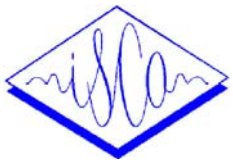

Speech Prosody 2004
Nara, Japan
March 23-26, 2004

A Preliminary Analysis of Focus and Ending in Chinese Intonation

Wen Cao

Center of the Chinese Proficiency Test, Beijing Language and Culture University, China
Department of Chinese Language & Literature, Peking University, China
tsao@blcu.edu.cn

Abstract

The present paper investigates the issue of focus types and the ending situations in Chinese declarative and interrogative intonations. Differing only in focus of each, 5 statement sentences and 5 interrogative sentences with the same words and the same syntax in Chinese are designed for the experiment. A male Standard Chinese speaker reads them in random order three times. The results show that there are two types of foci in Chinese intonation: H* and L*, subject to Chinese lexical tone features. Besides, both H* and L* can be identified by the D-value between the two H peaks. In particular, the shift of the focus location seems to have effects on boundary tone H%. The findings in this research seem to be able to make Chinese intonation curves somewhat predictable.

10.21437/SpeechProsody.2004-25

1. Introduction

There are four lexical tones in Standard Chinese, which have been traditionally described as /55/, /35/, /214/, and /51/ for Tones 1, 2, 3 and 4 respectively [1]. But in fact, /212/ and /211/ are more frequently uttered in speech for Tone 3, as mentioned in [1, 2]. That is why some phoneticians also analyze the four tones as H, R, L, F or H, LH, L, HL as well [3, 4, 5]. There is an atonic syllable, or "neutral tone" in Chinese, too [1, 5]; and it could be H or L. For example, when it follows H, LH, or HL, which have H feature, it is L; when it follows L, it becomes H.

Tone and intonation in Chinese have interactive effects on each other, and this has attracted much attention in phonetic researches [1, 3-14]. In their research, some authors took focus effect into account as well. Gårding modeled Chinese intonation with 'grids', qualitatively marking time-varying pitch ranges, and lexical tones fit into these ranges. She declared that focus was a combination of expansion and compression of such 'grids' [8]. Kratochvil set up six stages for modifications of suprasegmental features in Chinese sentence without considering focus at first in [6], but later in [7], he proposed four intonational devices in Chinese: channeling, tempo, focusing, and intonation carriers. He made a point similar to Gårding's, saying that the focus is reached by a gradual enlargement followed by a gradual diminishing of the channel [7]. Xu found that while the lexical tone acted as the most important factor for the f0 local contour of the syllable, focus modulated the global shape of the f0 curve, and f0 range after the focus was lowed as well as compressed [4]. J. Shen argued that there were two independent prosodic systems in Chinese intonation: the top line of a pitch contour for accent or focus and base line for speech act (declarative or interrogative) [9, 10].

From these analyses it is clear that focus has great effects on Chinese intonation, but there is still something unclear. For example, most researchers above stated that the focused syllable had expanded pitch range. Then, two questions could be raised: (1) Do all tones have the same quantitative expansion if they are emphasized? Or to say, are there any identifying properties of different tone foci? (2) In addition to the effect of post focus compression, is there any kind of influence of focus on boundary tones? The aim of this paper is to put forward a preliminary study intending to answer these two questions.

2. Method

2.1. Materials and procedure

Two groups of short SVO sentences of Chinese were designed for the experiment: one for statements, and the other for morphosyntactically unmarked yes-no questions. Every stimulus sentence is in fact composed of the same 3 words, totaling 5 syllables in the series: Tone3-Tone4-Tone3-Tone4-Neutral Tone, which could be simply featured as L-HL-L-HL-L. Lexically, the sentence is 马力买柚子'Ma3 Li4 mai3 you4zi' ('Ma Li buys shaddock'). But each group has 5 such sentences with different focus locations:

(1) Broad focus (FB). No intended emphasis on any word.
(2) Subject focus (FS). The accent falls on '*Li4*', the second syllable of the first word, carrying the given name.
(3) Verb focus (FV). The accent falls on the monosyllable word '*mai3*' (to buy).
(4) Object focus (FO). The accent falls on '*you4*', the first syllable of the last word (i.e. shaddock).
(5) Initial focus (FI). The accent falls on '*Ma3*', the very beginning syllable carrying the family name.

These sentences might be described with tone features as the following in table 1.

