THE INTERACTION OF INTONATION BOUNDARY AND LEXICAL TONES UNDER TWO FOCUS CONDITIONS IN THE WUZHI DIALECT OF CHINESE

Yuzhu Zhu[1,3], Zhiqiang Li[2], Aijun Li[3,4]
[1]Beijing International Studies University (BISU), China
[2] University of San Francisco, United States
[3]Institute of Linguistics, Chinese Academy of Social Sciences, China
[4]University of Chinese Academy of Social Sciences (UCASS), China
zhuyuzhu@bisu.edu.cn, zqli@usfca.edu, liaj@cass.org.cn

ABSTRACT

This study investigates the mechanism of boundary marking in tone languages through the interaction of intonation boundary and lexical tones under two focus conditions in Wuzhi. We analyze the prosodic realizations of the lexical tones at the right edge of string-identical statements and echo questions, both spoken with a broad or narrow focus that coincides with the nuclear stress on the final prosodic word. Results show that the final f_0 lowering is a robust indicator of statements, resulting in a falling contour in most tonal contexts. Questions display more intricate patterns, in which the H% target drives the surface tonal patterns in the two convex and two falling tones at the intonation boundary. Furthermore, the intonation boundary is signaled by longer duration and higher pitch register in questions than in statements. We conclude with a typological account of boundary marking for the simultaneous realization of tone and intonation in Chinese.

Keywords: intonation boundary, lexical tones, broad and narrow focus, Wuzhi

1. INTRODUCTION

In tone languages, tone and intonation are both expressed by the acoustic property of fundamental frequency (f_0) and its perceptual correlate of pitch. An essential aspect of studying intonation of Chinese boils down to how pitch contours that are associated with lexical tones are manipulated both locally and globally in response to different intonational meanings and focus conditions (See [17] for an overview). Of particular interest in this line of research is how lexical tones in the utterance-final prosodic word interact with the intonation boundary that serves to distinguish utterance types and emotions or affective states of utterances.

In one of the seminal works of tone and intonation in Mandarin Chinese, Chao [3] identified two modes of interaction in which lexical tones and intonational tones can be realized either simultaneously or successively. While the simultaneous addition may occur in any part of the utterance, the successive addition is often observed at the intonation boundary. The two modes of interaction shape terminal pitch contours differently at the boundary. In Mandarin, for example, the boundary tones associated with declarative and interrogative utterances modify the register or slope of the pitch contours at the right edge of the intonation phrase such that they do not alter the canonical shapes of the lexical contours [9, 10, 11]. In other words, the falling tone – Tone 4 in Mandarin – maintains its falling contour in questions, probably with a smaller pitch range, while the rising tone – Tone 2 in Mandarin – also retains its rising contour in questions, but its pitch register is likely elevated and expanded. The successive addition is found in the question intonation of Cantonese when a final rise is added to the pitch contour of the final tone in the utterance [6, 18], or in utterances that convey strong emotions such as disgust or anger in Mandarin [8].

In an attempt to understand how tone and intonation interact at the intonation boundary from a typological perspective, we ask the general question of how the two modes of interaction outlined above play out in different tonal systems other than the relatively well-studied ones in Mandarin or Cantonese. The present study compares the prosodic realizations of lexical tones at the intonation boundary in statements and echo questions (without question particles) under two focus conditions in a Jin dialect. We seek to address the following questions:

1. Are there any differences in the realization of lexical pitch contours at the intonation boundary in string-identical statements and echo questions?
2. How is the realization of pre-boundary lexical contours shaped by the introduction of narrow focus in comparison with broad focus?
3. How do different lexical tone systems respond to boundary marking differently?

Wuzhi is a Jin dialect of over 700,000 speakers in the north-west of Henan province, which borders the vast Mandarin speaking region. It is reported to have five lexical tones: two falling tones, two convex tones, and a checked tone [19]. Jin dialects, unlike Mandarin dialects, retain the checked tones in their lexical tone system, and Wuzhi serves as a typical example. Experimental work on their intonation has been lacking. In this study, we start with the lexical tone

system and the tone sandhi patterns before tackling intonational features, as Wu [11] suggested in his study of intonation in Mandarin. It is important to note that tonal alternations trigged by a phonological process is not intonational. A well-known tone sandhi is the Tone 3 sandhi in Mandarin: when two syllables in Tone 3 are juxtaposed, the first syllable changes to Tone 2 [3]. In this paper we focus on data and results of the intonation study, but will draw on findings from the tonal study too, which will be dealt with in more detail in [20].

