# IMITATION OF MANDARIN TONES BY L2 MANDARIN LEARNERS

Kiranpreet Nara[1]

[1]University of Toronto
kiranpreet.nara@mail.utoronto.ca

## ABSTRACT

Speakers have been shown to take on and retain phonetic properties (VOT, vowel quality, f0) of others in social and laboratory contexts. As much of the imitation work on f0 has investigated intonation, a question that remains is the extent to which imitation can be observed for lexical tones and how long lasting those effects are. In the current study, L2 learners of Mandarin were asked to imitate tones produced by native Mandarin speakers. The aims were to determine a) if imitation occurred, b) if the effects were retained into the post-imitation task, and c) if the effects were generalized to novel stimuli in the post-imitation task. The results of the Smoothing Spline ANOVA analyses showed that female L2 learners imitated the native speakers while the male L2 learners did not. Imitation effects were retained and also generalized to novel stimuli by the female L2 learners during the post-imitation task.

**Keywords**: Tone, Mandarin, L2 learners, Imitation

## 1. INTRODUCTION

Imitation of linguistic characteristics has been observed in both social [14, 19] and laboratory contexts [18, 23]. In particular, speakers reliably imitate various phonetic parameters such as vowel quality [1, 7], voice onset time (VOT) [23, 18], and fundamental frequency (f0) [2, 4, 11]. Researchers have studied general imitation of words in which participants heard words produced by a model speaker and repeated them [10, 17, 19]. The previously heard words were judged to be similar to the model's speech compared to words that were not heard, indicating that the participants imitated some phonetic properties of the model speaker. More specifically, there have been studies of imitation at the segmental level, for example, Delvaux and Soquet investigated vowel imitation in two different regiolects of French [7]. During the imitation task, the participants of one regiolect were exposed to those of the other regiolect via loud speakers and the researchers found that the vowel qualities of the participants changed to become more like the vowels of the talker heard through the loud speaker. Though the effects of imitation were retained to some extent during the post-imitation task, the participants had started returning to their natural vowel productions.

Researchers have investigated VOT in imitation experiments by presenting artificially lengthened VOT of American English voiceless stops to participants [23]. The participants lengthened their VOT to match the model speaker, providing evidence for the imitation of subsegmental phonetic properties. Nielsen [18] also conducted an imitation study utilizing artificially lengthened VOT (by 40ms) which found that the imitation effects were retained into the post-imitation task and generalized to novel stimuli that were not heard during the imitation task. However, the participants did not completely imitate the model speakers in this study as they increased their VOT by only about 10 ms.

Studies have also found evidence for f0 imitation at the intonation level; participants deviate from their own baseline average pitch toward that of the speaker to which they were exposed during the imitation task [22, 9, 11, 2]. At least two studies have looked at lexical tone imitation by speakers of non-tonal languages [3, 13]. In one study [3], native English speakers without any knowledge of a tonal language were asked to repeat mono- and trisyllabic words, /ra/ and /ra.ra.ra/, spoken by a native Mandarin speaker. The acoustic analysis showed native-like tone contours for the English speakers during imitation, however, the pitch range of the English speakers was not as wide as that of the Mandarin speakers. In the other study [13], English speakers, who had been learning Mandarin as an L2 for more than two years, were trained to produce tones in non-words using imitation methodology. Again, the L2 learners were able to imitate the tones in non-words.

The above-mentioned findings suggest that imitation occurs, though perhaps not completely [18], at multiple linguistic levels: word, segmental, subsegmental, and suprasegmental. Studies have also shown that imitation effects can be retained and generalized to novel stimuli during a post-task [7, 18]. For f0 imitation, retention and generalization of imitation effects has not been studied and in lexical tone studies, stimuli have been non-words.

There is also evidence that training, whether imitation or otherwise, can help learners produce and perceive sound contrasts of foreign languages. Training has focused on vowels [21], VOT [16], and tones [15], and has led to improvements in the

participants' perception and production of L2 sounds. The current study used the imitation methodology of a baseline task, imitation task, and a post-imitation task to determine if recent L2 learners of Mandarin are able to imitate Mandarin tones in real Mandarin words. The study also examined retention and generalization of imitation effects.