Table 1: *Part of the stimuli sentences with different foci. (the foci are in italics)*

	Tone series	Chinese	English meaning
FB	*L-HL-L-HL-L*	*马力买柚子*	*Ma Li buys shaddock*
FS	*L-HL*-L-HL-L	*马力*买柚子	*Ma Li* buys shaddock
FV	L-HL-*L*-HL-L	马力*买*柚子	Ma Li *buys* shaddock
FO	L-HL-L-*HL-L*	马力买*柚子*	Ma Li buys *shaddock*
FI	*L*-HL-L-HL-L	*马*力买柚子	*Ma* Li buys shaddoc*k*

In general, Chinese word stress falls on the final syllable. However, for the word containing a neutral tone, the stress falls on the forgoing syllable [1]. Normal in Chinese, the sentence accent is realized by word stress, but then any syllable except a neutral tone can be stressed to express unconventional meaning. That is why we see in the present paper that the nuclear prominences are distributed in the second syllable of the subject focus, the third for the verb, and the fourth for the object. As for the initial focus, it has special

emphasis on the family name, meaning 'it is Ma Li, not Sun Li, Liu Li, etc. who does something'.

The stimuli were read and recorded by a middle-aged male college teacher who speaks and teaches Standard Chinese. The recordings were carried out in a quiet environment, using Maya 5.1 USB Audiotrack SoundBlaster and Somic earphone which were connected to a laptop Toshiba Satellite 2410. The subject read the 10 sentences in random order three times. When a reading error occurred, the speaker was asked to read the sentence once more. The material comprises 30 correct utterances.

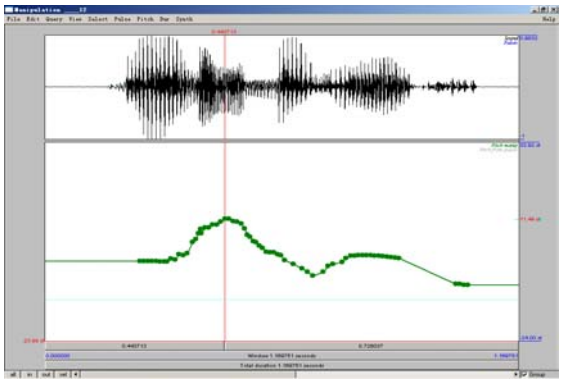

Figure 1: *Subject focused statement 'Ma Li mai youzi.' The nuclear accent falls on the second syllable 'Li', as the curser marks it.*

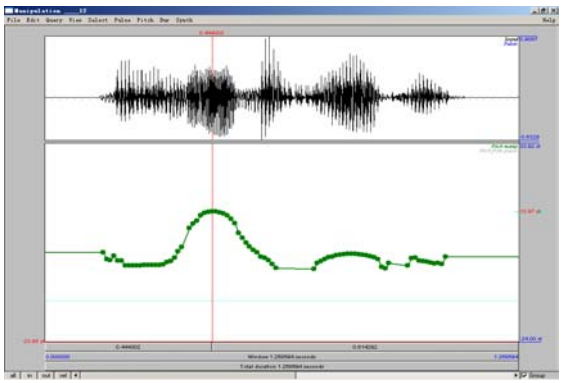

Figure 2: *Subject focused question 'Ma Li mai youzi?' The nuclear accent falls on the second syllable 'Li', as the curser marks it.*

2.2. Analysis

The data was saved as digital wave files at 22,050 Hz sampling rate, 16 bits resolution, mono, and was analyzed with the help of Pratt [http://fonsg3.let.uva.nl/praat/], as shown in Figure 1 and Figure 2. F0 values of each utterance were measured at 6 feature points: the beginning point, the lowest points of the first and third syllables, the highest points of the second and fourth syllables, and the ending point. The measurements were taken in semitones (st, f-reference=100 Hz) in order to make a perceptually relevant comparison between all data.

D-values between some points were calculated in order to check the focus expansion and following compression. The duration and intensity of each syllable was also recorded. However, the present paper is only investigating the F0 influence of focus shift on intonations, leaving the others for further work.

