Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20). 2008. Volume 1.
Edited by Marjorie K.M. Chan and Hana Kang. Columbus, Ohio: The Ohio State University. Pages 123-138.

Tone Production in Mandarin Chinese By American Students: A Case Study

Lijuan Guo
Harbin Institute of Technology

Liang Tao
Ohio University

This study reports finding on the developmental stages of tone production at the sentence level by American students in their first year of learning Mandarin Chinese. The study aims to identify the different developmental characteristics of each of the four tones in near-natural speech over one academic year.16 native speakers of American English in first-year Chinese language classes participated in this study. The data were collected at the end of each academic quarter for two non-consecutive quarters. The study analyzed learners' tonal performance and identified the different developmental characteristics of each of the four tones in both prepared and spontaneous oral speech during the first and third quarters of the 2006-2007 academic year. The study confirmed the difficulty of tone 3 in near-natural conversations.

1. Introduction

It is generally agreed upon that Mandarin Chinese (hereafter, Chinese) is one of the most difficult languages for English L1 learners to study (Ross, 2001). The syllable structures [(consonant) vowel (n/ng)] of Chinese are not particularly difficult for American learners of Chinese; instead, it is lexical tones that make the mastery of the language a challenging task (Shen, 1989). American learners of Chinese have very little difficulty in learning the perceptual discrimination of tones, but always experience difficulties in producing tonally correct speech (Chen. G. T., 1974). Moreover, as a Chinese instructor, I find that students can produce isolated tones correctly. However, they often have problems producing correct tones in connected speech, i.e. they mix up tones at the sentence level. For instance, students mispronounce tone 3 for tone 1, or mistakenly produce tones 1 or 4 instead of tone 2. They probably did not think they produced the wrong tones. I have often heard students say that they will never learn the tones, and even advanced level students still struggle with tones. This study intends to examine and analyze the tone production of 16 American learners of Chinese. More specifically, I will analyze their tonal performance and identify the different developmental characteristics of each of the four tones in both prepared and spontaneous

oral speech during two non-consecutive academic quarters (the first and third quarters in the 2006-2007 academic year).

2. Literature Review

As discussed above, Chinese tones can be described verbally by pitch contour (level, rising, falling) and pitch register (high, low, mid). For instance, tone 3 is a mid falling-rising tone, while tone 1 is a high level tone. Earlier studies such as Li and Thompson (1977), Tse (1978), and Yue (1980) claimed that in first language (L1, here-after) acquisition, the hierarchy of difficulty of tonal acquisition for native Chinese children was as follows: the level tone (i.e. tone 1) was easier than the falling (i.e. tone 4), the rising (i.e. tone 2) and the falling-rising (i.e. tone 3) tones; the falling tone was easier than the rising tone (referred to by Shen, 1989). Specifically, the acquisition order is: tone 1, tone 4, tone 2, and tone 3.

A number of auditory analysis and acoustic analysis studies on L2 learners' acquisition of Chinese tones (Chen. G. T., 1974; Chen. Q., 1997; Wang, 1995; Shen, 1989; White, 1981) have been conducted in order to establish American learners' hierarchy of difficulty of tonal acquisition as well as to explore the reasons for the tonal misproduction. An acoustic study by Shen (1989) claimed that tone 4 and tone 1 are the most problematic tones for American learners because they are less marked than tone 2 and tone 3. Shen (1989) claimed that L1 acquisition is different from L2 acquisition due to the different internal and external environments. Regarding the internal environment, unlike L1 learners, L2 learners have already acquired their native language before learning L2, indicating that they have already had the knowledge of another language. Concerning the differences of external environments, Shen (1989) goes further to point out that many L2 learners learn their second language from formal classroom instruction, while other L2 learners might pick up the language in different environments such as from their peers. Therefore, the differences of settings might bring different influence to the process of the L2 language acquisition (Shen, 1989).

Miracle's (1989) word-level acoustic analysis rejected the tonal acquisition order for American learners established in Shen's study. Miracle found that the students' tonal errors were evenly distributed among the four tones. The research revealed that although there is no significant difference, tone 2 might be particularly challenging for American learners.

Moreover, contrary to Shen's (1989) study, Lee et all (in press), Wang (1995), and Yue's (1986) studies displayed a different acquisition order for American learners. Yue's (1986) study found that the tonal acquisition order for L1 and L2 is similar: tone 1 is the easiest to acquire, followed by tone 4, tone 2 and tone 3 (referred to by Shen, 1989). Another word-level study on American learners' tonal acquisition order conducted by Wang (1995) supported Yue's findings on the hierarchy of difficulty. This study found that tones 2 and 3 are significantly more difficult for American learners to learn than

tones 1 and 4. Lee, Tao & Bond's (in press) tone perception study also displayed similar results.

From previous studies we can see that there is no agreement on the acquisition order. According to Peng & Hu (2006), the disagreement on the hierarchy of difficulty of the four tones might be mainly due to the differences of data types on which analyses were based. The method of data collection in Shen's study was to ask students to read a familiar text. The data collection at the sentence-level indicated that the investigated tones were not isolated tones but were merged in a phonological context, which inevitably affects the investigated tones. Whereas the data of other studies (Yue, 1986, referred to by Shen, 1989; Wang, 1995) were gathered from isolated words or phrases and, therefore, the surrounding phonological context had no effect on the participants' tone production.