This phenomenon of convergence in speech has been accounted for with memory models of perception and production, such as Episodic Theory [10, 23], and is also considered as a tool to manage social distance between the speaker and their conversation partner, as in the Communication Accommodation Theory [8]. Speakers may converge with their conversation partners to decrease the social distance, maintain their regular distance, or diverge to increase the social distance between them [5]. Studies have looked at and found social variables (such as showing an image of the model talker, the attractiveness of the model, or if the participant liked the model) that contribute to whether and how much participants imitate a model speaker [2]. The L2 learners of the current experiment were aware of the following variables: gender of the model speaker, that the models were native Mandarin speakers, and that they were hearing speech in their L2, Mandarin Chinese. Since the L2 learners were studying Mandarin at the time of the experiment, that might be a motivation for them to imitate the native speakers during the imitation phase of the experiment to sound more like the native speakers.

## 2. METHOD

### 2.1. Participants

Model tokens were produced by two native Mandarin speakers, one female (age 23) and one male (age 28), who were recruited from the University of Toronto. Both native speakers (NS) were born and raised in China, learned Mandarin as their first language, and spoke English as a second language.

The L2 learners had been studying Mandarin for less than eight months and they did not speak or know any other tonal language. They were recruited from an Introductory Mandarin course taught at the University of Toronto. Ten Mandarin learners (six female) took part in the study. The mean age of female participants was 19.5 and for the male speakers, the mean was 20.25. All participants provided written informed consent prior to taking part in the experiment.

### 2.2. Stimuli

The stimuli were real Mandarin words in CV syllables where the vowel was always /a/. The wordlists were slightly different depending on the phase of the experiment (baseline, imitation, and post-test), as summarized in figure 1. The consonants that were chosen for the baseline phase and imitation phase list were the following: /p f m t tʰ ʈʂʰ/. All four Mandarin tones (tone 1 'level', tone 2 'rising', tone 3 'falling-rising/dipping', and tone 4 'falling') occurred on a combination of each of these consonants with the vowel /a/. A different wordlist, called the 'complete wordlist,' was used during the post-test phase. The complete wordlist contained all the words from the baseline list as well as 16 novel words beginning with consonants /pʰ s ʂ ʈʂ k x/.

### 2.3. Procedure

The two native Mandarin speakers (NS) served as model talkers and read the complete wordlist. The words, written in Pinyin along with the gloss, were presented on a computer screen one at a time to NS using Microsoft PowerPoint.

**Figure 1**: Schematic of the experiment and stimuli.

| | |
|---|---|
| **Baseline:**<br>24 CV words using<br>/p f m t tʰ ʈʂʰ/ + /a/ | |
| **Imitation task:**<br>Baseline words pseudo-randomly ordered and alternating between both female and male NS | X 2 |
| **Post-test:**<br>Baseline list + 16 novel words using<br>/pʰ s ʂ ʈʂ k x/ + /a/ | |

L2 learners completed three phases in a single experimental session, see figure 1 for a schematic of the experiment. The first phase was the baseline task and involved reading the baseline wordlist which contained 24 Mandarin words written in Pinyin. The second phase of the experiment was the imitation task in which the L2 learners heard words from the baseline wordlist as spoken by the two native speakers and were asked to repeat the words in their normal speaking voice immediately after they heard the native speaker. The participants were not explicitly asked to imitate the talker. Stimuli were presented in a pseudo-randomized order and the female and the male speakers alternated so that the L2 learner never heard two or more words spoken in a row by the female or male NS. The breakdown of the words for each participant in the imitation experiment was as follows: 24 words * 2 repetitions of each word per NS * 2 NS = 96 tokens. Finally, the third phase of

the experiment was the post-task in which L2 participants read the complete wordlist which had 24 words from the baseline list and 16 novel words beginning with different consonants. The participants had a 10-minute break before moving on to phase three of the experiment. The participants were debriefed and compensated at the end of the experiment.

