Prosodic Realization of Focus and Interrogative Intonation in Kunming Mandarin

Siqi Lu[1, 2], *Hui Feng*[1], *Cong Zhang*[3]

[1]Tianjin University
[2]The University of Hong Kong
[3]Newcastle University
Siqi_lu47@connect.hku.hk, fenghui@tju.edu.cn, cong.zhang@newcastle.ac.uk

Abstract

This study examines the prosodic encoding of focus and interrogative intonation in Kunming Mandarin. Acoustic analysis of F0, duration, and intensity from twelve native speakers reveals that Kunming Mandarin is a post-focus compression (PFC) language. Significant F0 lowering and pitch range compression follow initial and medial focus in both declarative and interrogative functions. Focus is marked by pitch rising, pitch range expansion, duration lengthening, and significant pitch reduction in the post-focus domain. Among these cues, duration is the most consistent marker across all positions, while intensity serves a secondary role. Tone sandhi has no statistically significant influence on the prosodic realization of focus. Regarding interrogative intonation, an overall pitch raising combined with a sentence-final rise was observed. Compared to declaratives, focused constituents exhibit higher max F0 and longer durations in interrogations. Furthermore, focus and interrogation are simultaneously encoded: the PFC remains stable under the initial and medial focus, despite the pitch rise of question intonation at the end of the sentence.

Index Terms: focus, interrogative intonation, prosodic encoding, Kunming Mandarin

1. Introduction

Prosody conveys semantic and emotional meanings to fulfill its communicative functions [1]. Interrogative intonation serves as a means to mark questions. Focus refers to the component in a sentence that is emphasized based on the actual semantic or pragmatic needs [2]. It carries the important information in the sentence that needs to be emphasized and highlighted [1]. Focus can be marked by using syntactic means and can also be achieved through prosody.

Research on prosodic encoding has revealed that the prosodic realization of question intonation and focus can both be marked by acoustic parameters such as F0, duration, and intensity. Focused constituents exhibit rising pitch, lengthened duration, and energy boost [3], [4], [5], [6], [7]. In some languages, post-focus compression (PFC) is reported, where the pitch of the post-focus component decreases, the pitch range is compressed, and the energy is reduced [6], [8]. Languages like Mandarin [5], English [4], [9], Swedish [10], German [11], Uyghur [7], Tibetan [7], Nanchang Mandarin [7], Shanghai, Wuxi, Suzhou, and Ningbo Mandarin [12] are PFC languages. However, not every language has PFC. For example, Hong Kong Cantonese [13], Chongqing Mandarin [16], and ethnic minority languages such as De'ang & Wa (Mon-Khmer) [7], Li (Kra-Dai) [12], Tsat (Austronesian) [14], and Bai (Sino-Tibetan) [15] do not exhibit PFC. Specifically, Qin and Xu [16] found that in Chongqing Mandarin, except for a significant increase in mean F0 under initial focus, there are no significant changes in max F0, mean F0, duration, or intensity under other focus conditions, and no post-focus compression is observed. For interrogative intonation, it is marked by an overall elevation in pitch on the sentence level [17]. In contrast, others claim that the pitch-rising effect is more evident in the sentence-final position [18].

Focus often involves post-focus compression (lowering F0), and interrogative intonation shows a pitch rise. This creates a potential conflict when a sentence carries both narrow focus and interrogative intonation. Languages resolve this conflict differently. For instance, Uyghur (a Turkic language) sacrifices PFC to preserve the interrogative rise [19], whereas Mandarin maintains both by compressing post-focus pitch while elevating the boundary tone [20]. These varying approaches indicate that languages find ways to balance the competing demands of marking focus and signaling interrogation. However, the specific strategies employed to achieve this balance still lack sufficient research and understanding.

