

# The Production of Mandarin Coarticulated Tones by Inexperienced and Experienced English Speakers of Mandarin

Yunjuan He, Rairie Wayland  
Program in Linguistics  
University of Florida

## ABSTRACT

This study tested the effects of language learning experience, tonal environment, tonal context and syllable position on American English speakers' ability to produce Mandarin Chinese coarticulated tones in disyllabic words. Two groups of learners with different amount of classroom learning experience participated in the study. The results obtained indicated that (a) American learners with more learning experience were more accurate than less experienced learners in producing coarticulated Mandarin tones; (b) with increased experience, production of coarticulated tone becomes less affected by such phonological and phonetic factors as syllable position and tonal environment and tonal contexts; (c) tonal environment and tonal context only affected tone 1 production; and (d) syllable position affected tone 2 and tone 4 production accuracy.

## 1. Introduction

The effects of language learning experience have been examined on adult acquisition of second language sounds. One of the main research questions asked in these previous studies was whether the performance of experienced learners of the target language (L2) is better than that of the inexperienced learners in their ability to produce L2 sounds. Most previous research on segmental (vowels and consonants) and suprasegmental features (e.g., stress) suggested that the ability to produce L2 sounds may improve with learning experience or as learners' overall experience with the target language increased. (Flege et al. 1995; Flege et al. 1997; He et al, 2008).

The focus of this current study was on the effects of classroom learning experience on the production of Mandarin tones in disyllabic words, hereinafter referred to as "Mandarin coarticulated tones", among native speakers of American English.

## 2. Background

Mandarin has four lexical tones, which can be described using a five-level pitch scale, ranging from 1 to 5, lowest to highest. Tone 1 is a high level tone at pitch level 5. Tone 2 starts at pitch level 3 and rises to level 5. Tone 3 starts at level 2 and drops to level 1 before rising to level 4. Tone 4 is a falling tone which falls from pitch level 5 to level 1

Previous literature (e.g., Wang et al., 1999,

Wayland & Guion 2003, 2004) have shown that the ability to perceive lexical tone distinction improved with experience. Little attention, however, has been paid to the effect of experience on lexical tone production. Only one previous study (Wang et al. 2003) was conducted to examine Mandarin lexical tone production ability among native speakers of English learning Chinese. In this study, it was found that production accuracy of Mandarin tones in monosyllabic words significantly improved after a two-week long perception training. The effectiveness of the training was seen not only on the production of old words included in the training, but also on new words not previously encountered during training.

The goal of this current study was to extend this line of research to the production of coarticulated tones in Mandarin disyllabic words.

## 3. Research questions and hypothesis

The study was guided by two research questions.

*Research question 1:* Does American learners' production of Mandarin coarticulated tones improve with Mandarin learning experience?

Consistent with previous literature, learners with more years of classroom learning experience were expected to produce Mandarin coarticulated tones with a higher accuracy rate than relatively less experienced learners

*Research question 2:* How do such linguistic factors as tonal environment, total context and syllable position affect Mandarin coarticulated tone production accuracy among American English learners?

If the tones of the two syllables in a disyllabic word were both the same, the tonal environment was labeled *identical*, and *non-identical* otherwise. Since all the participants had already learned how to produce isolated Mandarin tones, they were expected to produce these tones more accurately in the identical tonal environment by simply reduplicating the tone in the tonal sequence. Therefore, tones in identical tone environment were expected to be produced with a higher accuracy rate than those in non-identical environment<sup>1</sup>.

---

<sup>1</sup>Tone 3 changed into Tone 2 when a Tone 3 occurs before another Tone 3. Therefore the tonal combination of Tone 3 + Tone 3 was not included in the study

Based on the definitions of tonal context in Xu's (1994) study, *compatible tonal context* means that the F0 value of the offset of the preceding tone and the F0 value of the onset of the following tone are similar, and *conflicting tonal context* means the F0 value of the offset of the preceding tone and that of the onset of the following tone are substantially different. The tones in compatible tonal context are connected smoothly. On the other hand, an F0 gap exists between the tones in the conflicting environment. It was expected that American learners would be able to accurately produce coarticulated tones in a compatible tonal context with a higher degree of accuracy than in a conflicting tonal context.

