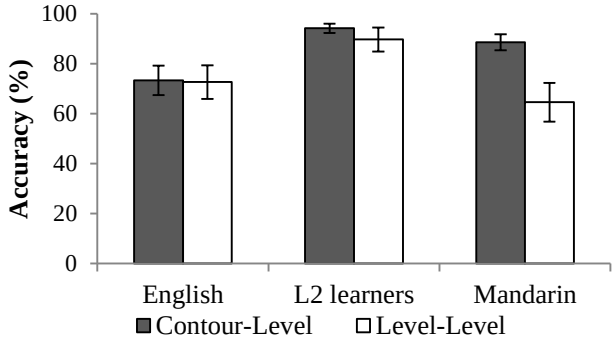
3. RESULTS

3.1. Contour-Level vs. Level-Level tone pairs

Accuracy of the three groups in discriminating contour-level and level-level tone pairs is illustrated in Figure 2.

A first logit mixed-effects model, focusing on the L2 learners, was conducted on the participants' accuracy. The results, as illustrated in Table 1, showed a significant Tone pair type x Group interaction for both the English group and the Mandarin group. This indicates that the L2 learners differed from the English and the Mandarin listeners in their discrimination of contour-level and level-level tone pairs. Post-hoc analysis showed that the L2 learners were better than English listeners [$z(2918)=12.4$, $p<.001$] and Mandarin listeners [$z(2918)=4.6$, $p<.001$] at using F0 direction to perceive the contour-level tone pairs. They also did better than English listeners [$z(2904)=10.4$, $p<.001$] and Mandarin listeners [$z(2904)=13.9$, $p<.001$] at using F0 height to perceive the level-level tone pairs.

Figure 2: Discrimination accuracy (and standard errors) of Contour-Level vs. Level-Level tone pairs by monolingual English listeners, L2 learners of Mandarin and monolingual Mandarin listeners.



A second logit mixed-effects model, focusing on the two monolingual groups, also showed a significant Tone pair type x Group interaction for the Mandarin listeners. Post-hoc analysis showed that between the two groups, the Mandarin listeners did better [$z(1844)=9.2$, $p<.001$] at using F0 direction to distinguish contour-level tone pairs while the English listeners did better [$z(1830)=4.3$, $p<.001$] at using F0 height to distinguish level-level tone pairs.

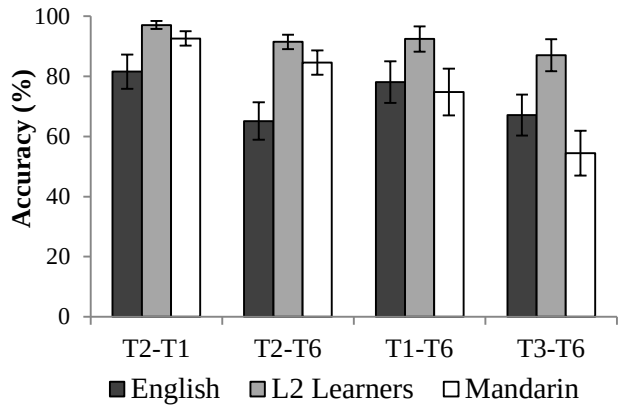
Table 1: Logit mixed-effects models on all participants' overall accuracy.

L2 learners' data (*df*=6125)				
Variable	**Est.**	**SE**	***z***	***p***
Group (E)	–1.9	0.15	–13	<.001
Group (M)	–0.8	0.17	–5	<.001
Tone Pair Type	–0.6	0.17	–4	<.001
Group (E) x Tone Pair Type	0.6	0.20	3	.002
Group (M) x Tone Pair Type	–1.0	0.21	–5	<.001
English vs. Mandarin (*df*=3977)				
Variable	**Est.**	**SE**	***z***	***p***
Group	1.2	0.12	9.2	<.001
Tone Pair Type	–0.1	0.10	–0.4	<.001
Group x Tone Pair Type	–1.6	0.16	–9.9	<.001

3.2. Individual tone pairs

Given the acoustically easy versus hard tone pairs, each tone pair was analyzed separately for each group, as illustrated in Figure 3.

Figure 3: Discrimination accuracy (and standard errors) of the four tone pairs for English monolinguals, L2 learners of Mandarin and Mandarin listeners.



A third logit mixed-effects model was conducted on the participants' accuracy on the two contour-level tone pairs. The results showed a significant effect of Tone pair [$z(3018)=-3.9$, $p<.001$], but no significant Tone pair x Group interaction was found. Similarly, a fourth logit mixed-effects model, focusing on the two monolingual groups, showed a significant effect of Tone pair [$z(1844) =-6.5$, $p<.001$], but no significant Tone pair x Group interaction was found. This indicates that T2-T1 was perceived more accurately than T2-T6 by all three groups, and the groups did not differ in their discrimination of the two contour-level tone pairs.