Kunming Mandarin, a variety of Southwest Mandarin, is a linguistically under-investigated language. While Standard Mandarin is a well-documented PFC language, it remains unclear whether Southwest Mandarin varieties share this prosodic feature or exhibit distinct patterns due to regional evolution. Qin and Xu [16] assumed that Southwest Mandarin as a whole does not have PFC, and Kunming Mandarin will provide empirical linguistic data to test this hypothesis. Furthermore, investigating how Kunming Mandarin resolves the conflict between PFC and interrogative rise contributes to the typological understanding of prosodic interactions in tonal languages. This research seeks to answer the following key questions: (1) What is the prosodic encoding of focus in Kunming Mandarin? Is Kunming Mandarin a PFC language? (2) What are the acoustic cues that encode interrogative intonation in Kunming Mandarin? (3) What is the interaction between focus and interrogative intonation?

2. Methodology

2.1. Participants and materials

Twelve native speakers of Kunming Mandarin (six females and six males; mean age = 24.41 years, *SD* = 3.54) were recruited. All participants were born and raised in Kunming, used the dialect daily, and passed a screening for phonological authenticity and speech clarity.

Kunming Mandarin has four tones: T1 (mid-level), T2 (mid-falling), T3 (high-falling), and T4 (low-dipping) [21].

Materials consisted of twelve sentences in a "*Person* + *Time* + Verb + *Place*" structure, covering three tonal combinations: T1-Tx, T3-Tx, and T4-Tx (x ∈ {1, 2, 3, 4}). These tonal combinations were selected to test potential tone sandhi effects in Kunming Mandarin.

Participants were provided with specific scenarios and were guided to produce sentences under declarative and interrogative conditions and four focus types: broad focus, initial focus (on "Person"), medial focus (on "Time"), and final focus (on "Place").

Table 1: *Sample materials (from* INTO-CASS [22]*)*

Tone	Focus	Targeted Statements	Targeted Questions
T1T1	Initial	**张刚**初三飞舟山。 zhang1 gang1 chu1 san1 fei1 zhou1 shan1. **"Zhang Gang** flew to Zhoushan on the third lunar day."	**张刚**初三飞舟山？ zhang1 gang1 chu1 san1 fei1 zhou1 shan1? **"Zhang Gang** flew to Zhoushan on the third lunar day?"
	Medial	张刚**初三**飞舟山。 zhang1 gang1 chu1 san1 fei1 zhou1 shan1. "Zhang Gang flew to Zhoushan **on the third lunar day**."	张刚**初三**飞舟山？ zhang1 gang1 chu1 san1 fei1 zhou1 shan1? "Zhang Gang flew to Zhoushan **on the third lunar day**?"
	Final	张刚初三飞**舟山**。 zhang1 gang1 chu1 san1 fei1 zhou1 shan1. "Zhang Gang flew to **Zhoushan** on the third lunar day."	张刚初三飞**舟山**？ zhang1 gang1 chu1 san1 fei1 zhou1 shan1? "Zhang Gang flew to **Zhoushan** on the third lunar day?"
	Broad	张刚初三飞舟山。 zhang1 gang1 chu1 san1 fei1 zhou1 shan1. "Zhang Gang flew to Zhoushan on the third lunar day."	张刚初三飞舟山？ zhang1 gang1 chu1 san1 fei1 zhou1 shan1? "Zhang Gang flew to Zhoushan on the third lunar day?"

2.2. Data collection and analysis

Recordings were conducted in a quiet room using a Plantronics Blackwire 3220 headset at a sampling rate of 44.1 kHz (16-bit, mono-channel). Participants were provided time to practice before recording to familiarize themselves with the tasks.

A total of 1,152 tokens (12 sentences × 4 focus types × 2 moods × 12 speakers) were collected and manually segmented in Praat [23]. Acoustic parameters, including F0, intensity, and duration, were extracted using the ProsodyPro script [5]. F0 values were converted to semitones (ref: 60 Hz) using (1) and normalized using Z-scores (2) [24] to minimize individual variation among speakers. Statistical analyses were performed using Linear Mixed-Effects Models (LME) with the *lme4* and *lmerTest* packages in R [25].

$$\mathrm{F0(st)} = 12 \times \left(\log_2(\mathrm{F0}) - \log_2(\mathrm{F_{ref}})\right) \quad (1)$$

$$z = \frac{x - \mu}{\sigma} \quad (2)$$

3. Results and discussions

3.1. Analysis of prosodic realization of focus

3.1.1. F0 Analysis

Figure 1 illustrates the F0 contours of statements. Under the broad focus condition, the overall F0 shows a downward trend from the beginning to the end of the sentence, and this is particularly evident in the sentence-final position.