With regard to the effect of syllable position, there was insufficient information from previous research to generate a firm prediction as to which tone on which syllable would be more accurately produced.

	Identical environment	Nonidentical environment	Compatible context	Conflicting context
T1+T1	X		X	
T1+T2		X		X
T1+T3		X		X
T1+T4		X	X	
T2+T1		X	X	
T2+T2	X			X
T2+T3		X		X
T2+T4		X	X	
T3+T1		X		X
T3+T2		X	X	
T3+T4		X		X
T4+T1		X		X
T4+T2		X	X	
T4+T3		X	X	
T4+T4	X			X

Table 1. Mandarin tonal combinations in different tonal environments and contexts.

### 3 The Experiment

#### Participants

Nine 'inexperienced' American learners of Mandarin with three months of Mandarin learning experience and nine 'experienced' American learners who were in an intermediate level and had studied Mandarin for twelve months participated in the study. All of them were tested on their ability to perceive and produce isolated Mandarin tones and obtained an accuracy score of at least 67%. This step was taken to obtain baseline data on their ability to produce and perceive Mandarin tones in monosyllabic words.

#### Stimuli and Procedure

To minimize the effects of lexical knowledge on production accuracy, the 45 tested words used in the study were Mandarin pseudo (nonsense) words. Participants were asked to read the target words from a wordlist.

To ensure that the participant produced coarticulated tones (as opposed to two isolated tones), they were instructed to produce the two syllables as a single word with no pause between them. The 45 stimulus words were preceded by five

practice words that were not analyzed. Each word was produced three times in random order. The productions were collected using a solid state recorder (Marantz, PMD660/U3B) with a professional microphone (Audio-Technica AT4041 Car) and were transferred to a PC for accuracy judgment.

#### Production Accuracy Judgment

Two native speakers of Mandarin with Beijing accent transcribed the tones produced by participants based on their perception. Answer sheets with target words printed in pinyin without a tonal diacritic were provided. Judges were asked to provide a tonal diacritic corresponding to the tone they hear (Wang 2003). Native speakers' judgment was also checked against the pitch contours generated by the speech analysis software PRAAT.

### 4 Results and Analysis

#### Effects of Experience

As shown in FIG 1, overall, tonal production by experienced American learners of Mandarin were judged to be more accurate by native listeners than that of inexperienced learners [ $F(1,106) = 54.93, p < .001$ ] (79.17% vs. 52.92%). In addition, production accuracy varied significantly across the four tones [ $F(3,104) = 44.999, p < .001$ ], Post hoc analyses (Bonferroni adjusted  $p < .05$ ) showed that Tone 1 production was significantly more accurate than all other tones. In addition, Tone 2 and Tone 4 were significantly more accurate than Tone 3. Production accuracy of Tone 2 and Tone 4 were, however, comparable.

A significant interaction between Group and Tone [ $F(1,104) = 3.702, p = .014$ ] was also obtained. This was due mainly to the fact that experienced learners were more accurate than inexperienced learners in their production of Tone 1 ( $p < .001$ ), Tone 3 ( $p < .001$ ) and Tone 4 ( $p < .001$ ), but not Tone 2 ( $p = .210$ ).