3. Results

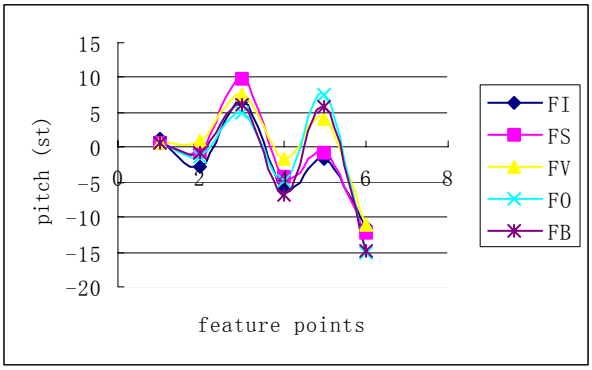

Figure 3: *Statement trends with different foci.*

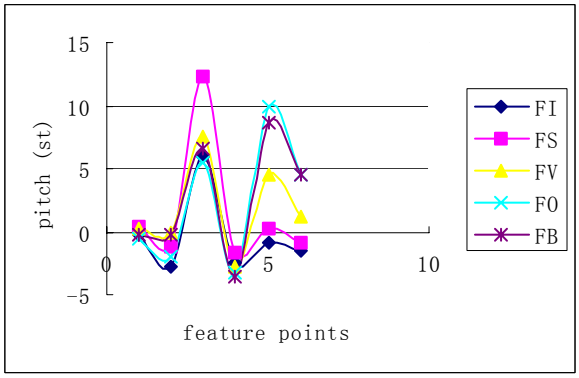

Figure 4: *Question trends with different foci.*

Figures 3 & 4 compare traced feature points for statement and interrogative intonations with different foci.

In fig. 3 we find that the final point has the lowest pitch in each trace, and all traces have a declination on the L tones. The results have somewhat proved the existence of a downtrend base line in statement intonation, whatever the focus is.

On the contrary, the top line changes according to the focus. As mentioned above, in order to check the focus expansion and following compression, we calculated some D-values. They could be categorized into several types: D1, between points 1 &2; D2, between points 2 & 3; D3, between points 3 & 4; D4, between points 4 & 5; D5, between points 5 & 6; D6, between points 3 and 5. Unfortunately, among these types of D-values, only D6 was found to have constancy and adequate explainability. Comparisons using paired samples t-tests show a significant difference in D6 values between all pairs except the pair of FB and FO ($t(2)=3.53$, $p>0.05$). This indicates that: (1) the D-value between pitch peaks is a good

metric to locate different Chinese statement intonation trends that have different foci; (2) the broad focus utterance (FB) and the object focused utterance (FO) are perceptually the same, although they have slight acoustical differences shown in fig. 3. The latter phenomenon has been found in both Chinese and English intonations [4, 9, 10, 17, 18, 19, 20].

Moreover, in fig. 3, all statement curves start with almost the same value at the f0 beginning point, without significant effect of focus on it ($F(4,10) = 1.20$, $p>0.05$). As for the final point, although four trends end with different values in average, they do not show significant difference ($F(4,10)= 2.03$, $p>0.05$), either.

Seen in fig. 4, the final point is obviously not the lowest any longer in question trend, but it is still lower in pitch than the forgoing HL tone in the trace, keeping its lexical L feature. Besides, it seems that there is a significant effect of focus location on the final point ($F(4,10) =13.05$, $p<0.01$). Paired samples t-tests show a significant difference between the ending f0 values of FI and FV ($t(2)=-3.85$, $p<0.05$), FI and FO ($t(2)=-4.88$, $p<0.05$), FI and FB ($t(2)=-5.86$, $p<0.05$), FS and FO ($t(2)=-9.43$, $p<0.05$), FS and FB ($t(2)=-4.60$, $p<0.05$), but not between those of FI and FS ($t(2)=-1.02$, $p>0.05$), FS and FV ($t(2)=-1.68$, $p>0.05$), FV and FB ($t(2)=-3.54$, $p>0.05$), FO and FB ($t(2)=-0.68$, $p>0.05$). It can be interpreted as follows: (1) the broad focused and the object focused utterances may have the same ending; (2) when the focus shift to the right neighboring position, e.g. from initial the first syllable to subject the second syllable, or from subject the second syllable to the verb the third syllable, etc., the interrogative utterance may have the similar ending as its original.

Just as shown in statements, the results of the interrogatives also show that the focus location has significant influence on the D-values between two H peaks ($F(4,10) = 52.12$, $p<0.01$). Paired samples t-tests show a significant difference of D-values between all pairs except FB and FO ($t(2)=-1.24$, $p>0.05$). This indicates that: (1) the broad focus interrogative (FB) and the object focused interrogative (FO) are still almost the same in intonation; (2) the D-value between pitch peaks plays the decisive role in locating different Chinese interrogative intonation trends that have different foci.

Comparing between the data of the statements and questions, we find that the question curves have lower beginning pitch than the statement ones (see table 2).

Table 2: *Pitch values of beginning points in statements and questions.*

	in statements	in questions
FB	1.17st	-0.17st
FI	0.68st	0.51st
FS	0.84st	0.17st
FV	0.84st	-0.53st
FO	0.84st	-0.17st
Average	0.88st	-0.03st

In this experiment, we have neither seen increased initial f0 in the 'unmarked question' curve brought up by X. Shen in [12], nor 'paralleled higher interrogative intonation curve' set forth by Yuan, et al. [14].