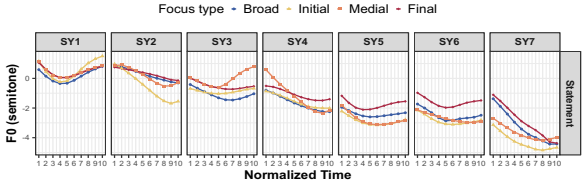

Figure 1: *F0 contours under broad, initial, medial, and final focus in statements.*

In narrow focus conditions, the medial and final focuses tend to achieve acoustic prominence by a higher F0, resulting in higher pitch than the surrounding non-focused words and the corresponding positions in the broad focus condition. The statistical analysis showed that focused constituents in medial ($p < .01$) and final ($p < .05$) positions exhibited a significantly higher F0 compared to the broad focus baseline. However, for the initial focus, the F0 did not differ significantly from the broad focus ($p = .59$).

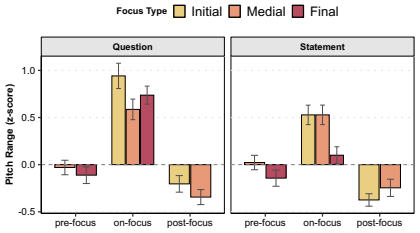

Figure 2: *Pitch range under initial, medial and final focus in statements.*

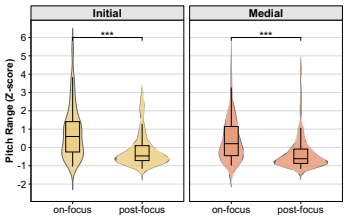

Figure 3: *Comparisons between pitch range of on-focus position and post-focus position in statements*

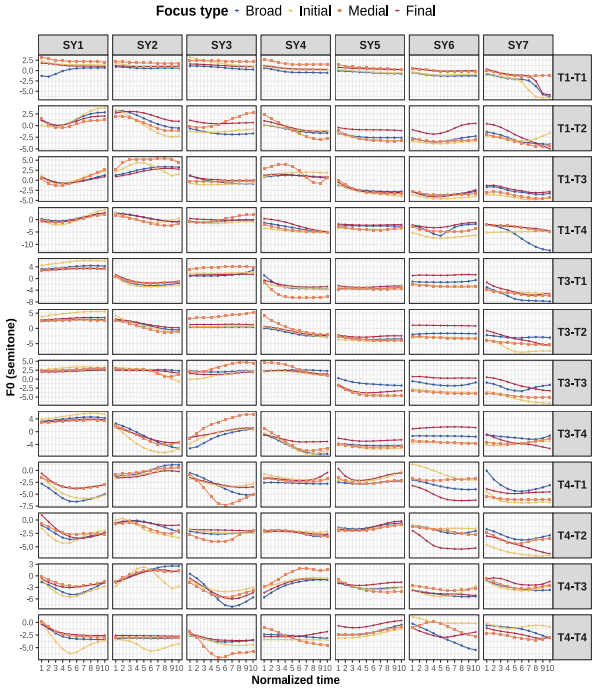

Figure 4: *Tone effect on F0 contours of statement in broad, initial, medial, and final focus conditions.*

However, in all focus conditions, post-focus compression was observed, with a significant pitch drop and pitch range compression in the post-focus positions ($p < .001$) (See Figure 2 and Figure 3). These results confirm that Kunming Mandarin uses PFC as a strategy for focus marking. This consistency

contrasts sharply with the non-PFC Southwest Mandarin branch Chongqing Mandarin [16], where post-focus regions show no PFC. However, the presence of PFC in Kunming aligns typologically with languages like Mandarin [5].

To further examine how focus interacts with lexical tones (see Figure 4), an LME was conducted. No significant interaction was found between Focus and Tone Combination (p = .94). This result indicates that the prosodic encoding of focus in Kunming Mandarin is stable. Tonal combinations do not influence the realization of focus.

