COMPARATIVE STUDY ON CHINESE TONE PERCEPTION: A REPORT ON FALLING F0 CONTOURS

Wen Cao[a] *& Qiuyu Wang*[b]

[a]Center for Studies of Chinese as a Second Language,
Beijing Language and Culture University, China;
[b]Department of East Asian Languages and Civilizations, Harvard University, USA
tsao@blcu.edu.cn; wang50@fas.harvard.edu

ABSTRACT

The main task of this study is to establish if a separate low-falling tone could be recognized as Tone-3 in Mandarin. Sixty listeners with different language backgrounds take part in the experiment. Their perception of twelve end-fixed falling pitch contours is reported. The results show that Chinese participants do not perceive low-falling tones as Tone-3; however, both Thai and the other foreign listeners do.

Keywords: tone, perception, low-falling, tone-3, Mandarin

1. INTRODUCTION

The four lexical tones in Mandarin have been traditionally described in the five pitch level system as /55/, /35/, /214/ and /51/, respectively [1, 4, 11]. However, since Chao himself stated that it was mostly realized as low-falling or low-level [4], the value of Tone-3 (hereafter referred as T3) has been a great point of contention recently. For example, Wang [15] argued that T3 was "basically" a low-level tone, and the beginning falling part and the ending rising part were subordinate. However, some recent researches [2, 3, 16] disagreed with the low-level view. In addition, Gårding, et al [6] showed that an introductory fall was important for T3. Both Lin [9] and Shen [13] found that the low-falling pitch was the most common form of T3, and it could even exist before a pause or at the end of a sentence. Some TCFL teachers nowadays insist on low-falling T3 as its basic form to start with when they are teaching Chinese tones to foreigners [12, 14, 16].

The disagreement on the basic form of T3 in phonetics has also been reflected in phonology, in which T3 has been described differently as L, LL and ML separately by different researchers [8, 10, 17].

It is known that, as a T3 sandhi rule in SC, /214/ is to be /21/ when it precedes a non-T3 syllable. Now the question is:

- Can a separate low-falling tone (e.g. /21/) be recognized as T3 in Mandarin?

Being relative to teaching and learning, two questions require further concern:

- How do learners of Chinese as a second language (L2) perceive the low-falling tone?
- To what extent does experience with tone in their first language (L1) affect their perception?

The main goal of the present study is trying to find the answer of these questions.

2. METHOD

2.1. Listeners

Sixty university students, aging from 20 to 30, took part in the test. Twenty were Chinese natives (CNS), twenty were Thai students (THS), and the other twenty were non-tonal language nation students (NTS). They were from France, Germany, Russia and Korea.

Both THS and NTS had learnt Chinese in Beijing for 1.5 - 2 years.

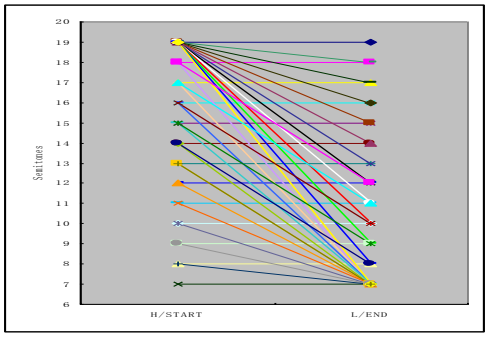
2.2. Stimuli

The original syllable for synthesizing stimuli in this experiment was "*kai* [k^hai]" with a high-falling tone valuing from 19.2 st_{100} (i.e. 304 Hz) to 9.6 st_{100} (i.e. 174 Hz). It was pronounced by a female Thai Chinese teacher whose average pitch range was from 152 Hz low to 300 Hz high, almost reaching an octave.

A sum of 41 *kai* with different pitch contours (PCs), which included 13 level ones and 28 falling ones (showing in Fig. 1), were synthesized with the help of Praat [*http://www.fon.hum.uva.nl/praat/*]. Among the falling contours, twelve were end fixed (i.e. at 7 st_{100}) and beginning varied at 1 st_{100} interval; another twelve were beginning fixed (i.e. at 19 st_{100}) and end varied at 1 st interval; the other six were sloping rate fixed (falling with 6 st_{100} down trend)[1].