Although previous studies did not arrive at an agreement on American learners' tonal acquisition order, studies agree American learners do have difficulties in learning Chinese tones. The interesting question is what, then, causes the difficulties? Previous studies claim that American students' difficulties in learning Chinese tones are caused by the physical differences of the two languages. For instance, G. T. Chen (1974) conducted an acoustic study that compared the pitch range between native Chinese speakers and native English speakers. He tested both word-level and sentence-level pitch, and found the average pitch range of Chinese speakers in their native language was 1.5 times wider than that of native English speakers speaking English. G. T. Chen (1974) thus, suggested that native speakers of non-tonal languages should learn to widen their normal pitch range in order to successfully produce Chinese tones. White (1981) provided another answer to the question of why learning Chinese tones is a challenging task for American learners. White (1981) found that American learners' tonal errors are not caused by randomly replacing one tone with another; rather, they are systematic errors that probably stem from speakers' transfer of English intonation patterns into Chinese sentences. To explain the observed tonal errors from a linguistic perspective, White (1981) compared the system of Chinese tones to that of English intonations. White found two major differences between Chinese tones and English intonations. First is the significantly wider pitch range of Chinese. The second difference is that mechanisms of stress in the two languages are different. He thus concluded that the systems of Chinese tones and English intonations include phonetic, phonological, and syntactic levels, while the system of English intonations also includes a pragmatic level. A tonal error analysis by Q. Chen (1997) also confirmed White's (1981) account of what causes American learners' tonal errors. Q. Chen's research clearly showed the interference of English intonation in both the elicited-speech test and the tonal perception task. The cause of the difficulty was not only learners' narrower pitch ranges, but also the negative transfer of English intonation. Studies by Chao (1980, referred to by Shen 1989) and Zhao (1989, cited in Shen, 1989) indicated that American learners' tonal misproduction stems from pitch register instead of pitch contour. It needs to be pointed out that even though Shen's study displayed a

different acquisition order from the other studies, her auditory and acoustic analysis of eight American learners, was consistent with the claims by Chao (1980) and Zhao (1989) that American learners' tonal misproduction was attributed to L1 interference. In addition, Gui's (2000) study, which focused on the interference of English intonation in Chinese tone production, analyzed common tonal errors of American learners. The study further proved that the main cause to those errors laid on the interference of English intonations in Chinese tones.

To those possible explanations, Wang, Jongman & Sereno's (2001) contrastive study in the perceptive domain provided more evidence about the reason why American learners have difficulties in acquiring Chinese tones. This study claimed that Chinese tones are predominantly processed in the left cerebral hemisphere by native Chinese speakers, suggesting that tones are processed as linguistic units, just like segmental properties. This study also revealed that, unlike native speakers of Chinese, American learners with no previous exposure to any tonal languages processed Chinese tones bilaterally. The causes of the difficulties encountered by American learners have been explored previously by the studies. However, many of those studies discussed have been done only at the word level. I believe that a more relevant way of detecting tone production improvements is beyond the lexical level. Moreover, according to White (1981), American learners' tonal errors mainly occur in sentences or in combinations of words, not in isolated syllables.

This study attempts to analyze students' tonal production at the sentence level so the effects phonetic context might have on tone production can be discovered. Moreover, no previous study has measured American learners' tonal production improvements. As a longitudinal study, this research aims at examining American learners' development of tonal production over two academic quarters. Thus, the developmental process of students' tone production will be tracked in order for an instructor to employ specific techniques. All of the previous studies in this area were conducted using purposefully elicited data or data from perception tasks. However, the data for this study is based on near-natural speech, which will help give us a better picture of American learners' acquisition of Chinese tones. At the time, the participants were unaware that their oral presentations would be used for a research project; therefore their tone production was not effected by additional stress.

3. Research questions

This study is based on the following research questions:

1) Are the four tones learned equally well? Which tones are produced with less accuracy, which tones are produced with more accuracy? 2) Are there any general improvements of the four tones in near-natural language use? 3) Do the participants try to avoid producing certain tones, such as tone 3?

4. Hypotheses

To answer these questions, based on the previous studies and my own language instruction experience, I have formulated three hypotheses that I will be testing throughout my research.

1) The four tones are not learned equally well. Tones 2 and 3 are produced with less accuracy than tones 1 and 4. 2) General improvements of the four tones can be detected in near-natural language use. 3) Students will not avoid producing certain tones.

5. Methodologies

5.1 Participants

The participants consisted of 16 native speakers of American English who studied first-year Chinese in the 2006-2007 academic year at Ohio University (OU). The reason I chose the 16 participants out of the 56 first year students is that these students are taking Chinese courses right now at OU, which makes it is easy for me to ask them to sign the consent form. They had four contact hours each week with their instructors during the 2006-2007 academic year. The 16 participants from different majors studied Chinese in three different classes. All three teachers co-taught the first quarter but not the third quarter. None of the students had any previous experience in learning Chinese before taking the first-year class.