3.1.2. Duration Analysis

Figure 5 illustrates that focused constituents have the longest duration across all sentence positions compared to any other focus condition. To statistically verify this, separate LMEs were conducted for Word 1, Word 2, and Word 3 under statement mood.

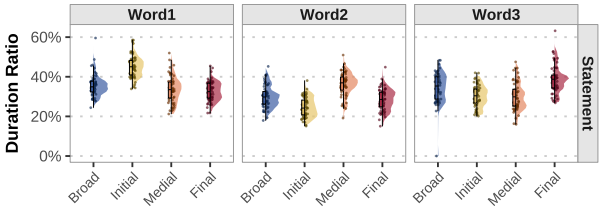

Figure 5: *Duration ratio of broad, initial, medial, final and focus under statement*

Statistical results showed that in all focus conditions, the duration of the focused constituents is the longest. For Word 1, the duration under the initial focus condition was significantly longer than in broad focus (β= -1.02, p < .001), medial focus (β = -0.86, p < .001), and final focus (β = -1.18, p < .001). Similarly, Word 2 achieved longest duration under the medial focus condition, which is significantly longer than in broad, initial, and final focus conditions (ps < .001). For Word 3, the final focus condition also has the longest duration, significantly longer than the broad (p < .01), initial (p < .001), and medial (p < .001) focus conditions.

These results show that, compared to broad focus, the duration ratios in the other three focus types all significantly increased. Additionally, there is a one-to-one mapping between focus position and duration peak: Word 1 is longest in initial focus, Word 2 in medial focus, and Word 3 in final focus. This demonstrates the effect of the focus position on the duration.

3.1.3. Intensity Analysis

The intensity of the three keywords is presented in Figure 6. From the graph, focused words exhibit slightly higher intensity compared to their non-focused counterparts under the final focus condition. However, this trend is not pronounced in initial and medial focus conditions. To examine the effects of focus type and word position on intensity, separate LMEs were conducted.

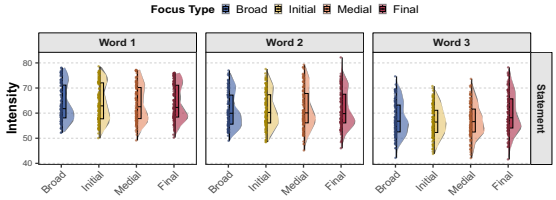

Figure 6: *Intensity of initial, medial, final and broad focus under statement*

For Word 3, the final focus condition showed the highest intensity, significantly higher than the broad (β = -0.40, p < .001), initial (p < .001), and medial (p < .001) focus conditions. This indicates that intensity is a primary cue for final focus. In contrast, Word 1 (initial focus) and Word 2 (medial focus) did not yield statistically significant differences (ps > .05).

3.2. Analysis of prosodic realization of interrogative intonation and focus

3.2.1. F0 Analysis

Figure 7 is a comparison of pitch contours between declarative and interrogative intonation. In interrogative sentences, the pitch contour is characterized by post-focus compression followed by a sentence-final rise: after the F0 is compressed in the post-focus region, it rises again at the sentence-final position to signal question. This suggests that Kunming Mandarin adopts a strategy of simultaneous marking: PFC and the sentence-final interrogative rise co-exist stably within sentences.

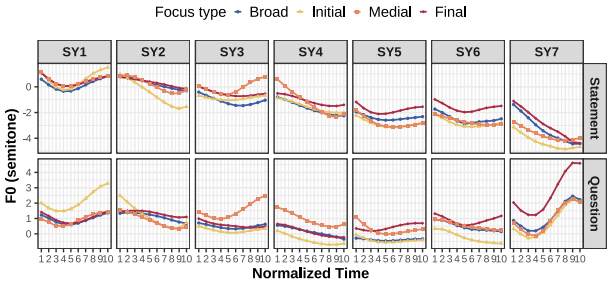

Figure 7: *Comparison between focus realization under interrogative and declarative mood*

Compared to declarative sentences, interrogative sentences show a global upward trend. As shown in Table 2, the max F0 of focused words is higher in questions than in statements. This implies that in the interrogative context, to ensure the clear transmission of focus information, the speaker adopts a more exaggerated F0 strategy.