4. Discussion

Although there are different opinions on ToBI system use in Chinese intonation labeling [5, 14, 15, 16], I still intend to apply it in this paper. According to my own observation, it is descriptively convenient for the research of Chinese intonation.

4.1. H* and L*

It is well established that there are H* and L* in English, which are subject to attitudinal expressions [21, 22, 23, 24]. There are H* and L* in Chinese, too. However, unlike in English, H* or L* in Chinese is not decided only by intonation expression, but also by lexical tones. In brief, only Tone 3 that is distinctively featured with L could realize L*. In the present paper, the initial focused and the verb focused utterances show L*, because their foci are on a syllable of L each; while the subject focused and the object focused utterances show H*, because their foci are on a syllable of HL (non-L) each.

As mentioned above, many authors believe that a focused syllable in Chinese intonation has an expanded range of pitch: the H becomes higher, and the L becomes lower [1, 4-8, 17, 18]. It seems that we can find some pitch range expansion on some focused syllables in this experiment. For example, in either the FI statement or the FI interrogative curve, the focused initial syllable which has L tone is the lowest among all at point 2 (see fig.3 & fig. 4), and the focused second syllable which has H tone is the highest at point 3. However, we also see that the focused third syllable in the FV statement curve, which has L tone too, becomes the highest among all at point 4. Then, there must be some other metric consistent to identify accents in Chinese intonation.

Some researchers argue that the post-nuclear syllable be raised when the L tone is accented [9, 10, 17]. That seems more likely the truth, but it is still inconsistent. Statistics show that the D-value between the values of point 2 and point 3 (D3, as mentioned before) in statements do not have significant difference between FI and FS ($t(2)=-2.43$, $p=0.136>0.05$), or between FI and FV ($t(2)=3.54$, $p=0.07>0.05$).

I suggest D-values between the two H peaks play the most important role in identifying accents. As a rule in Chinese phonology, two L tones can not appear in neighboring [25, 26, 27]. So in most cases, there must be H peaks in Chinese Utterances. Generally, H* occupies the highest pitch point, but we could not say that the highest f0 means H*. For example, both in the initial focused utterance and in the mid-focused utterance in figures 3-4, the highest peak does not qualify as H*. Results have already shown that we could clarify the nature of the peak by D-values. According to present results, the D-values could be categorized into 4 types: Big, Middle, Small, and Negative or Zero. The Big D-value is about one octave, as in the subject focused utterance, precisely, 11st in the statement and 12st in the question. The Middle D-value is 6 or 7 semitones, as in the initial focus utterance. The Small D-value is about 3st, as in the middle verb focused utterance. The Negative or Zero D-value is found in the broad focus or the object focused utterance. So, the suggestion is that if the D-value between the first peak and the next peak is Big, the first peak acts as H* itself; if Middle, the low tone after it acts as L*; if Small, the low tone before it is L*, and if Zero or Negative, the next H tone occupies the nuclear accent.

4.2. H% and L%

Unlike H* and L*, H% and L% are not bound to lexical tones in Chinese. They are purely labels of distinctive intonation endings. Theoretically, both the L tone and the tone with H, like HH/LH/HL, could have L% or H%. For instance, the final syllable "zi" in all utterances is marked L as an atonic syllable following HL tone, but in the object focus question, it is raised to about as high as the first H. Perhaps, the curve could be simply described like:

L—HL—L—HL—L
* H%

or

L—HL—L—H*L—L (H%)

Taking carry over effect suggested by Xu [4] into account, we could describe the intonations of the 10 utterances in this experiment as the following:

L*—H(L)—L—H(L)—L (L%)	FI-statement
L—H*(L)—L—H(L)—L (L%)	FS-statement
L—H(L)—L*—H(L)—L (L%)	FV-statement
L—H(L)—L—H*(L)—L (L%)	FB/FO-statement
L*—H(L)—L—H(L)—L (H%)	FI-statement
L—H*(L)—L—H(L)—L (H%)	FS-statement
L—H(L)—L*—H(L)—L (H%)	FV-statement
L—H(L)—L—H*(L)—L (H%)	FB/FO-statement

Furthermore, results also indicate a influence of focus shift on boundary tones. Although final pitches in differently focused statement curves do not show a significant difference, final pitch values in the questions do. Accent type being the same, L* or H*, the later the focus appears, the higher the H%. In addition, it can be seen from Figure 4 that all trends have the tendency of final increase, to a value higher than the forgoing L, but lower than the neighboring H.