All recordings were made on a SoundDevices 722 digital audio recorder and DPA 4011 microphone using a sampling frequency of 44.1 kHz. Sennheiser HD280 headphones were used during the imitation session to play the native speakers' tokens. All recordings took place in the sound attenuating booth in the Phonetics Lab at the University of Toronto.

## 3. RESULTS

Smoothing Spline ANOVA (SS-ANOVA) analyses were conducted for each tone to examine differences between the L2 and native speakers of the same gender. The smoothing splines function in the SS-ANOVA connects discrete data points to create curves or 'splines' and calculates 95% Bayesian confidence intervals, which are represented as dotted-line curves above and below the main spline. Two splines are deemed significantly different from one another if the curves of their confidence intervals do not intersect [6].

These analyses were done in R using the *gss* and *ggplot2* packages [20, 12, 24]. The legend of the SS-ANOVA charts is as follows: 'P1' refers to phase one (baseline task), 'P2' is phase two (the imitation task) of the experiment. 'P3_Old' refers to the third phase (post-task) of the experiment and only contains the data for the words used duing the baseline and imitation tasks. 'P3_Novel' category contains only the novel words that were produced by the L2 speakers during phase three. Finally, 'NS_F/NS_M' refers to the average f0 of the one of the two native speakers of Mandarin. The x-axis represents the eight intervals at which f0 was measured in the vowel duration.

### 3.1. Female L2 learners

Tone 1 pitch for the female NS was around 200 Hz and for the female L2 learners, it was around 260 Hz in phase one, see figure 2. During the imitation phase, the female L2 learners lowered their pitch, indicating that imitation had occurred. The imitation effects were retained and generalized into phase three (P3_Old and P3_Novel) of the experiment, though the L2 learners had started going back to their original tone 1 pitch produced during phase one. A similar pattern was observed for Tone 2 for which the pitch range of the female L2 learners lowered to match the average pitch of the female NS during the imitation phase. The L2 learners imitated the average pitch of NS. The imitation effects for tone 2 were retained and generalized to novel stimuli in the post-task. For tone 3, again, the female L2 learners showed deviation from their baseline pitch during the imitation phase as well as generalization of the imitation effects to novel stimuli. Finally, tone 4 was also imitated during the phase two of the experiment but the effects were not retained or generalized during the third phase of the experiment, as the splines for phase three overlap with the phase one spline, see figure 3. The tonal contours of all four Mandarin tones for female L2 learners were like those of the female NS speaker.

**Figure 2**: Tone 1 of female L2 compared with female NS.

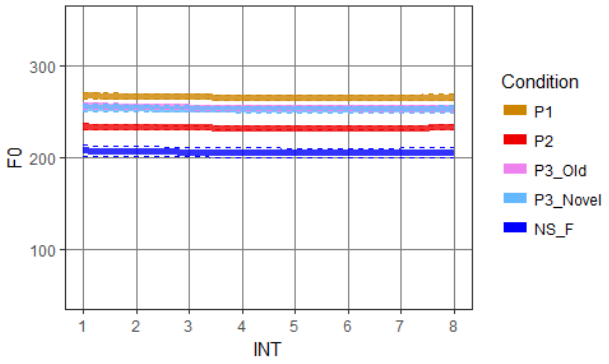
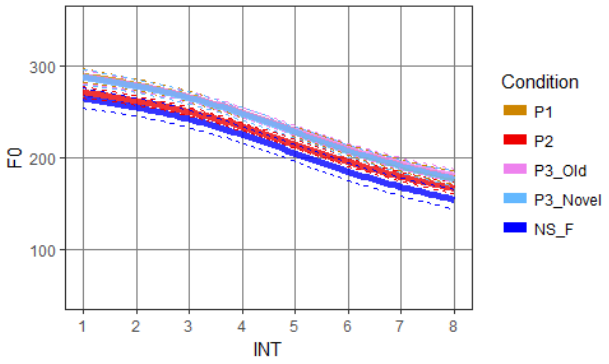