Table 2: *Max F0 (semitones) of three keywords under broad, initial, medial, and final focus in statements and questions*

	Focus Condition	Word 1	Word 2	Word 3
Question	Broad	58.7	57.5	59.1
	Initial	**61.4**	58.0	58.8
	Medial	58.8	**60.8**	58.8
	Final	59.0	58.0	**61.0**
Statement	Broad	58.4	56.4	55.3
	Initial	**59.3**	56.5	54.8
	Medial	58.5	**58.8**	54.8
	Final	58.3	56.7	**55.9**

To verify the statistical significance of PFC in interrogative sentences, LME was conducted. Similar to statements, focused constituents in medial (p < .01) and final (p < .01) positions showed significant F0 raising compared to the broad focus baseline. However, for the initial focus, the expansion was not statistically significant (p = .09). Despite this, it is observed that

the pitch range in the post-focus domain was significantly compressed compared to the focused constituent ($p < .001$) (See Figure 8). These findings statistically confirm that Kunming Mandarin maintains PFC as a focus marker even when in questions.

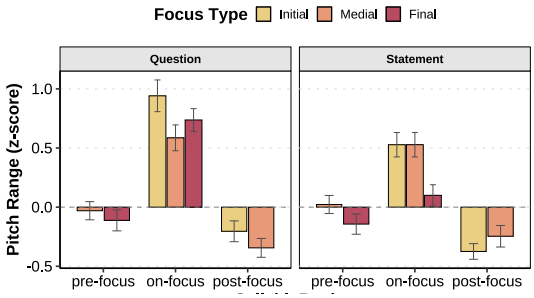

Figure 8: *Pitch range of initial, medial, and final focus under statement and question*

3.2.2. Duration Analysis

Consistent with the findings in declarative sentences, focus in interrogative sentences also has a significant influence on duration, as shown in Figure 9. To statistically verify this, LMEs were fitted for the three target words under the interrogative functions.

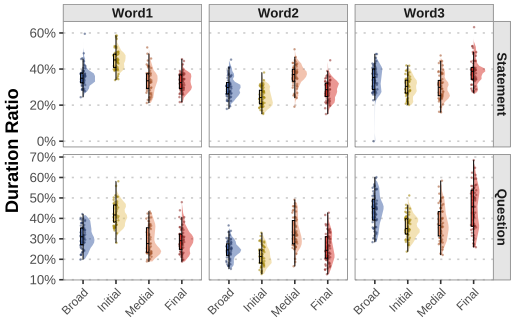

Figure 9: *Duration ratio of initial, medial, final and broad focus in statements and questions*

There is a consistent pattern for duration in all focus conditions. Specifically, for Word 1, the duration under initial focus was significantly longer than in broad, medial, and final focus conditions ($ps < .001$). Similarly, Word 2 exhibited the longest duration under medial focus, significantly exceeding all other focus conditions ($ps < .001$). For Word 3, the final focus condition still elicited a significantly longer duration compared to broad focus ($p < .01$), initial focus ($p < .01$), and medial focus ($p < .05$). Overall, the duration of the focused constituent is longer than the non-focused counterparts.

Comparing the two moods, LME analyses revealed that the duration was longer in interrogative sentences across all positions ($ps < .05$). Whether at the sentence initial, medial, or final position, the focused constituents in questions exhibited significantly longer durations than their counterparts in statements. This suggests that in interrogative sentences, speakers rely more on duration to distinguish focus.

3.2.3. Intensity Analysis

Figure 10 demonstrates the comparison between the intensity of statement and question, and the focus's effect on question intonation.

Partly consistent with statements, focus does not exhibit a significant effect on intensity. For Word 3, the final focus condition showed the highest intensity, significantly exceeding the broad focus (β = -0.43, $p < .001$) and the initial focus condition ($p < .001$). In contrast to statements where medial focus was not significantly louder than the broad focus, Word 2 (medial) in questions exhibited a higher intensity compared to broad focus (β = -0.41, $p < .01$). For Word 1 (initial), however, the intensity did not differ significantly from the broad focus baseline ($p = .14$). This suggests that intensity is not a stable or consistent cue to signal focus in Kunming Mandarin; speakers rely more on pitch and duration to mark focus.