2. METHODS

2.1. Materials and speakers

The recording materials consist of three parts: 64 monosyllables, 837 disyllables and 25 sentences. The word lists were constructed based on the historical origins of tonal categories in the Middle Chinese and tonal combinations. 25 disyllabic combinations out of five lexical tones were embedded into a carrier sentence "张军初三飞___" (Zhangjun flied to ___ on the third day of the month), in which "飞" is the verb "fly to". All 25 combinations are city or country names that occur in the sentence-final object position and receive the nuclear stress when the sentence has a broad, neutral focus or when a narrow focus is placed on the object noun [5]. Three mini-dialogues were created similarly for each of the 25 target sentences, as follows:

1. A: 你刚刚说什么？ (What did you just say?)
   B: 张军初三飞英国。(Zhangjun on the third day of the month flied to England.)
2. A: 你刚刚问什么？ (What did you just ask?)
   B: 张军初三飞英国? (Zhangjun on the third day of the month flied to England?)
3. A: 张军初三飞【英国】? (Zhangjun on the third day of the month flied to **England**?)
   B: 张军初三飞【法国】。(Zhangjun on the third day of the month flied to **France**.)

In Dialogue 1, the response to the question receives a broad focus. In Dialogue 2, B repeats the question in response to A. The response in B has the form of an echo question (without question particles) that has the pragmatic effect of expressing surprise, requesting confirmation, or making an indirection negation [14]. An echo question in Chinese have the same form as a statement. In Dialogue 3, the object noun receives a narrow focus in A and B when the focus of the question is on the location noun.

12 speakers of the Wuzhi dialect, six males and females, participated in the experiment. They reported no hearing or speaking difficulties. Their average age is 52 (SD = 15).

2.2. Recording and data extraction

The recording was carried out in the quiet bedrooms at the participants' homes. An in-house recording program X-recorder [13] was used to collect data on a PC laptop computer with a "Sound Blaster Play! 3" external sound card and a Philips SHM 1900 headset. The sampling rate and resolution were set to be 22.05 KHz and 16 bits respectively.

For the monosyllables, 11 speakers read 64 syllables three times, generating 2112 recordings (64×3×11). For the disyllables, 11 speakers read 837 syllables once for a total of 9207 recordings (837×11). 12 speakers were asked to read the three mini-dialogues for each target sentence, producing four utterance types: a statement and question in broad focus (1B and 2B), and a statement and question in narrow focus (3B and 3A). We ended up with 1200 utterances (12 speakers, 25 target sentences, four utterances for each target sentence). One speaker's data were excluded from the analysis because of their limited ability to read Chinese characters.

A Praat-based [1] automatic segmentation and transcription tool was employed to provide segmental annotations for syllabic and subsyllabic elements (initials and finals) [12]. Pitch contour was extracted for each syllable, with spurious pitch cycles manually modified. 11 pitch values were extracted at equal intervals on each syllable final. Durations of syllable initials and finals were also obtained. An example of segmentation and pitch extraction is shown in Fig 1.

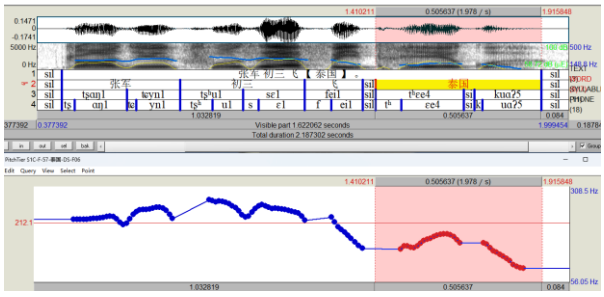

Figure 1: Segmentation, transcription and pitch tracks.

3. RESULTS

3.1. Lexical tone system in Wuzhi

Table 1 provides a summary of tone values on a five-point scale [2] and their corresponding tonal features. Tones 1 and 4 are convex tones that have three that tonal targets MHM or LHL, differentiated by pitch registers. Meanwhile, Tones 2 and 3 are falling tones in low and high registers, respectively. Lastly, Tone 5 is a mid-level checked tone.