**Figure 3**: Tone 4 of female L2 compared with female NS.

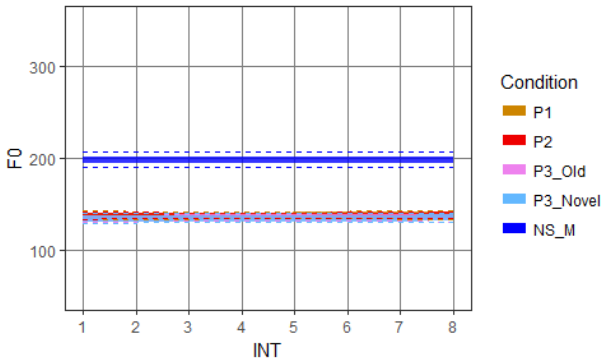



### 3.2. Male L2 learners

Tone 1 of male L2 learners did not deviate from its pitch around 140Hz to become more like the native speaker, see figure 4. The splines for the three phases of the experiment for L2 learners overlap suggesting no imitation effects. Tones 2 also followed the same pattern in which the male L2 learners did not imitate the native speaker. Tone 3 was produced with similar pitch range and contour for both L2 and NS speakers. The L2 learners did imitate the pitch height of tone 4 in its second half, however, the effects disappeared in

the third phase of the experiment. The tonal contours of the male L2 learners were like those of the native speaker but the f0 range for male learners' tones 2 (90-130 Hz) and 4 (75-150 Hz) were not as large as the native speaker's tone 2 (105-180 Hz) and tone 4 (125-240 Hz) range, respectively.

**Figure 4**: Tone 1 of male L2 compared with male NS.




## 4. DISCUSSION

Female L2 learners of Mandarin diverged from their natural lexical tone height to become more like the native speakers they heard during the imitation task for all four tones. Male speakers, however, only imitated part of tone 4 pitch height but not tone 1 or 2. Tone 3 was already similar for male L2 and NS speakers. With the exception of L2 male speakers' rising and falling tones, the tone contours and pitch range of the L2 learners were already native-like. This suggests that after about eight months of learning, the L2 speakers had created native-like tonal categories with the appropriate contours and pitch range. The L2 learners did not completely imitate the average pitch of the native speakers, but in tones 1, 2, and 3 of the female L2 learners and tone 4 of the male L2 learners, the difference between the baseline and the imitation phase was significantly different. The small degree of imitation was expected for the current experiment and also observed in Nielsen [18]. These results are in line with previous laboratory studies which have shown imitation of average pitch height at the level of intonation or overall conversation pitch [9, 11, 2] as well as lexical tone [13, 3]. Another similarity to Bent's [3] study was that the L2 learners produced native-like contours, however, unlike in Bent, the learners in the current study did not have smaller pitch ranges, except for tones 2 and 4 of L2 males, than the native speakers perhaps because Bent's English speakers did not have any experience with Mandarin whereas the participants in the current study had been studying it for eight months.

These results also replicate the retention and generalization of imitated features from previous non-f0 imitation studies such as vowel quality [7] and VOT [23, 18]. The effects of imitation were retained to some extent into the post-test. Retention was also observed in Delvaux and Soquet [7] and Nielsen [18]. Delvaux and Soquet found more robust imitation during the imitation task than during the post-task, as was also the case in the present study. This suggests that imitation effects are stored for some time in memory but do start to fade after the imitation task, which is not surprising as the imitation task in the current study took less than ten minutes to complete and was not very linguistically meaningful because the target words were read in list form. The female L2 speakers also generalized the imitation effects to stimuli beginning with novel consonants during the post-task and these findings were in line with Nielsen [18].

The lack of convergence on the part of the males could be that the male native speaker's natural pitch range was unusually wide (between 90 Hz and 250 Hz) compared to the L2 males (between 75 Hz and 150 Hz). Either the male learners did not want to imitate the native speaker's high pitch, or they could not physiologically widen their pitch range. The male native speaker's pitch might have contradicted with the participants' beliefs regarding typical male voice pitch and prevented them from converging to him. This would be in accordance with Communication Accommodation Theory because gender and social roles play a role in accommodation [8, 17]. Female L2 learners, on the other hand, might have imitated the native speakers to sound like the native speakers who have more experience with the language, and given that the L2 learners were taking the Mandarin course at the time of the experiment, it might have been another motivation for the females to converge to the native speakers.