Each stimulus was kept as long as the original *kai*, i.e. 235 ms (150 ms in rime).
Being randomized first, all these *kai* were combined into one wave file, with 3 second silence in between.

Figure 1: Synthesized pitch contours for stimuli.



2.3. Procedure

All participants listened to the stimulus wave file with headphones in a classroom. They were asked to identify the tone of each syllable and write it down on paper during each 3 second silence interval. 1 stood for T1, 2 for T2, 3 for T3 and 4 for T4. The wave file was played twice.

Before the experiment began, a 10-item practice session (10 trials randomly selected) was presented to familiarize listeners with the procedure.

Table 1: Pitch values of the twelve falling (HL) tones which are to be reported in the present paper.

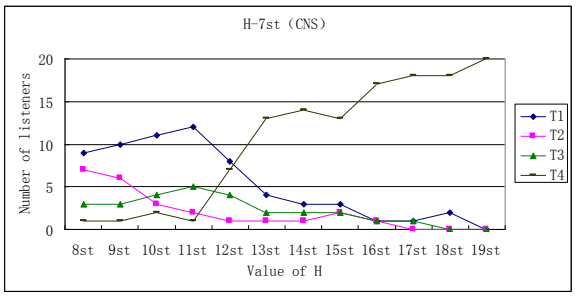
No.	H	L	FPL	No.	H	L	FPL
①	8 st	7st	**11**	⑦	14 st	7st	**31**
②	9 st	7st	**11**	⑧	15 st	7st	**41**
③	10 st	7st	**21**	⑨	16 st	7st	**41**
④	11 st	7st	**21**	⑩	17 st	7st	**51**
⑤	12 st	7st	**31**	⑪	18 st	7st	**51**
⑥	13 st	7st	**31**	⑫	19st	7st	**51**

Note: FPL=Five Pitch Level

3. RESULTS AND ANALYSIS

Figures 2-4 show falling tone perception results of CNS, THS and NTS, respectively.

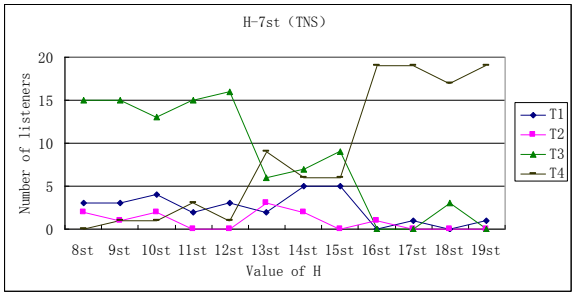
Figure 2: Falling tone identification by CNS.



Apparently the PCs with their H higher than 12 st tend to be regarded as T4 by Chinese natives (see Fig. 2). Statistical tests show that this tendency is fairly significant ($\chi^2_{13st}(3)=7.6$, $P \approx 0.05$; $\chi^2_{14st}(3)=8.7$, $P<0.05$; $\chi^2_{15st}(3)=6.6$, $0.05<P<0.10$), or very significant ($\chi^2_{16st}(3)=14.5$, $P<0.001$; $\chi^2_{17st}(3)=17.7$, $P<0.001$; $\chi^2_{18st}(3)=18.6$, $P<0.001$; $\chi^2_{19st}(3)=24$, $P<0.001$). However, when the H(s) ≤ 12 st, no pitch contour but the 11-7 st (H=11 st) shows fairly significant tendency ($\chi^2_{11st}(3)=6.8$, $0.05<P<0.10$), and it tends to be regarded as T1 (Not T3).

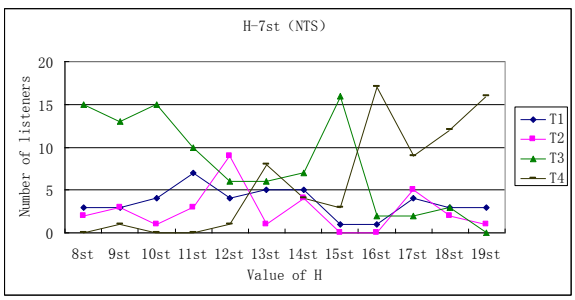
This result leads to two findings: (1) Chinese natives tend to perceive a tone higher than /31/ (including a phonetically high /31/) as T4, which is commonly described as "high falling tone"; (2) A low falling tone, which is lower than /31/ (even a phonetically lower /31/) does not have tendency to be perceived as T3 by Chinese.