Tones	T1	T2	T3	T4	T5

shapes	/343/	/31/	/51/	/242/	/3/
features	MHM	ML	HL	LHL	M

Table 1: Lexical tones and features in Wuzhi.

Fig. 2 illustrates the major variants averaged over 11 speakers. All tones in the citation form exhibit tremendous pitch movement over the course of the duration, even in the checked tone. Comparing the two convex tones, Tone 1 has a relatively longer falling tail while Tone 4 has a much longer rising portion. The two falling tones tend to level off or even go up in pitch toward the end of the syllable in some variants. Most variants conform with tonal shapes in Table 1, especially when they occur in disyllabic combinations. The falling tail is more likely to occur when the tone is in the final position.

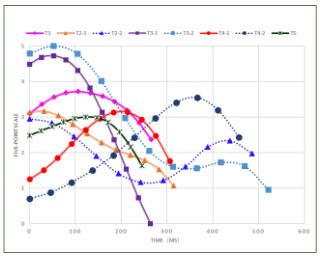

Figure 2: Averaged pitch contours of tones in Wuzhi.

3.2. Disyllabic patterns: Tone 4

The disyllabic patterns are given in Fig. 3. In each of the panels, X is one of the five tones. Since Tone 5 patterns with Tone 1, "X+T5" is omitted. The f_0 plots were generated after averaging the f_0 values at each of the 11 points where measurements were taken.

Tonal contrasts are preserved in the disyllabic patterns. For example, panels labelled "X+T2" and "X+T3" show that the contrast between the Tones 2 and 3 is clearly maintained in the word-final position. Their register distinction is also retained in the word-initial position, as in the panel of "X+T1".

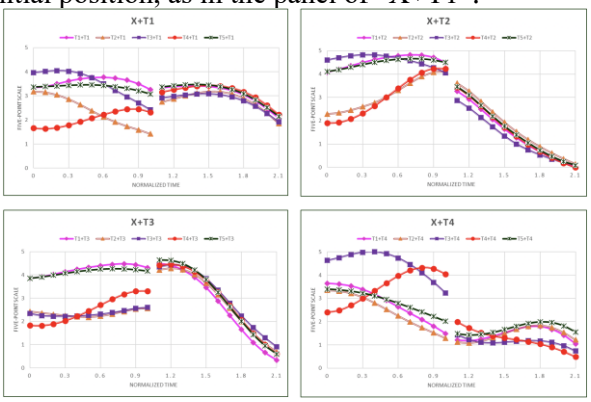

Figure 3: Disyllabic patterns of tones in Wuzhi.

3.3. Effects of intonation boundary on lexical contours

Regarding the effect of intonation boundary on lexical contours, three situations can be identified in statements under two focus conditions. Firstly, Tone 2 and Tone 3 are realized as falling contours before the intonation boundary, regardless of the focus condition. In this context, the contour shape and the register specification are faithfully maintained as in disyllabic combinations. Secondly, Tones 1 and 4 are realized as a low or low fall in the utterance-final position. Both tones are influenced by the ending f_0 of the preceding tone, which suggests the presence of a coarticulatory effect [15]. This means that the tonal target for both tones is most likely L. Thirdly, in Wuzhi, Tone 5 is a checked tone whose duration is generally shorter than long tones. However, its surface contour shape appears to align with that of Tones 1 and 4 in statement intonation under both focus conditions. An example is provided in Fig. 4. The final syllable is in Tone 1 and the penultimate syllable is in one of the five tones labelled as X. In almost all statements produced by the speakers, the final f_0 lowering has been a robust indicator of statement intonation, resulting in a falling contour in most tonal contexts.

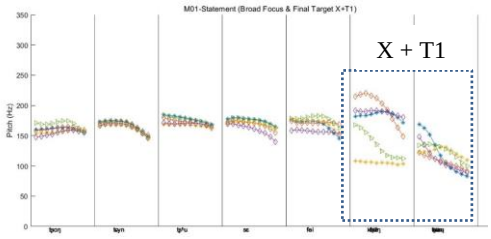

Figure 4: Pitch contours of a statement in broad focus.

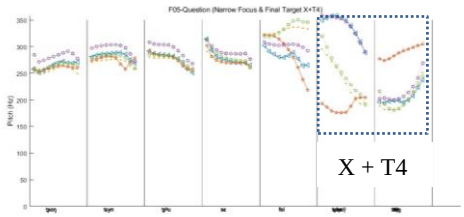

Figure 5: Pitch contours of a question in narrow focus.