A fifth logit mixed-effects model was conducted on the participants' accuracy on the two level-level tone pairs. The results showed a significant effect of Tone pair [$z(2904) =-3.0$, $p=.003$], but no significant Tone pair x Group interaction was found. A sixth logit mixed-effects model, focusing on the two monolingual groups, showed a significant effect of Tone pair [$z(1830) =-4.3$, $p<.001$], as well as a significant Tone pair x Group interaction [$z(1830) =-2.1$, $p=.04$]. The results showed that L2 learners of Mandarin did not differ from the other two groups in their discrimination of the two level-level tone pairs. Although the English and Mandarin listeners differed regarding the two level-level tone pairs, a post-hoc analysis showed that both the English group [$z(771)=-4.8$, $p<.001$] and the Mandarin group [$z(765)=-8.3$, $p< 001$] showed a significant effect of Tone pair. This indicates that T1-T6 was perceived more accurately than T3-T6 by all the groups.

4. DISCUSSION

This study investigated how linguistic experience and acoustic factors affect perception of L3 tones. Overall, our results showed that both factors had an impact on the perception of L3 tones.

In terms of linguistic experience, we found that L2 experience did modulate listeners' use of F0 cues in the perception of L3 tones. Our results showed that the L2 learners had a pattern which was different from either that of the English monolinguals or Mandarin listeners in their discrimination of Cantonese tones. The L2 learners performed better than the English monolinguals and Mandarin listeners at discriminating the contour-level tone pairs. This suggests that L2 experience increased L2 learners' sensitivity to F0 direction in the perception of L3 tones. Additionally, the results showed that the L2 learners, like the English monolinguals, were better than the Mandarin listeners in using F0 height to discriminate the level-level tone pairs. This suggests that L1 experience still played a role. Thus L2 influence did not override the L1 influence in the perception of L3 tones.

In terms of acoustics, we found that all listeners showed higher accuracy on acoustically easy pairs than acoustically hard pairs within the level-level as well as contour-level tone pairs. This suggests that acoustic properties of the tonal stimuli affected listeners' perception regardless of their language backgrounds. Because the acoustically easy versus hard pair within each comparison in our study is based on F0 height differences, it would be interesting to include an easy versus hard pair involving contour differences (e.g., big rise-level vs. small rise-level) to explore speaker group and tone pair interactions in a further study.

Overall, the present study found the effect of L2 influence given the similarity between L2 and L3 in terms of lexical tones. Converging with studies of L3 morpho-syntax acquisition [e.g., 12], our study contributes to the fields of both L2 and L3 acquisition by providing evidence supporting the influence of L2 on the acquisition of L3 in the speech domain.

5. CONCLUSION

Both linguistic experience and acoustic factors were found to affect the perception of L3 tones in this study. First, the acoustics of the tones affected all listeners' discrimination. Second, in the perception of Cantonese tones, the L2 learners of Mandarin showed a pattern different from either that of English monolinguals or Mandarin listeners in their discrimination of contour-level and level-level tone pairs. Hence, L2 experience as well as L1 experience was found to modulate perception of L3 tones.

6. REFERENCES

[1] Baayen, R. H. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics*. Cambridge: Cambridge University Press.

[2] Burnham, D., Mattock, K. 2007. The perception of tones and phones. In: O.-S Bohn & M. J. Munro (eds), *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*. Amsterdam: John Benjamins, 259-280.

[3] Flynn, S., Foley, C., Vinnitskaya, I. 2004. The cumulative enhancement model for language acquisition: Comparing adults and children's patterns of development in first, second and third language acquisition of relative clauses. *International Journal of Multilingualism* 1, 3–16.

[4] Francis, A. L., Ciocca, V., Ma, L., Fenn, K. 2008. Perceptual learning of Cantonese lexical tones by tone and non-tone language listeners. *Journal of Phonetics* 36, 268-294.

[5] Gandour, J. T. 1983. Tone perception in far Eastern languages. *Journal of Phonetics* 11, 149–175.

[6] Gass, S. M., Selinker, L. 1983. *Language Transfer in Language Learning.* Rowley, MA: Newbury House.

[7] Hombert, J. M. 1978. Consonant types, vowel quality, and tone. In: Fromkin, V. (eds), *Tone: A linguistic survey*. New York: Academic Press, 77-111.

[8] Ladd, D. R. 2008. *Intonational phonology*. Cambridge: Cambridge University Press.

[9] Mok, P., Zuo, D., Wong, P. 2013 Production and perception of a sound change in progress: tone merging in Hong Kong Cantonese. *Language Variation and Change* 25, 341-370.

[10] Odlin, T. 1989. *Language Transfer. Cross-linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.

[11] Rothman J. 2011. L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research* 27, 107–27.

[12] Rothman J., Cabrelli Amaro. J. 2010. What variables condition syntactic transfer?: A look at the L3 initial state. *Second Language Research* 26, 189–218.

[13] Xu, Y. 2013. ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis. *Proc. of Tools and Resources for the Analysis of Speech Prosody* (TRASP 2013) Aix-en-Provence, 7-10.