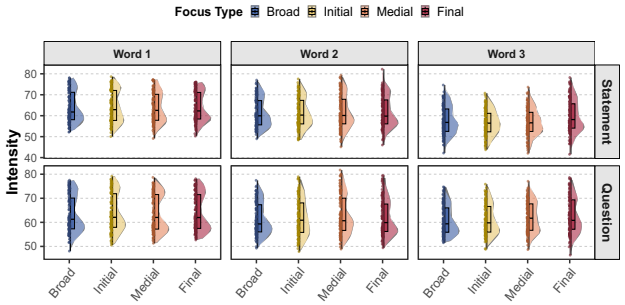

Figure 10: *Intensity of initial, medial, final, and broad focus under statement and questions*

4. Conclusions

This study investigated the prosodic encoding of focus and interrogative intonation in Kunming Mandarin, a Southwest Mandarin variety. Qin & Xu [16] raised a hypothesis that Southwest Mandarin as a whole lacks PFC, with evidence from Chongqing Mandarin. Findings in the current paper show that Kunming Mandarin exhibits PFC hence PFC presence cannot be assumed across all dialects in this family.

Focus in Kunming Mandarin is realized through pitch range expansion, F0 raising, and duration lengthening. Among all acoustic parameters, duration is the most consistent focus marker, showing significant effects across all focus positions (initial, medial, final) and both moods (declarative, interrogative). Initial focus shows weaker F0 effects compared to medial and final positions, a pattern likely reflecting competition between focus-induced pitch raising and sentence-initial F0 declination. Intensity plays only a secondary role, primarily effective in the final position. Importantly, tone sandhi does not influence focus realization, confirming that focus encoding operates independently of the lexical tone system.

Focus and interrogation are simultaneously encoded in Kunming Mandarin. Interrogative intonation is marked by global raising and a sentence-final rise. Compared to declarative sentences, higher maximum F0 values and longer durations characterize focused constituents in interrogative contexts, suggesting that speakers employ more exaggerated prosodic cues to maintain focus salience in interrogative sentences. Additionally, PFC remains stable under the initial and medial focus, despite the pitch rise of question intonation at the end of the sentence. This suggests that, similar to Mandarin, Kunming Mandarin adopts a strategy of preserving both PFC and sentence-final question intonation. The rising intonation in questions primarily occurs in the sentence-final region.

5. Acknowledgements

This work was supported by the Independent Topic Selection Project of Cultural Experts and "Four Batches" of Talents awarded to Aijun Li, and also Key Laboratory of Linguistics, Chinese Academy of Social Sciences (2024SYZH001).

6. References

[1] D. R. Ladd, *Intonational Phonology*. Cambridge, U.K.: Cambridge University Press, 2008.

[2] D. L. Bolinger, "A theory of pitch accent in English," *WORD*, vol. 14, nos. 2–3, pp. 109–149, 1958, doi: 10.1080/00437956.1958.11659660.

[3] W. E. Cooper, S. J. Eady, and P. R. Mueller, "Acoustical aspects of contrastive stress in question–answer contexts," *The Journal of the Acoustical Society of America*, vol. 77, no. 6, pp. 2142–2156, 1985.

[4] S. J. Eady and W. E. Cooper, "Speech intonation and focus location in matched statements and questions," *The Journal of the Acoustical Society of America*, vol. 80, no. 2, pp. 402–415, 1986.

[5] Y. Xu, "Effects of tone and focus on the formation and alignment of f0 contours," *Journal of Phonetics*, vol. 27, no. 1, pp. 55–105, 1999.

[6] Y. Xu, S. Chen, and B. Wang, "Prosodic focus with and without post-focus compression: A typological divide within the same language family?" *The Linguistic Review*, vol. 29, no. 1, 2012, doi: 10.1515/tlr-2012-0006.