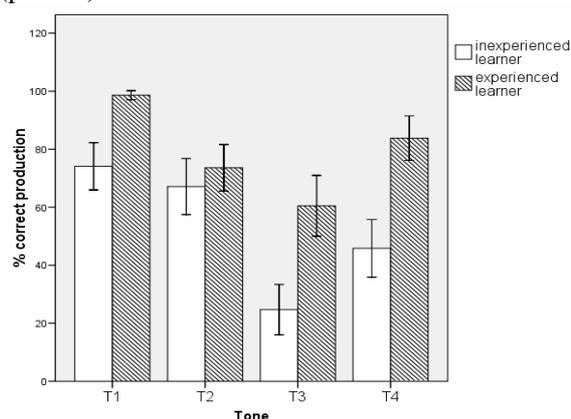


FIG 1. Mean percentage of accurate production of coarticulated tone by inexperienced and experienced American learners of Mandarin. Error bars represent standard error.

#### Effect of tonal environment

FIG 2 shows, overall, American speakers tend to produce tones better in identical tonal environment (76.53%) than in nonidentical tonal environment (68.52%). A three way repeated

ANOVA yielded significant main effects of Tonal Environment [ $F(1, 34) = 4.842, p < .05$ ], Group [ $F(1, 34) = 11.692, p = .002$ ] and Tone [ $F(2, 33) = 20.181, p < .001$ ]. Post hoc analyses (Bonferroni adjusted  $p < .05$ ) showed that Tone 1 was produced more accurately than Tone 2 and Tone 4, but that production accuracy of Tone 2 and Tone 4 was not significantly different.

Additionally, a three-way interaction between Group, Tonal Environment and Tone was significant [ $F(2, 33) = 3.848, p < .05$ ]. Inexperienced learners' production of Tone 1 was found to be significantly more accurate in the identical environment than in the non-identical environment [ $t(1, 70) = 2.737, p < .01$ ]. In contrast, there was no significant difference in the production accuracy between the two tonal environments for Tone 1 among experienced learners [ $t(1, 70) = 1.271, p = .208$ ].

There was no significant difference on accuracy rate of Tone 2 or Tone 4 production between the two tonal environments for either the inexperienced learners [Tone 2:  $t(1, 70) = -.943, p = .349$ ; Tone 4:  $t(1, 70) = 0.270, p = .788$ ] or the experienced learners [Tone 2:  $t(1, 70) = .610, p = .544$ ; Tone 4: [ $t(1, 70) = -1.497, p = .139$ ].

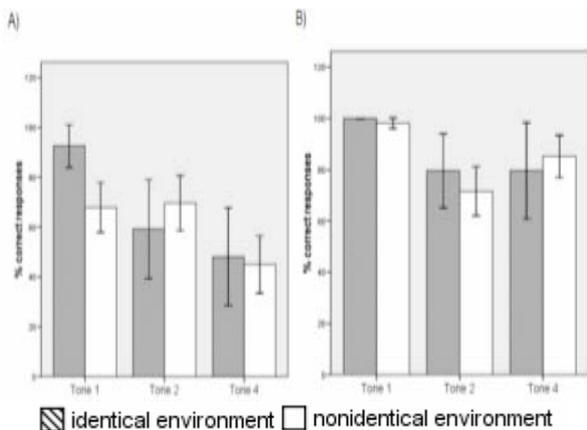


FIG 2. Mean percentage of accurate production of coarticulated tones in identical tone environment vs. non-identical tone environment by (A) Inexperienced and (B) Experienced groups. Error bars represent standard errors.

### Effects of Tonal Context

FIG 3 shows that, overall, American speakers' production of Mandarin coarticulated in both compatible tonal environment (66.67%) and conflict tonal environment (62.03%) was equally accurate. A three way repeated ANOVA yielded a significant main effect of Tone [ $F(3, 32) = 50.928, p < .001$ ] and Group [ $F(1, 32) = 26.263, p < .001$ ]. However, The effect of Tonal Context [ $F(1, 34) = 2.635, p = .114$ ] did not reach significance. Tone was found to interact marginally significantly with Tonal Context [ $F(3, 34) = 2.686, p = .063$ ].