5.2 Instruments

The data are the existing video taped oral presentations as part of the final exams of the three first-year classes in two academic quarters: the fall quarter and the spring quarter. Students gave the three-minute presentations in pairs. Not all of the 16 participants were paired with one another. There are two parts of each pair of presenters' oral presentation. One part is a prepared conversation between the two speakers. The other is a spontaneous question-and-answer part, when the audience (the whole class except for the presenters and the teacher) were requested to ask questions according to each conversation; the presenters were supposed to answer the questions raised by the audience. Each student was requested to ask at least three questions throughout all of the presentations. Certain grammar patterns and vocabulary the students had learned were assigned to be used in the prepared presentation. All the 16 participants took the two final oral presentations.

5.3 Procedure

The video taped oral speeches were converted into DVD for the convenience of listening and watching before the transcription was conducted. I transcribed each participant's oral presentations. With the transcriptions ready, I started to work on the rating of the accuracy of tones. One more native speaker of Chinese was invited to evaluate the oral presentations in order to confirm my evaluations. This second rater is a Chinese scientist who has never been exposed to nonnative speakers of Chinese. The reason I

asked one more native speaker rather than a veteran teacher is to avoid possible bias from Chinese instructors who are more or less familiar with American students' tone production. My second rater and I worked independently. The criteria for evaluation were established according to the 1 to 5 numerical scales for the normal pitch range of a native Chinese speaker. I trained my second rater to use the 1 to 5 numeral scale-criteria before he started to work on the evaluation to help him determine whether a tone is indeterminate or unacceptable. Therefore, his judgment was based on the 1 to 5 numeral scale-criteria as well as a native Chinese speaker's intuition. The second rater was allowed to listen to each utterance as often as he wanted. My judgment was based on the 1 to 5 numeral scale-criteria, on the linguistic training I have received, and on my past language teaching experience. I counted the total number of each of the four tones produced by each participant. For example, a participant produced tone 1 ten times, and produced tone 3 nine times. Then my second rater and I evaluated each participant's tone production for each tone according to the three categories: unaccepted tones, indeterminate tones and acceptable tones. For instance, the participant produced the first tone five times acceptably, three times indeterminately and two times unacceptably. Lastly, I calculated the percentage of each category for data analysis. So, the participant produced the first tone acceptably 50 percent of the time, indeterminately 30 percent and unacceptably 20 percent. To answer my research questions, I will only focus on the acceptable tones for data analysis for the time being; the indeterminate tones and unacceptable tones will be taken into consideration in follow-up studies.

5.4 Data Analysis

I carried out statistical analysis to see whether there are any differences between the two raters' evaluations. The agreement between the two raters' evaluations was tested using the General Linear Model Test. The result reflects that there is no significant difference between the two sets of evaluations in the categories of acceptable tones ($p>.1$) and that of unacceptable tones ($p>.1$), confirming that the instructor's judgment was based on the standard perception of native speakers. Therefore, the data discussion of this study is based on my own evaluation.

6. Results And Discussion

In the following discussion, *Test One* will be used to stand for the participants' oral speeches in the fall quarter, while *Test Two* stands for their oral speeches in the spring quarter.

Figures 1, 2 and 3 show the raw scores of Test One and Test Two. In these figures, *AC* stands for the number of acceptable tones, *TN* is used for the total number of the participant's tone production for each tone, *AC (%)* refers to the percentage of acceptable tones, and the *average* stands for the 16 participants' average percentage of acceptable tones.

Subjects	Tone 1			Tone 2			Tone 3			Tone 4		
	AC	TN	AC(%)	AC	TN	AC(%)	AC	TN	AC(%)	AC	TN	AC(%)
1	2	7	28.571	4	12	33.333	2	13	15.385	4	7	57.143
2	10	11	90.909	6	7	85.714	13	23	17.391	9	14	64.286
3	8	12	66.667	3	4	75	4	8	50	11	11	100
4	7	11	63.636	1	7	14.286	7	17	41.176	15	24	62.5
5	6	9	66.667	5	8	62.5	10	18	55.556	7	14	50
6	6	7	85.714	10	13	76.923	13	21	61.905	8	13	61.538
7	4	6	66.667	12	16	75	16	19	84.211	5	6	83.333
8	8	10	80	1	8	12.5	9	21	42.857	8	16	50
9	5	5	100	3	13	23.077	6	25	24	10	15	66.667
10	7	10	70	3	3	100	8	21	38.095	5	8	62.5
11	10	12	83.333	9	14	64.286	12	17	70.588	12	15	80
12	13	15	86.667	4	4	100	8	20	40	13	26	50
13	6	7	85.714	8	20	40	15	26	57.692	13	19	68.421
14	12	16	75	4	8	50	10	21	47.619	9	17	52.941
15	12	17	70.588	4	8	50	9	16	56.25	8	15	53.333
16	16	19	84.211	2	4	50	9	14	64.286	5	9	55.556
		174			149			300			229	
average			75.272			57.039			47.938			63.639