Figure 3: Falling tone identification by THS



The perception of falling tones by the Thai listeners is obviously different from that of Chinese. And three categories can be induced: (1) when the H(s) ≤ 12 st, the PCs tend to be regarded as T3 ($\chi^2_{8st}(3)=11.8$, $P<0.01$; $\chi^2_{9st}(3)=10.8$, $P<0.05$; $\chi^2_{10st}(3)=7.6$, $P \approx 0.05$; $\chi^2_{11st}(3)=11.8$, $P<0.01$; $\chi^2_{12st}(3)=13.9$, $P<0.005$); (2) When the H(s) ≥ 16 st, the PCs tend to be regarded as T4 ($\chi^2_{16st}(3)=\chi^2_{17st}(3)=\chi^2_{19st}(3)=20.8$, $P<0.001$; $\chi^2_{18st}(3)=17.0$, $P<0.005$); (3) When the 12 st <H<16 st, no PC shows significant tone tendency ($\chi^2_{13st}(3)=3.0$, $P>0.25$; $\chi^2_{14st}(3)=1.7$, $P>0.5$; $\chi^2_{15st}(3)=6.2$, $P>0.10$).

Figure 4: Falling tone identification by NTS.



The result of NTS shows more complexity (see Fig. 4).

The first three falling PCs, in which H ≤ 10 st, tend to be perceived as T3 by NTS ($\chi^2_{8st}(3)=11.8$, $P<0.05$; $\chi^2_{9st}(3)=7.2$, $0.05<P<0.10$; $\chi^2_{10st}(3)=12.8$, $P<0.05$). They are all low-falling tones, and can be described in FPL system as /11/, /11/ and /21/.

And for PCs in which 11 st ≤ H ≤ 14 st, tone identification is nearly at chance level ($\chi^2_{11st}(3)=7.5$, $0.05<P<0.10$; $\chi^2_{12st}(3)=4.0$, $P>0.25$; $\chi^2_{13st}(3)=3.4$, $P>0.25$; $\chi^2_{14st}(3)=0.6$, $P>0.95$).

When the H ≥ 15 st, the result of NTS's identification of the PCs seems irregular. 16 out of 20 listeners (or 80%) regard the 15-7 st tone as T3, tendency being significant ($\chi^2_{15st}(3)=13.9$, $P<0.005$); and more (17, or 85%) regard the 16-7 st tone as T4, which is yet again significant ($\chi^2_{16st}(3)=15.5$, $P< 0.001$). However, when the H becomes higher, e.g. 17 st, the rate of being regarded as T4 decreases, and it does not show significant tendency of any tone perception ($\chi^2_{17st}(3)=5.2$, $P>0.10$). Neither does the 18-7 st PC ($\chi^2_{18st}(3)=2.5$, $P>0.25$), though T4 perception rate increases somehow. The only exception is the 19-7 st PC, which 16 out of 20 NTS listeners take as T4, and the tendency is significant($\chi^2_{19st}(3)=13.9$, $P<0.005$). This situation can only be explained by checking the stimulus order.

For instance, by looking at the appendix of this paper, it can be found that the stimulus before 15-7 st is the 19-12 st PC, which is surely a high-falling tone. So, when the NTS listeners heard the 15-7 st, they might have compared it with the former high-falling pitch, and be affected to perceive it as a low-falling tone. As a contrast, stimulus 16-7 st follows 15-15 st (/44/ in FPL) which is so easily regarded as a high level tone (i.e. T1) that the tone with higher beginning pitch (i.e. 16-7 st) is mis-confirmed by the listeners, and it is regarded as T4 with a rate even higher than that of 19-7 st (/51/). The other results can be explained similarly.