Follow-up tests showed that the production of Tone 1 was nearly significantly more accurate in compatible context than in conflicting context [ $t(1, 142) = 1.719, p = .088$ ]. On the contrary, no significant difference was observed for the production of Tone 2 [ $t(1, 142) = 0.295, p = .768$ ],

Tone 3 [ $t(1, 106) = 0.866, p = .389$ ] and Tone 4 [ $t(1, 142) = 0.000, p = 1$ ].

Further investigation into the effects of tonal context revealed that, for inexperienced learners, Tone 1 was produced with a higher accuracy rate in the compatible context than in the conflicting context, and this difference is marginally statistically significant [ $t(1, 70) = 1.852, p = .068$ ]. For experienced learners, there was virtually no difference in accuracy rate in the two tonal contexts [ $t(1, 70) = .583, p = .562$ ].

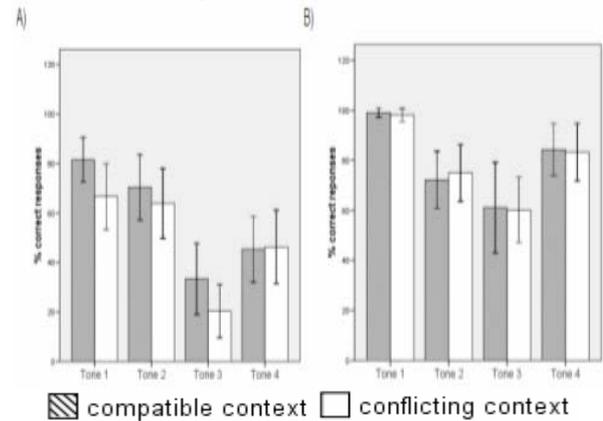


FIG 3. Mean percentage of accurate production of coarticulated tones in compatible vs. conflicting tonal context by (A) Inexperienced and (B) Experienced groups. Error bars represent standard errors.

### Effects of Syllable Position

FIG 4 shows that, overall, American speakers produce tones more accurately in final syllable position (68.53%) than in initial syllable position (62.95%). A three way repeated ANOVA yielded a significant main effect of Syllable Position [ $F(1, 52) = 4.435, p < .04$ ], Group [ $F(1, 52) = 42.715, p < .001$ ] and Tone [ $F(3, 50) = 45.033, p < .001$ ]. Tone was found to interact significantly with Syllable Position [ $F(3, 50) = 45.033, p < .001$ ] and with Group [ $F(3, 50) = 3.636, p < .02$ ].

Follow-up tests on the significant interaction between Tone and Syllable Position showed that the production of Tone 2 was significantly more accurate in final syllable position than in initial syllable position [ $t(1, 142) = -3.870, p < .001$ ]. On the contrary, Tone 4 production was significantly more accurate in initial syllable position than in final syllable position [ $t(1, 142) = 2.001, p < .05$ ]. However, Tone 1 and Tone 3 productions were equally accurate in both syllable positions, [Tone 1,  $t(1, 142) = .300, p = .765$ ; Tone 3,  $t(1, 104) = -1.239, p = .218$ ].

Further investigation into the effects of syllable position showed that Tone 2 was produced significantly more accurately in final syllable position by both inexperienced learners [ $t(1, 70) = -2.49, p = .015$ ] and experienced learners [ $t(1, 70) = -3.05, p = .001$ ]. Inexperienced learners produced Tone 4 significantly better in initial syllable position [ $t(1, 70) = 2.21, p = .030$ ], but syllable position did not affect experienced learners' production accuracy of this tone [ $t(1, 70) = .84, p = .41$ ].