Figure 1: Test One

Subjects	Tone 1			Tone 2			Tone 3			Tone 4		
	AC	TN	AC(%)	AC	TN	AC(%)	AC	TN	AC(%)	AC	TN	AC(%)
1	23	35	65.714	8	34	23.529	21	57	36.842	36	58	62.069
2	8	17	47.059	14	18	77.778	31	39	79.487	16	30	53.333
3	4	4	100	8	16	50	16	23	69.565	26	30	86.667
4	2	3	66.667	7	14	50	16	25	64	9	20	45
5	19	32	59.375	17	24	70.833	20	54	37.037	15	40	37.5
6	7	13	53.846	10	21	47.619	23	28	82.143	25	33	75.758
7	11	15	73.333	9	16	56.25	34	53	64.151	15	31	48.387
8	19	29	65.517	9	12	75	15	45	33.333	18	34	52.941
9	9	13	69.231	20	30	66.667	13	24	54.167	11	23	47.826
10	9	13	69.231	13	30	43.333	14	39	35.897	21	25	84
11	7	13	53.846	13	14	92.857	22	32	68.75	9	14	64.286
12	18	25	72	7	12	58.333	29	44	65.909	22	33	66.667
13	13	16	81.25	11	17	64.706	22	32	68.75	19	26	73.077
14	12	19	63.158	14	28	50	31	53	58.491	16	19	84.211
15	3	6	50	11	12	91.667	22	29	75.862	6	14	42.857

16	4	9	44.444	4	12	33.333	11	52	21.154	10	21	47.619
		262			310			629			451	
average			64.667			59.494			57.221			60.762

Figure 2: Test Two

Subjects	Tone 1		Tone 2		Tone 3		Tone 4	
	Test1	Test2	Test1	Test2	Test1	Test2	Test1	Test2
1	7	35	12	34	13	57	7	58
2	11	17	7	18	23	39	14	30
3	12	4	4	16	8	23	11	30
4	11	3	7	14	17	25	24	20
5	9	32	8	24	18	54	14	40
6	7	13	13	21	21	28	13	33
7	6	15	16	16	19	53	6	31
8	10	29	8	12	21	45	16	34
9	5	13	13	30	25	24	15	23
10	10	13	3	30	21	39	8	25
11	12	13	14	14	17	32	15	14
12	15	25	4	12	20	44	26	33
13	7	16	20	17	26	32	19	26
14	16	19	8	28	21	53	17	19
15	17	6	8	12	16	29	15	14
16	19	9	4	12	14	52	9	21

Figure 3: Total Number of the 16 Participants' Tone Production

6.1 Test One

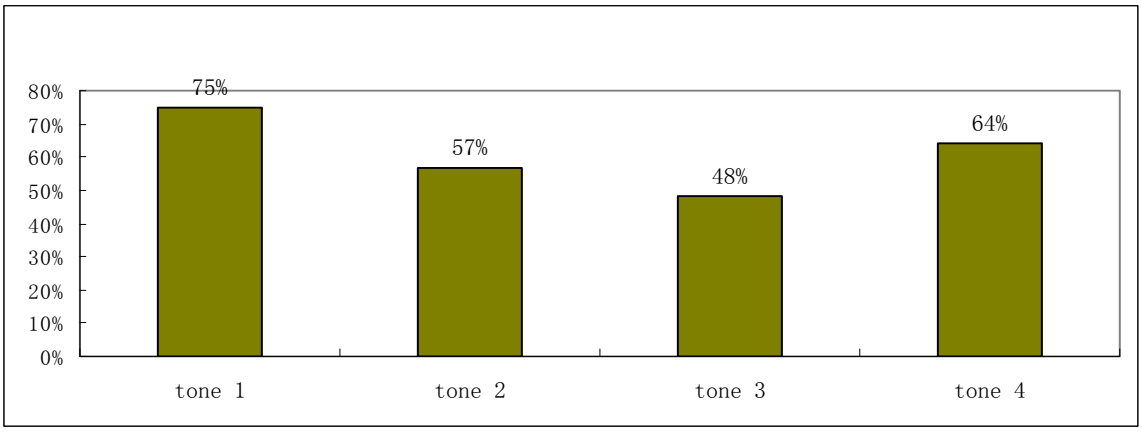

Figure 4: Percentage Distribution of Acceptable Tones in Test One

Figure 4 shows the order of accuracy rate of the four tones: tone1 (75%), tone 4 (64%), tone 2 (57%) and tone 3 (48%). The statistical analysis from Multivariate Tests did not indicate significant differences in the accuracy rates of tone 1 (75%) and tone 2 (57%), and tone 4 (64%) and tone 2, which I did not expect. However, the statistical analysis shows significant differences in the accuracy rates of tones 1 and 3 (48%), and tones 4 and 3, which I did expect. The difference between tones 1 and 3 is 27.33 % ($p<.05$), and the difference between tones 4 and 3 is 15.70 % ($p<.05$). The significant difference between tone 3 and tones 1 and 4 highlights the fact that tones 1 and 4 were produced with the greatest accuracy, while tone 3 was produced with the least accuracy.