What is discovered here is that NTS listeners are subject to stimulus order or context when they listen to a Chinese pitch contour, even when the silence in between reaches 3000 ms (cp. Hallé, et al [7]).

It can also be said that NTS have not yet mastered Chinese tonemes. Otherwise, they would not be affected by the order or context after such a long silence.

4. DISCUSSION

Although low-falling tones such as /21/, /11/ and even /31/ somehow have more chances to be perceived as T1 by Chinese listeners (shown in Fig.2), statistical test tells that they do not have clear perception tendency. Therefore, the first question raised previously in this paper can be answered right now.

- A separate low-falling tone (e.g. /21/) can not be recognized as a T3 in Mandarin.

This finding suggests that our knowledge and description on T3 shall come back to confirm the traditional saying (if we admit that it is an identical tone and can be uttered as independent as the other tones do). T3 is basically a "dipping tone", just like /214/. A single/separate T3 can not be simplified as /21/, /11/, or /LL/, /ML/.

T3 is a little more complicated than the other three tones in Mandarin, but this is the truth.

However, how do we read the result that a majority of the THS and NTS tend to perceive the low-falling tones as T3?

Gandour & Harshman [5] did a research to study the cross-language differences in tone perception. They found that non-tone language speakers (e.g. Americans in the experiment) tended to rely on Average Pitch and Extreme Endpoint, and Thai participants mainly relied on Average Pitch, Direction and Slope. They concluded that people with different language backgrounds perceived tones on different acoustic cues.

The result that THS and NTS tend to perceive the low-falling tones as T3 can partly be explained with the theory in [5]. Assuming the four tones of Mandarin being the most common /55/ (T1), /35/ (T2), /214/ (T3) and /51/ (T4), their Average Pitches will be /5/, /4/, /2/ and /3/, respectively. Thus, T3 is the lowest. Among 12 end-fixed (at 7 st) falling tones in the present experiment, there are four PCs of which the Average Pitch is lower than /2/, i.e. 8-7 st (/11/), 9-7 st (/11/), 10-7 st (/21/) and 11-7 st (/21/). Therefore NTS perceive the former three PCs as T3. As for the 11-7 st, because of its higher Extreme Endpoint, the rate of perceiving it as T3 goes down. However, being more sensitive to Average Pitch than the NTS, Thai listeners in the present experiment perceive 11-7 st -- and even 12-7 st (/31/) -- as T3. As for the other two /31/ tones, i.e. 13-7st and 14-7st, although the Average Pitch of them each is /2/, they are not to be perceived as T3 by Thai. It could be the effect of bigger Slope according to [5].

What should be noted is that the subjects in the present experiment are different from the listeners in [5] on their language backgrounds. Both THS

and NTS have learned Chinese (as L2) for at least one year, but none of the listeners in [5] was a L2 learner. Different experiments had similar findings. This indicated that THS and NTS in the present experiment might still adhere to the tactics in perceiving tones they had had before they learned Chinese. Since the results were quite different from CNS, it seemed depressing for Chinese teachers that students could make little progress in tone perception with even more than one year 's experience of studying in the target language environment.

Nevertheless, learning experience might have ALREADY influenced their perception as well.

Both THS and NTS have stayed in China for more than one year, they are sure to test in a real life setting what they have learned in the classroom. However, because of the existence of the T3 sandhi rule, the variant low-falling T3 are much more heard than the full and dipping T3 in practical Chinese speech. Therefore, their awareness of "T3=the lowest tone" would be strengthened, and it could eventually be "T3=low-falling tone". In other words, THS and NTS might have perceived the variant T3 instead of the full or its prototype.

5. CONCLUSION

Now we can conclude to answer the three questions inquired at the beginning of this paper.

First of all, as is said in discussion, falling tone perception result of Chinese natives proves that: A separate low-falling tone can not be regarded as T3 in Mandarin. Results suggest indirectly that T3 is basically a dipping tone, just as it had been described in traditional literature.