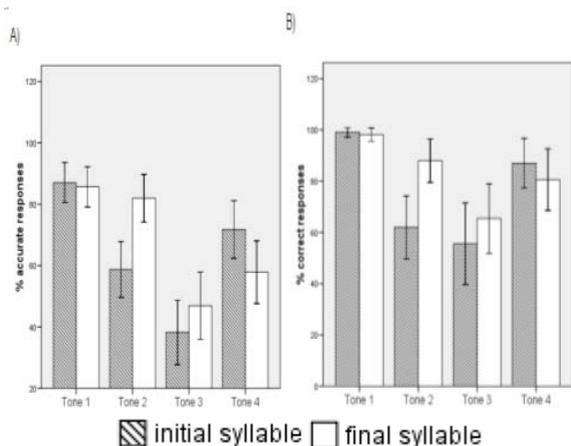


FIG 4. Mean percentage of accurate production of coarticulated tone of coarticulated Mandarin tone in different syllable positions by (A) inexperienced learners and (B) experienced learners. Error bars represent standard errors.

## 5 Discussion

The purpose of the study was to test whether American learners' pronunciation of Mandarin Chinese coarticulated tones would improve with classroom language learning experienced and whether tonal environment, tonal context and syllable position would affect their production accuracy.

### Effects of learning experience

The results showed that experienced learners produced coarticulated tones with a higher accuracy rate than inexperienced learners. This result confirmed the findings from previous studies that the ability to perceive and produce L2 sounds improves with experience.

### Effects of tonal environment and tonal context

There was no effect of tonal environment or tonal context on Tone 2, Tone 3, or Tone 4 production accuracy by either group of learners. However, there was an effect of tonal environment and tonal context on the accuracy rate of Tone 1 production by inexperienced learners, but not by experienced learners. Inexperienced learners produced Tone 1 with a higher accuracy rate in the identical tonal environment than in the non-identical tonal environment. In addition, their production of Tone 1 was more accurate in compatible tonal context than in conflicting tonal context.

### Effects of syllable position

Syllable position affected how accurately Tone 2 was produced by both experienced learners and inexperienced American learners. Both groups showed a higher production accuracy rate for Tone 2 in the final syllable position. This result may be attributed to the rising intonation used to form questions in English. On the other hand, the factor of syllable position only affected inexperienced learners' production of Tone 4. For inexperienced learners, the accuracy rate of Tone 4 production on the initial syllable was higher than that on the final

syllable. No clear explanation can be offered to explain why Tone 4 was produced more accurately in initial syllable position. Syllable position, however, did not affect production accuracy of Tone 1 or Tone 3 production.

## 6 Conclusion

This study revealed that American learners with more learning experience were more accurate in producing coarticulated Mandarin tones in disyllabic words. Furthermore, it was found that with increased experience, production of coarticulated tone becomes more resistant to such phonological and phonetic factors as tonal contexts and syllable position. Considering the effect of linguistic factors on the production of tones on disyllabic words by American learners, it was found that tonal environment only affected tone 1 production and syllable position affected tone 2 and tone 4 production accuracy. Neither tonal environment or syllable position has an influence on the accuracy rate of production of Tone 3.

## Reference

- Flege, J., Takagi, N. & Mann, V. (1995). Japanese adults can learn to produce English /r/ and /l/ and accurately, *Language and Speech*, 38, 25-55.
- Flege, J., Bohn, O.S., & Jang, S. (1997). The effect of experience on non-native subjects production and perception of English vowels. *Journal of Phonetics*, 25, 437- 470.
- He, Y., Wang, Q, & Wiltshire, C. (2008). Production of English lexical stress by inexperienced and experienced learners of English. *Canadian Acoustics*, Vol.36. No.3.128-129.
- Wang, Y., Jongman, A., and Sereno, J. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America*, 113, 1033-1043.
- Wang, Y., Spence, M., Jongman, A., and Sereno, J. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America* 106, 3649-3658.
- Wayland, R.P. and Guion, S.G. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report.. *Language Learning*, 54, 681-712.
- Wayland, R.P. & Guion, S.G. (2003). Perceptual discrimination of Thai tones by naïve and experienced learners of Thai. *Applied Psycholinguistics* 24, 113-129.
- Xu, Y. (1994). Production and perception of coarticulated tones. *Journal of the Acoustical Society of America* 95: 2240-2253.