6.2 Test Two

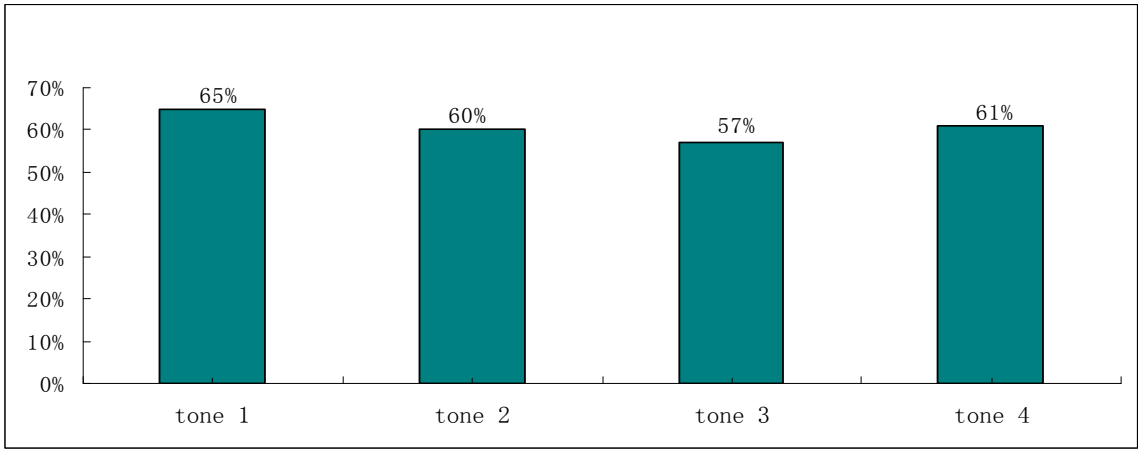

Figure 5: Percentage Distribution of Acceptable Tones in Test Two

Figure 5 also displays an order of the accuracy rate in Test Two: tone 1 (65%), tone 4 (61%), tone 2 (60%) and tone 3 (57%). However, the statistical analysis from the Multivariate Tests did not show any significant differences in the accuracy rate among the four tones. This implies that the differences in the accuracy rate have been minimized. In other words, there is tendency of moving toward uniformity. The next chart (Figure 6) further supports my view.

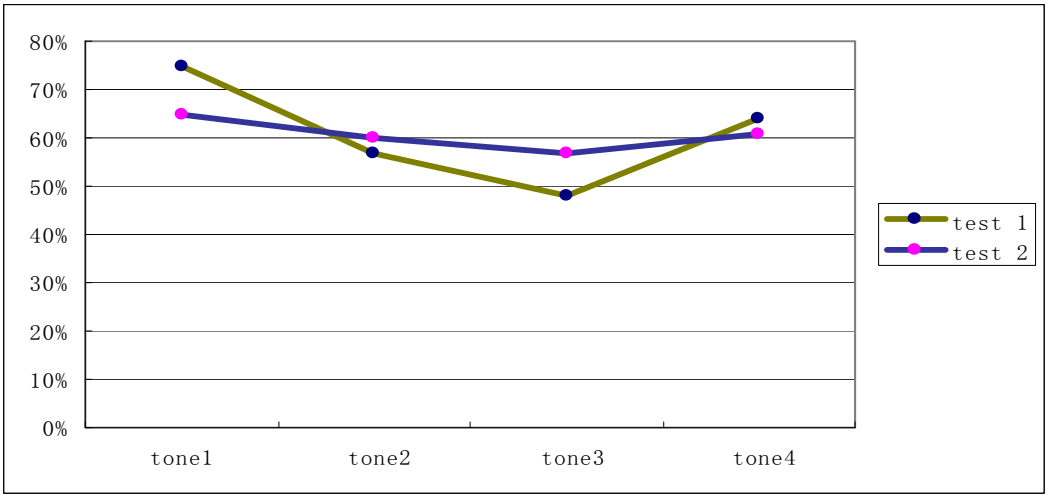

Figure 6: Percentage Distribution of Acceptable Tones

Figure 6 revealed that in Test One, the accuracy rate of the four tones forms a V curve, indicating the significant differences between tone 3 and tones 1 and 4. In other words, tones 1 and 4 were produced with greatest accuracy, while tone 3 was produced with the least accuracy. However, in Test Two, the distribution of the accuracy rate levels, indicating that the development of the four tones has moved toward uniformity. Moreover, it is obvious that the acceptable rating of tones 2 and 3 has increased in Test Two. The results also displayed that 10 (62.5%) out of 16 participants' accuracy rates for tone 3 increased, and 8 (50%) out of 16 participants' accuracy rates in tone 2 increased (see figures 1 & 2). Given that students themselves might pay more attention to the two difficult tones, this result is not at all surprising. In addition, the fact is likely a result of the teachers' emphasis. I also interviewed the instructors to see whether they made extra efforts to alert students about tones 2 and 3 in the spring quarter. I was told they did in both the winter and the spring quarters. One the other hand, this might indicate that students can produce tones in shorter sentences correctly. However, they often have problems producing correct tones in longer sentences. For instance, participants 2 and 16's percentage of acceptable tones in tone 1 are both above 80 % in Test One (see figure 1), while in Test Two those decreased to below 50% (see figure 2). Regarding tone 4, participant 3's percentage of acceptable tones is 100% in Test One (see figure 1), while it declined to 86% in Test Two (see figure 2). This trend can also be seen in participant 7's tone production. The percentage of her acceptable tones in tone 4 is 83% in Test One (see figure 1), while it decreased to 48% (see figure 2) in Test Two. The decline can serve as a reminder for both students and teachers that students, especially at the early stages of their study, need to pay more attention to each of the four tones, even though they might think tone 1 and tone 4 are easier. On the other hand, although the accuracy rates of tones 1 and 4 appear to decrease, the ratings are higher still than those of tones 2 and 3.