Questions display more intricate patterns under the two focus conditions. Tones 2 and 3 are realized with a rising tail toward the end of the syllable, leading to a somewhat concave f_0 configuration in some cases. A very intriguing case is seen in Tones 1, 4 and 5 when the boundary is associated with the question intonation. As Fig. 5 shows, when the final prosodic word was focused, Tone 4 was realized as a rise in the question. The final rise in Tones 1, 4 and 5 appears to be the predominant pattern across speakers.

The pitch register specification participates in the computation of the surface tonal contours, as we see lower rise in Tone 4, but higher rise in Tones 1 and 5.

3.4. Prosodic cues for utterance types and focus conditions

GLMM was conducted on the pre-boundary tones in the two utterance types under two focus conditions in order to determine the prosodic cues that can be used effectively to distinguish statements and questions, and different focus types. It was found that the intonation boundary was signaled by longer duration (Fig. 6a, $p<0.05$) and higher pitch register (Fig. 6b-d, $p<0.05$) in questions than statements. The maxima, minima and average values of f_0 in the final syllable are all significantly higher in questions than statements ($p<0.05$ for all). The pitch range of the final syllable is significantly smaller in questions than statements (Fig. 6e).

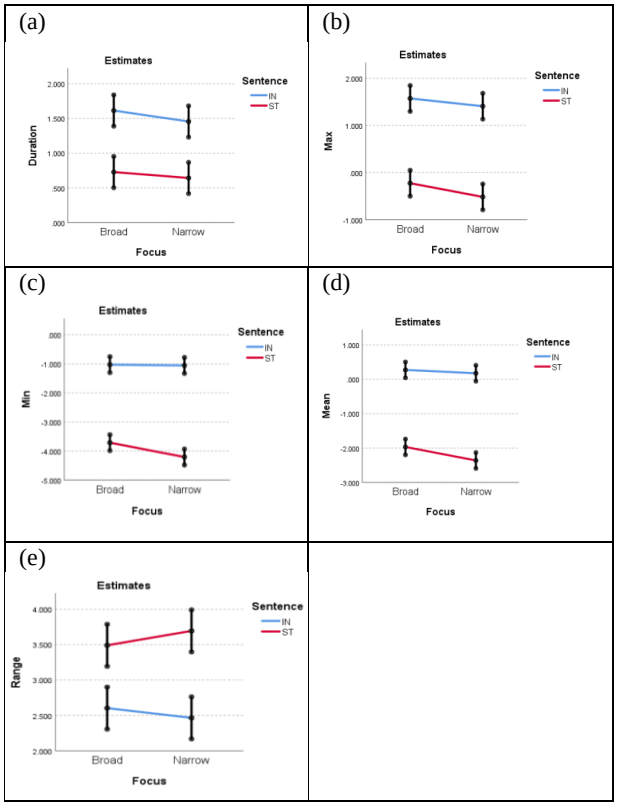

Figure 6. GLMM analysis results for five prosodic cues across focus conditions and utterance types.

The difference in duration under the two focus conditions is not significant (Fig. 6a, $p=0.558$). However, we notice that the domain of the narrow focus is the prosodic word consisting of the final and penultimate syllables, not just the final syllable.

In statements, the pitch range of the final syllable was found to be larger when it was under the narrow focus. However, the opposite seems to be true in questions (Fig. 6e).

Significant interactions between utterance type and focus type were found in f_0 minima ($p=0.027$), and marginal significant interactions were found in average f_0 values ($p=0.056$), but no interactions for other cues ($p>0.05$).

4. DISCUSSION

In tone languages, intonation meanings are expressed by modifying pitch contours associated with lexical tones, especially at the intonation boundary. The boundary marking mechanism in tone languages appears to be subject to both universal and language-specific constraints of intonation. For example, in numerous languages, the question tune ends with a rising pitch movement, represented by the L-H% boundary tone in the AM model of intonation [7]. In languages where word meanings are expressed in tonal contrasts, intonation features typically do not alter tonal identity, as we have seen in Mandarin [9, 10, 11]. Nonetheless, a boundary tone can manifest itself even without a segmental host, such as in the example of Cantonese question intonation without the particle [6, 18].