In conclusion, the results of the current study provide evidence for imitation of lexical tones by L2 learners of Mandarin. Not only did imitation occur, but the effects were retained, though they had started diminishing, in memory during the post-task but also applied to novel stimuli, indicating that a short imitation session can be used to alter aspects of speakers' lexical tone productions.

## 5. REFERENCES

[1] Babel, M. 2009. Selective vowel imitation in spontaneous phonetic accommodation. UC Berkeley *Phonology Lab Annual Reports* 5.

[2] Babel, M., & Bulatov, D. 2012. The role of fundamental frequency in phonetic accommodation. *Language and speech* 55, 231–248.

[3] Bent, T. 2005. Perception and production of non-native prosodic categories. Unpublished Ph. D. thesis,

Department of Linguistics, Northwestern University, Evanston, IL.
[4] Bosshardt, H. G., Sappok, C., Knipschild, M., & Hölscher, C. 1997. Spontaneous imitation of fundamental frequency and speech rate by nonstutterers and stutterers. *Journal of psycholinguistic research* 26, 425–448.
[5] Bourhis, R. Y., & Giles, H. 1977. The language of intergroup distinctiveness. *Language, ethnicity, and intergroup relations* 13, 119–135.
[6] Davidson, L. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance a. *The Journal of the Acoustical Society of America* 120, 407–415.
[7] Delvaux, V., & Soquet, A. 2007. The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica* 64, 145–173.
[8] Giles, H., & Coupland, N. 1991. *Language: Contexts and consequences.* Thomson Brooks/Cole Publishing Co.
[9] Goldinger, S. D. 1997. Words and voices: Perception and production in an episodic lexicon. *Talker variability in speech processing* 33–66.
[10] Goldinger, S. D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological review* 105, 251.
[11] Gregory, S. W., Dagan, K., & Webster, S. 1997. Evaluating the relation of vocal accommodation in conversation partners' fundamental frequencies to perceptions of communication quality. *Journal of Nonverbal Behavior* 21, 23–43.
[12] Gu, C. 2014. Smoothing spline ANOVA models: R package gss. *Journal of Statistical Software* 58, 1-25.
[13] Hao, Y. C., & de Jong, K. 2016. Imitation of second language sounds in relation to L2 perception and production. *Journal of Phonetics* 54, 151–168.
[14] Kim, M., Horton, W. S., & Bradlow, A. R. 2011. Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory phonology* 2, 125–156.
[15] Leather, J. 1990. Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. *New sounds* 90, 72–97.
[16] McClaskey, C. L., Pisoni, D. B., & Carrell, T. D. 1983. Transfer of training of a new linguistic contrast in voicing. *Perception & psychophysics* 34, 323–330.
[17] Namy, L. L., Nygaard, L. C., & Sauerteig, D. 2002. Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology* 21, 422–432.
[18] Nielsen, K. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics*, 39, 132–142.
[19] Pardo, J. S. 2006. On phonetic convergence during conversational interaction. *J. Acoust. Soc. Am.* 119, 2382–2393.
[20] R: Core Team. 2016. R: A Language and Environment for Statistical Computing, Retrieved from, <https://www.R-project.org>.
[21] Rojczyk, A. 2012. Phonetic imitation of L2 vowels in a rapid shadowing task. *Proc. 4th Pronunciation in Second Language Learning and Teaching Conference* 66–76.
[22] Scherer, K. R. 1982. Methods of research on vocal communication: Paradigms and parameters. *Handbook of methods in nonverbal behavior research* 136–198.
[23] Shockley, K., Sabadini, L., & Fowler, C. A. 2004. Imitation in shadowing words. *Perception & Psychophysics* 66, 422–429.
[24] Wickham, H. 2009. ggplot2: Elegant Graphics for Data Analysis Springer-Verlag New York.