Therefore, the reason for the decease might be that the students thought tones 1 and 4 were easier compared to tones 2 and 3. They might have thought they knew how to produce them after the first quarter's study, they, thus, did not have to study more on them later on. Or it might be because the teachers emphasized the study of tones 2 and 3 rather than tones 1 and 4.

6.3 Improvements: Test One and Test Two

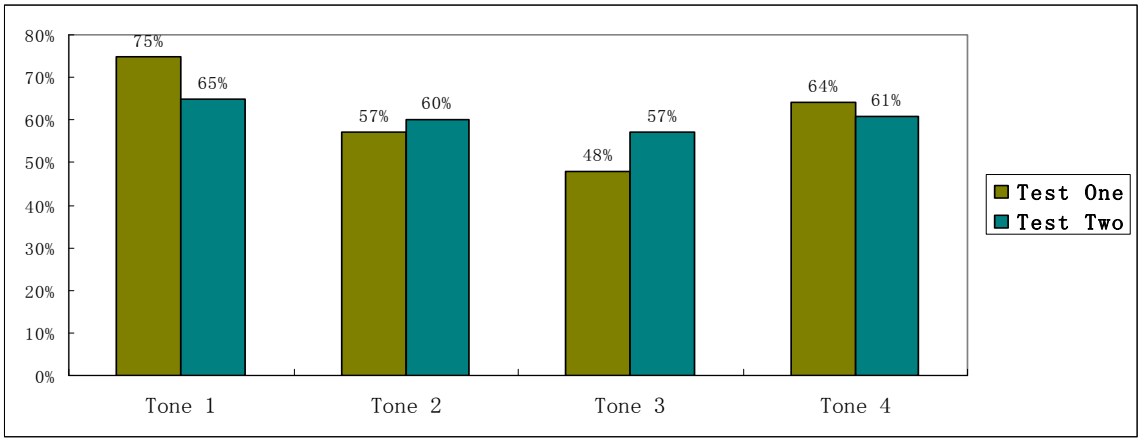

Figure 7: Percentage Distribution of Acceptable Tones

Figure 7 shows that the orders of accuracy rate in these two tests are the same: tone 1, tone 4, tone 2 and tone 3. However, as discussed above, there are no significant differences in the accuracy rates among the four tones in Test Two. Even though there is a trend of improvement of tones 2 and 3, the results from statistical analysis did not show any significant improvement. Instead of improvements, there are decreases of accuracy rates in tones 1 and 4.

Even though I did expect improvements, unfortunately, the statistical analysis from the Independent T-test did not show any significant improvements between the two tests. However, I believe there is overall improvement beyond the accuracy rate of the tones because there are a few other factors that influence students' tone production. First of all, the mean lengths of utterance (MLU, hereafter) were greater in Test Two than in Test One (see Figure 8).

Subjects	Test One		Test Two	
	number of utterances	number of syllables	number of utterances	number of syllables
1	12	39	17	124
2	16	55	23	104
3	9	35	15	73
4	15	59	16	62
5	11	49	26	150
6	14	54	20	95
7	13	47	17	115
8	18	55	19	120
9	17	58	16	90
10	13	42	23	119
11	15	58	14	73
12	16	65	19	114
13	16	72	15	91
14	13	62	15	119
15	15	56	20	61
16	13	46	18	94
total number	226	852	293	1664

Figure 8: Numbers of Syllables and Utterances

I tabulated the 16 participants' tone productions in order to calculate the MLU in each test. The results showed that the MLU is 3.8 in Test One and 5.7 in Test Two. Moreover, there is no significant increase in the number of utterances overall. The total numbers of the 16 participants' utterances in Test One is 226 and 293 in Test Two while there is a significant increase in the number of syllables. The total number of the 16 participants' syllables in Test One is 852; however the total number of their syllables in Test Two is 1664 (see figures 3 & 8).

The second factor that needs to be noticed is that a speed-accuracy trade-off may have influenced the students' speech. Although the accuracy rates of tones 1 and 4 declined, both the speed of their speech and the number of their syllables increased, and the sentence structures got more complicated. I calculated the 16 participants' total syllables for each tone in the two tests. Figure 9 shows the distribution of the 16 participants' syllables of each tone of the two tests.