5. Conclusion and further research

In sum, I found some rules in Chinese intonation as following:

(1) Chinese statement intonation has downtrend base line and the ending pitch is almost consistent.
(2) The D-value between the highest two peaks is a significant metric to identify different foci.
(3) Focus location affects the ending pitch value in Chinese question intonations. Type being the same, H* or L*, the later the focus appears, the higher the utterance ends.

These findings seem to be able to make Chinese intonation curves somewhat predictable.

It is planned to investigate the foci and endings in more and longer utterances and to develop a somehow complete model of global trends in Chinese intonation. Further research would also have perceptional experiments.

6. Acknowledgements

This research was supported by NOCFL grant HBK01-05/17. I would like to thank Prof. Xiaoqing Xie and Dr. Jimin Wang for advice on statistics.

7. References

[1] Chao, Y. R., 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
[2] Lin, T., 1990. Hanyu yunlv tezheng he yuyin jiaoxue [Prosodic features of Chinese and phonetics teaching], in *Collected Works on Phonetics Exploring*, 158-168. Beijing: Beijing Language Institute Press.
[3] Wu, Z., 1982. Rules of intonation in Standard Chinese. In *Reprints of Papers for the working Group on Intonation ICL* 13. Tokyo, 95-103.
[4] Xu, Y., 1999. Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetic*, 27, 55-105.
[5] Peng, S. et al., A pan-mandarin ToBI. http://deall.ohio-state.edu/chan.9/MToBI.htm
[6] Kratochvil, P., 1968. *The Chinese Language Today*, London: Hutchinson & Co. (publishers) Ltd..
[7] Kratochvil, P., 1998. Intonation in Beijing Chinese. In *Intonation Systems: A Survey of Twenty Languages*, Hirst and Cristo (eds.). Cambridge: CUP, 417-431.
[8] Gårding, E., 1987. Speech act and tonal pattern in Standard Chinese:constancy and variation. *Phonatica* 44: 13-29.
[9] Shen, J., 1985. Beijinghua shengdiao de yinyu he yudiao [Pitch range of tone and intonation in Beijing dialect]. In *Beijing Yuyin Shiyanlu*, Lin and Wang (eds)., Beijing: PUP, 73-130.
[10] Shen, J., 1994. Hanyu yudiao gouzao he yudiao leixing [Intonation structure and intonation types of Chinese]. *Fangyan*, 3, 221-228.
[11] Shih, C., 1988. Tone and intonation in Mandarin. *Working Papers of the Cornell Phonetics Laboratory* 3: 83-109.
[12] Shen, X., 1990. *The Prosody of Mandarin.* Berkley: UCP.
[13] Liao, R., 1994. Pitch Contour Formation in Mandarin Chinese: A Study of Tone and Intonation. Ph. D. dissertation, OSU, Ohio.
[14] Yuan, J. et al., 2002. Comparison of Declarative and Interrogative Intonation in Chinese , http://prosodies.org/papers/2002/interrogative2002.pdf
[15] Li, A. et al., 1999. A National Database Design and Prosodic Labelling for Speech Synthesis. Paper presented at the *Oriental COCOSDA'99*, Taiwan.
[16] Wight, C., 2002. ToBI or not ToBI? http://www.lpl.univ-aix.fr/sp2002/pdf/wightman.pdf
[17] Kwon, Y., 2001. Sentence Stress in Chinese Intonation. Ph. D. dissertation, PkU, Beijing.
[18] Jin, S., 1996. An Acoustic Study of Sentence Stress in Mandarin Chinese. Ph. D. dissertation, OSU, Ohio.
[19] Ladd, D., 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
[20] Cruttenden, A., 1997. *Intonation*. Cambridge: CUP.
[21] Pierrehumbert, J., 1980/87. The phonology and phonetics of English intonation. Ph. D. dissertation, MIT. Published by Indiana University Linguistics Club (1987)
[22] Beckman, M.; Elam, G., 1997. Guidelines for ToBI labeling, version 3.0. The Ohio State University Foundation.
[23] Liberman, M.; Pierrehumbert, J., 1984. Intonational invariance under changes in pitch range and length. In Aronoff Oehrle, eds. *Language Sound Structure: Studies in Phonology Presented to Morris Halle by His Teacher and Students*. The MIT Press.
[24] Tench, P., 1996. *The Intonational Systems of English.* London: CASSELL.
[25] Cheng, C., 1973. *A Synchronic Phonology of Mandarin Chinese.* The Hague: Mouton.
[26] Duanmu, S., 1990. A formal study of syllable, tone, stress and domain in Chinese languages. MIT Ph. D. dissertation.
[27] Lin, H., 1996. *Mandarin Tonology*. Taipei: Pyramid Press.