[7] B. Wang, L. Wang, and T. Qadir, "Prosodic realization of focus in six languages/dialects in China," in *Proc. International Congress of Phonetic Sciences (ICPhS)*, 2011, pp. 144–147.

[8] Y. Xu, "Post-focus compression: Cross-linguistic distribution and historical origin," in *Proc. International Congress of Phonetic Sciences (ICPhS)*, 2011, pp. 152–155.

[9] Y. Xu and C. Xu, "Phonetic realization of focus in English declarative intonation," *Journal of Phonetics*, vol. 33, no. 2, pp. 159–197, 2005.

[10] M. Heldner and E. Strangert, "Temporal effects of focus in Swedish," *Journal of Phonetics*, vol. 29, no. 3, pp. 329–361, 2001.

[11] C. Féry and F. Kügler, "Pitch accent scaling on given, new and focused constituents in German," *Journal of Phonetics*, vol. 36, no. 4, pp. 680–703, 2008.

[12] Q. Wu, B. Wang, X. Zhang, and B. Wang, "Does silent pause facilitate focus perception in a non-PFC language? An experimental study of Li language," in *Proc. 18th Conference of the Spoken Language Communication Association (COCOSDA)*, 2015.

[13] W. Wu and Y. Xu, "Prosodic focus in Hong Kong Cantonese without post-focus compression," in *Proc. International Conference on Speech Prosody*, 2010, pp. 1–4.

[14] B. Wang, C. Li, Q. Wu, X. Zhang, B. Wang, and Y. Xu, "Production and perception of focus in PFC and non-PFC languages: Comparing Beijing Mandarin and Hainan Tsat," in *Proc. Interspeech*, 2012, pp. 663–666.

[15] L. Liu, B. Wang, and X. Li, "The Prosodic Realization of Focus and Boundary in the Bai Language," *Minority Languages of China*, vol. 6, pp. 71–79, 2020.

[16] Z. Qin and Y. Xu, "Lack of Prosodic Focus in Chongqing Dialect and Possible Historical Sources," in *Proceedings of the 10th International Conference on Speech Prosody (Speech Prosody 2020)*, Tokyo, Japan, May 25-28, 2020, pp. 630-634, doi:10.21437/SpeechProsody. 2020-129.

[17] A. Ho, "Intonation variation in a Mandarin sentence for three expressions: Interrogative, exclamatory and declarative," *Phonetica*, vol. 34, no. 6, pp. 446–457, 1977, doi: 10.1159/000259916.

[18] J. Ma, V. Ciocca, and T. L. Whitehill, "Quantitative analysis of intonation patterns in statements and questions in Cantonese," in *Proc. 3rd International Conference on Speech Prosody*, 2006, pp. 277–280.

[19] K. Qadir and B. Wang, "Weiwueryu zhong Yiwen he Jiaodian dui Yudiao de Gongtong Tiaojie Zuoyong [Interaction between interrogative and focus in Uyghur]," *Journal of Computer Applications*, vol. 33, no. 3, pp. 784–788, 2013.

[20] Y. He and S. Jing, "Beijinghua Yudiao de Shiyan Tansuo [Experimental exploration of intonation in Beijing Mandarin]," *Language Teaching and Linguistic Studies*, no. 2, pp. 71–96, 1992.

[21] X. Chen, "Study on the origin and change of the phonetic system of Yunnan dialect," Ph.D. dissertation, Nankai Univ., Tianjin, China, 2013.

[22] A. Li and Z. Xiong, "INTO-CASS: A Corpus for the Study of Intonation and Prosody in Chinese Dialects and Ethnic Language," In *24th Conf. Oriental COCOSDA Int. Comm. for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)*, 2021.

[23] P. Boersma and D. Weenink, "Praat: Doing phonetics by computer," Version 6.1.53. [Online]. Available: http://www.praat.org/, 2021.

[24] P. Rose, "Considerations in the normalisation of the fundamental frequency of linguistic tone," *Speech Communication*, vol. 6, no. 4, pp. 343–352, 1987.

[25] R Core Team, "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria. [Online]. Available: https://www.R-project.org/, 2024.