In Wuzhi, the H% appears to drive the surface tonal patterns at the intonation boundary in the question intonation, and the mechanism here is different from Mandarin and Cantonese. In the case of Tones 1, 4 and to some extent Tone 5, the occurrence of the final rise is only associated with echo questions. The disyllabic tone sandhi data [20] suggest that the convex tone is better represented as R+L. Using Tone 4 as an example, it appears as a rising tone in the first syllable and a low tone in the second. While the H% target in question intonation does not directly manifest itself as it does in Cantonese, it does influence the emergence of the R tone. In the case of the two falling tones, the rising tail toward the end of the syllabic host provides a cue for the H% boundary tone.

The concurrent realization tone and intonation at the intonation boundary can described in terms of simultaneous addition and successive addition [3]. Our anslysis suggests that Wuzhi displays the successive addition pattern in both statements and questions, but the specific mechanism differs from Mandarin and Cantonese.

In conclusion, we have seen that lexical pitch contours at the intonation boundary are modified differently in statements and echo questions, due to the different boundary tones aligned with the boundary. Focus, especially narrow focus, generally modifies pitch contours in the same way as in other tone languages [16].

5. ACKNOWLEDGEMENTS

This work was supported by the Project of Cultural Experts and "Four Batches" of Talents awarded to Aijun Li and the CASS Innovation Project. We would also like to thank Ziyu Xiong, Liang Zhang and Xiaoyan Zhang for their contributions to data collection and comments.

6. REFERENCES

[1] Boersma, P., & Weenink, D. 2017. *Praat: Doing phonetics by computer* (version 6.0.24). Amsterdam: University of Amsterdam.

[2] Chao, Y. R. 1930. A system of 'tone-letters'. *Le Maître Phonétique*, 30, 24–27.

[3] Chao, Y. R. 1933. Tone and Intonation in Chinese. *Bulletin of the Institute of History and Philology Academia Sinica*, 4(2), 121-134.

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

[5] Duanmu, S. 2007. *The Phonology of Standard Chinese* (2nd). New York: Oxford University Press.

[6] Ge, C., Li, A. 2019. Intonation of Cantonese interrogative sentences with and without sentence final particle. *Proc. 19th ICPhS* Melbourne, 2455–2459.

[7] Ladd, D. R. 2008. *Intonational Phonology*. Cambridge: Cambridge University Press.

[8] Li, A., 2013. Successive addition boundary tone of Chinese emotional intonation: production and perception. In: Cheng, C. (ed.), 206-308.

[9] Lin, M. 2006. Interrogative mood and boundary tone in Chinese. *Journal of Chinese Linguistics*. Vol. 4. 364-384.

[10] Lin, M., Li, Z. 2011. Focus and boundary tone in Chinese intonation. *Proc. 17th ICPhS* Hong Kong, 1246–1249.

[11] Wu, Z. 1990. Basic contour patterns of intonation in Mandarin Chinese. In: *A Festschrift for Wang Li*. Beijing: Commercial Press.

[12] Xiong, Z. 2019. xSegmenter: a tool for automatic segmentation and labelling. *Journal of Chinese Phonetics* 11. 27-34.

[13] Xiong, Z. "X-recorder: recording tool for speech corpus," https://mp.weixin.qq.com/s/ewBC2hDg4seuV1YWoPVi-A.

[14] Xu, J. 2016. Linguistic forms of echo questions. In: Eom I., Zhang, W. (eds.), *Language Evolution and Changes in Chinese*. Project on Linguistic Analysis, University of California, 206-215.

[15] Xu, Y. 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics* 25. 61-83.

[16] Xu, Y. 1999. Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics* 27. 55-105.

[17] Xu, Y. 2019. Prosody, tone and intonation. In: Katz, W. F., Assmann, P. F. (eds), *The Routledge handbook of Phonetics*. New York: Routledge, 314–356.

[18] Zhang, L., 2014. Segmentless sentence-final particles in Cantonese: an experimental study. *Studies in Chinese Linguistics*, 35(2), 47–60.

[19] Zhu, Y. 2018. *A Study of Rime Change in the Wuzhi Dialect*. Doctoral dissertation at Tsinghua University.

[20] Zhu, Y., Li, A. Submitted. A phonetic study of disyllabic tones in the Wuzhi dialect.