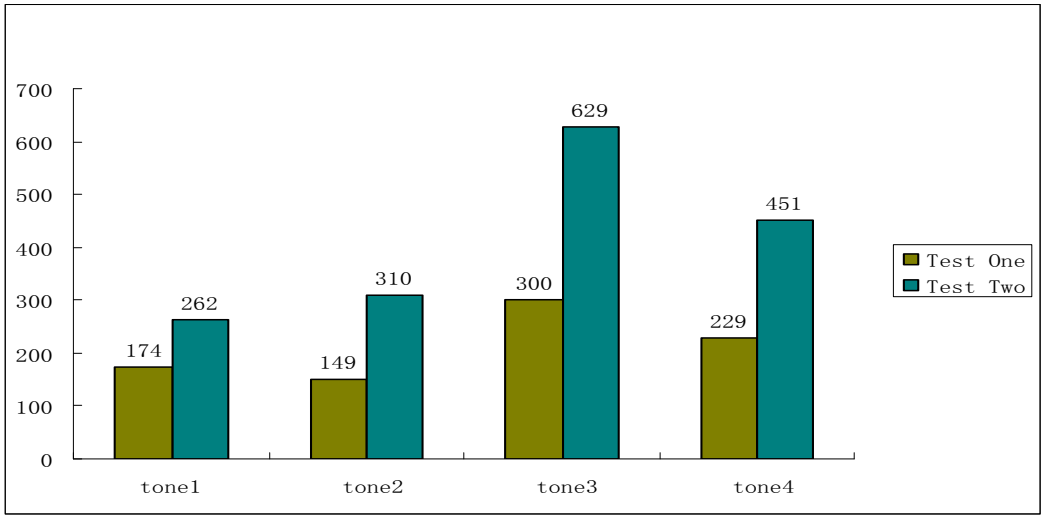

Figure 9: 16 Participants' Total Number of Syllables

Figure 9 convincingly highlights the fact that the total number of syllables for all the four tones significantly increased in Test Two. Moreover, in Test Two, the complexity and prosody of the sentences greatly increased. Therefore, the reduced accuracy of tones 1 and 4 in Test Two may have reflected tone production in conversations at the sentence level; whereas in Test One many tones may have been produced at the lexical level.

In addition, in Test Two, students produced longer sentences at a faster pace. For example, in Test One, participant 2 produced a six-syllable utterance for about 3.5 seconds indicating she used 0.58 second to produce one syllable; while in Test Two, she produced a seven-syllable utterance for about 2.7 seconds meaning she produced one syllable in 0.39 second (see figures 10 & 11).

Figures 10 and 11 also displayed that in Test One, the student divided the six-syllable sentence into 3 units. In addition, it is easily seen that the participant paused longer between each syllable and each unit in Test One; however, she did not divide the seven-syllable sentence into units in Test Two, and she did not pause between each syllable either. As discussed above, in Test One, students might unconsciously produce tones at the lexical level rather than the sentence level to make the utterance easier to produce. However, they were more confident in producing tones in Test Two, so they pronounced them more naturally in terms of rhythm.

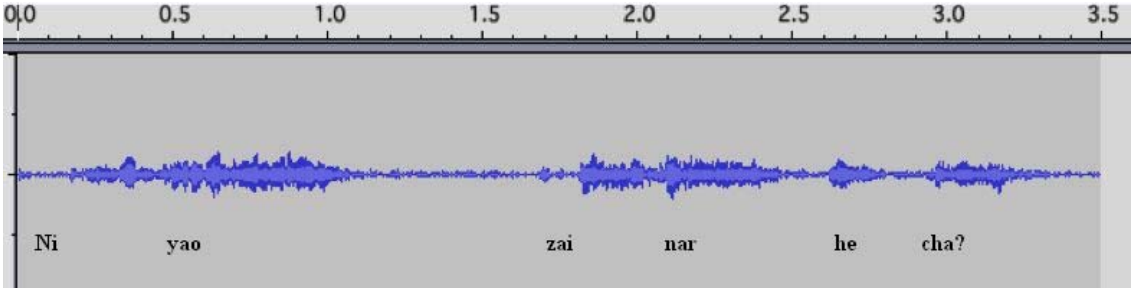

Figure 10: Participant 2's Six-syllable Sentence in Test One

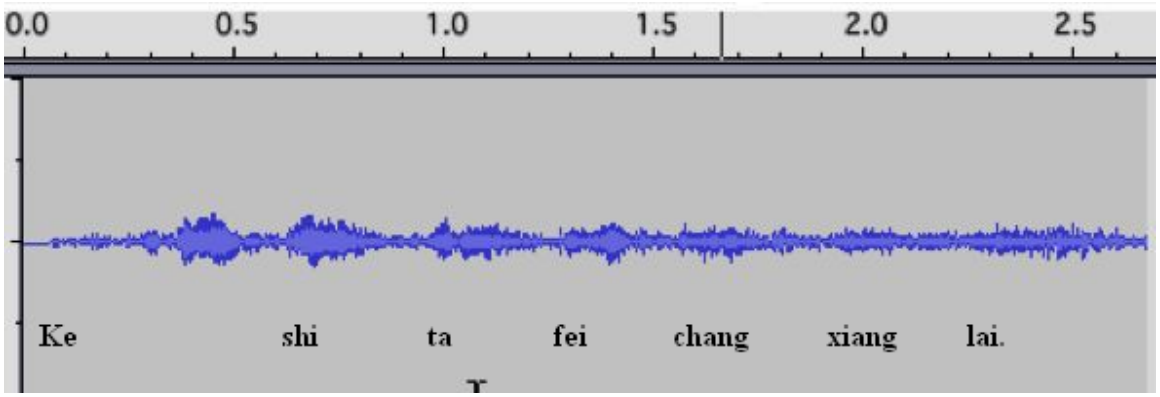

Figure 11: Participant 2's Seven-syllable Sentence in Test Two

I interviewed some of the participants. They said that they paid more attention to syntactical structures and semantic meanings than to tones because they preferred to express themselves meaningfully than to pronounce accurately. Therefore, this may not indicate a decrease or decline in the quality or ability of speech; rather it is a normal phenomenon or stage at this level of tone production.