Secondly, L2 learners perceive low-falling tones as well as mid-falling tones and high-falling tones differently from Chinese. They tend to regard low-falling tone -- one of the variants of T3-- as T3 itself. There can be two reasons for this: (1) habitual perception tactics on pitch, generated from L1 experience; (2) high frequency of variant hearing and using in L2 circumstances.

It is also found that being with or without tone in L1 does influence the perception results:

- L1 being tonal language, Thai people perceive more low-falling tones as T3 than NTS.
- Students with non-tonal L1 are more easily to be affected by stimulus order.

6. APPENDIX

No.	Contours		No.	Contours		No.	Contours	
	H st	L st		H st	L st		H st	L st
1	**19**	**10**	*15*	**12**	**12**	*29*	**13**	**13**
2	**19**	**7**	*16*	**9**	**9**	*30*	**10**	**10**
3	**19**	**12**	*17*	**16**	**16**	*31*	**11**	**7**
4	**15**	**7**	*18*	**13**	**7**	*32*	**15**	**15**
5	**18**	**18**	*19*	**14**	**8**	*33*	**16**	**7**
6	**11**	**11**	*20*	**19**	**9**	*34*	**17**	**17**
7	**19**	**15**	*21*	**17**	**7**	*35*	**8**	**7**
8	**18**	**7**	*22*	**12**	**7**	*36*	**8**	**8**
9	**9**	**7**	*23*	**19**	**17**	*37*	**19**	**19**
10	**16**	**10**	*24*	**19**	**18**	*38*	**19**	**16**
11	**17**	**11**	*25*	**18**	**12**	*39*	**19**	**11**
12	**14**	**7**	*26*	**15**	**9**	*40*	**19**	**13**
13	**7**	**7**	*27*	**19**	**8**	*41*	**19**	**14**
14	**14**	**14**	*28*	**10**	**7**			

7. REFERENCES

[1] Bai, D. 1934. *Pekingnese Tones and Their Changes*. (Quoted) In Luo, C., Wang, J. *Outline of General Phonetics*. Beijing: Science Press. 125-127.

[2] Cao, W. 2008. A comparative study on tone perception. *Proc. PCC2008* Beijing, 203-206.

[3] Cao, W. 2010. Mandarin tone perception. *Zhongguo Yuwen* 6, 536-543.

[4] Chao, Y.R. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

[5] Gandour, J., Harshman, R. 1978. Cross-language differences in tone perception: A multidimensional scaling investigation. *Language and Speech* 21, 1-33.

[6] Gårding, E., Kratochvil, P., Svantesson, J.-O., Zhang, J. 1986. Tone 4 and Tone 3 discrimination in modern Standard Chinese. *Language and Speech* 29(3), 281-293.

[7] Hallé, P., Chang, Y.C., Best, C. 2004. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *J.Phonetics* 32(3), 395-421.

[8] Lin, H. 1996. *Mandarin Tonology*. Taibei: Pyramid Press.

[9] Lin, T. 1979. Phonetics teaching and syllable teaching. *Yuyan Jiaoxue yu Yanjiu* 4, 6-15.

[10] Lin, Y. 2007. *The Sounds of Chinese*. Cambridge: CUP.

[11] Luo, C., Wang, J. 1957. *Outline of General Phonetics*. Beijing: Science Press.

[12] Ma, Y. 1994. Tone errors of the L2 students. *Working Papers of Beijing Normal University* 4, 98-110.

[13] Shen, X.N. 1991. On Mandarin T3 teaching. *Proc. 3rd Int. Con. Chinese Teaching* Beijing, 172-178.

[14] Wang, A. 2006. On Chinese tone feature teaching. *Yuyan Jiaoxue yu Yanjiu* 3, 70-75.

[15] Wang, L. 1979. Some issues on modern Chinese analysis. *Zhongguo Yuwen* 4.

[16] Wang, Y., Li, M. 2010. The effects of tone pattern and register in perceptions of Tone 2 and Tone 3 in Mandarin. *Acta Psychologica Sinica* 9, 899-908.

[17] Xu, Y. 1999. Effects of tone and focus on the formation and alignment of f0 contours. *J. Phonetics* 27, 55-105.

[1] Due to the limit of papers, only the results of end-fixed falling tones were reported here.