6.4 Tone 3

The production of tone 3 is significantly more frequent than that of tone 1, tone 2 and tone 4 in both of the two tests (See figures 3 & 9). Moreover, there is a significant increase in the number of the participants' productions of tone 3, from 300 in Test One to 629 in Test Two. Furthermore, I calculated the number of the 16 participants who produced tone 3 more than the other tones. The results shows that 11 out of 16 (68.8%) participants produced tone 3 the most in Test One, 12 out of 16 (75%) participants produced tone 3 the most for Test Two; figure 9 displays that the frequency of usage of tone 3 ranks second in Test One, and first in Test Two. The results convincingly highlight the fact that no correlation exists between the frequency of usage of a particular tone and its degree of difficulty. Obviously, the participants did not try to use avoidance as a

strategy, to be specific, to avoid pronouncing tone 3 because it is the most problematic one.

To sum up, my first hypothesis was partially confirmed. That is, tones 1 and 4 are produced with the greatest accuracy. However, in Test two no significant improvements were found, which means my second hypothesis was rejected. My third hypothesis was confirmed, indicating the participants produced tone 3 most frequently. It might be interesting to investigate the frequency of the four tones' production in my follow up studies.

7. Conclusion

By collecting data from oral presentations, this study analyzed American learners' tonal production at the sentence level in order to investigate learners' tonal progress over the two non-consecutive academic quarters.

This analysis and investigation identifies student problem areas regarding tonal production. The results of this study may provide insights into the difference of the four Chinese tones, and how American students learn the four tones so that Chinese teachers can better understand students' specific difficulties in order to help them improve tonal production. That is important because early monitoring and practicing of tones from instructors is a crucial, beneficial aspect of tone production by American students. In addition, the study attempts to remind Chinese instructors that their primary task is not merely to teach learners how to produce these tones in isolation. A teacher's fundamental task is to know the learners' tonal performance at the sentence level and to help them consciously overcome L1 interference.

REFERENCES

Chao, Yuen Ren.1968. A grammar of spoken Chinese. Berkeley: University of California Press.

Chen, Gwang-tsai. 1974. The pitch range of English and Chinese speakers. Journal of Chinese Linguistics, 2, 159-171.

Chen, Qinghai. 1997. Toward a sequential approach for tonal error analysis. Journal of the Chinese Language Teachers Association, 32, 21-39.

Gui, Mingchao. 2003. Zailun meiguoyingyu yudiao dui meiguo xuesheng xuexi hanyu shengdiao de ganrao. (The interference of American English intonation patterns in Mandarin tone studies revisited), Journal of Yunnan Normal University, 1, 38-43.

Gui, Mingchao.2000. Meiguoyingyu yudiao dui meiguo xuesheng xuexi hanyu shengdiao de ganrao**.** Shijie hanyu Jiaoxue, 51, 89-93 **(**The interference of American English intonation.) The World Chinese Teaching, 51, 89-93.

http://www.chinese-outpost.com/language/pronunciation/tones-of-mandarin-chinese.asp

Kwock-Ping Tse, John. 1978. Tone Acquisition in Cantonese: A Longitudinal Case Study. Journal of Child Language 5: 191-204

Lee, Chaoyang, Tao, Liang and Bond, Zinny. in press. Identification of Acoustically Modified Mandarin Tones by Native Speakers. Journal of Phonetics.

Li, Charles N. and Thompson, Sandra A. 1977. The Acquisition of Tone in Mandarin Speaking Children. Journal of Child Language, pp. 185-199.

Miracle,W. Charles. 1989.Tone production of American students of Chinese: A preliminary acoustic study. Journal of the Chinese Language Teachers Association, 24, 49-65.

Peng, Yukang, Hu, Yuanyuan. 2006 Dui wai hanyu shengidiao jiaoxue yanjiu huiwang. Jinnan Daxue Huanwen Xueyuan Xuebao, 4, 18-26. (An review of studies on teaching tons in Chinese as a second language. The Journal of Chinese College in Jinan Univesity, 4, 18-26)

Ross, Claudia. 2001. Evaluating intermediate Chinese textbooks. *Journal of the Chinese Language Teachers Association*, 36, 1-22.

Shen, Xiaonan Susan. 1989. Toward a register approach in teaching Mandarin tones. Journal of the Chinese Language Teachers Association, 24, 27-47.

Tao, Liang, Bond, Zinny & Beutner, Michael. 2006. Speech recognition technology in the instruction of Mandarin Chinese. Journal of the Chinese Language Teachers Association, 41(3), 51- 88.

Wang, Yunjia. 1995. Ye tan meiguoren xuexi hanyu shengdiao. Yuyan Jaoxue Yu Yanjiu, 2, 126-140. (On American learners' tone acquisition. Language Teaching and Research, 2, 126-140.)

Wang, Yue, Jongman, Allard and Sereno, Joan A. 2001. Dichotic perception of Mandarin tones by Chinese and American listeners. Brain and Language, 78(3), 332-348

White, C.M. 1981. Tonal pronunciation errors and interference from English intonation. Journal of the Chinese Language Teachers Association, 16(2), 27-56.

Yue-Hashimoto, A. O. 1986 Shengdiao jiaofa de Shangque. Diyijie Guoji Hanyu Jiaoxue Taolunhui Lunwenxuan, pp. 229-235