Interspeech 2024
1-5 September 2024, Kos, Greece

Key Acoustic Cues for the Realization of Metrical Prominence in Tone Languages: A Cross-Dialect Study

Yiying Hu[1,2], *Hui Feng*[1*]

[1]School of Foreign Languages, Tianjin University, China
[2]Research Centre for Language, Cognition, and Neuroscience, Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong SAR, China
huyiying@tju.edu.cn, fenghui@tju.edu.cn

Abstract

Few studies have quantitatively discussed the acoustic realization of metrical prominence across tone languages. This study addresses this issue in word-level prosodic units, aiming to identify effective acoustic cues for cross-language analysis based on the realization of metrical prominence. Findings include: (1) Within-dialect analyses reveal robust acoustic cues for realizing metrically strong units: duration, pitch height, and pitch slope for both left- and right-dominant prominence, and intensity, pitch range and cumulative pitch dynamics, an often-overlooked cue, for right-dominant one; (2) Cross-dialect analyses by Boruta algorithm identify that all these six acoustic cues are significant for the classification of Chinese dialects according to their realization of metrical prominence. Linear Discriminant Analysis emphasizes duration, intensity, and cumulative pitch dynamics as key discriminators across dialects.

Index Terms: acoustic realization, metrical prominence, Chinese dialects, machine learning

1. Introduction

1.1. Metrical prominence across Chinese dialects

Metrical structure refers to the prominence of prosodic elements [1]. Following the relative prominence initially proposed by Liberman and Prince [1], Hayes [2] introduced the concepts “trochaic” and “iambic” into linguistics and focused on the differences between the trochaic types and the iambic types among world languages.

Based on prosodic elements at the word level, the world languages can be categorized into prototype systems including tone languages and stress-accent languages [3]. In tone languages, metrical prominence is related to the word-level tonal patterns in the prosodic unit [4, 5, 6]. Metrical prominence of Chinese dialects can be generally classified into the left-dominant and the right-dominant metrical structure [4, 5, 7, 8, 9].

The differences between the typically left- and right-dominant metrical structures can partially be observed in the surface tonal realization. In the left-dominant structure, a strong-weak pattern is presented by the initial syllable retaining the citation tone and the non-initial syllable undergoing the sandhi tone [1, 5, 9], such as Chengdu Mandarin [10], Changzhou Wu [9], and Dongkou Xiang [9]. In the right-dominant metrical structure, conversely, the weak-strong pattern is presented by the final syllable retaining the citation tone and the non-final syllable undergoing the sandhi tone [4, 5, 7, 9], such as Xiamen Min [11], Fuzhou Min [12], Kunming Mandarin [13], Tianjin Mandarin [14], and Changsha Mandarin [5]. Previous research has noticed that the tone sandhi usually occurs in the metrically weak unit at the word level [5], also referred to as the metrical tone sandhi [15].

1.2. Acoustic realization of metrical prominence

Acoustic correlates of metrical prominence have always been a focus of research. For languages with stress, stress is related to longer duration, stronger intensity, higher pitch, more peripheral vowels, and spectral tilt [16, 17, 18, 19, 20]. For tone languages, studies concerning the acoustic realization of metrical prominence mainly focus on cues including duration, intensity, and pitch realization [2, 5, 6, 9, 19].

The role of duration in the realization of metrical prominence is inconsistent in previous research. Several studies on tone languages indicate duration as a reliable phonetic parameter associated with stress [17], such as Thai [21] and Kurtöp [22]. In Chinese dialects, duration has been regarded as a correlate of prosodic strength [23, 24]. Specifically, a short-long pattern can be found in the right-dominant metrical structure, regarded as a way to realize the trochaic structure [25], such as the Fuzhou Min and Xiamen Min [4, 5]. And a long-short pattern is found in the left-dominant metrical structure, such as Changsha Mandarin [5]. In contrast, duration is insensitive to metrical prominence in certain Chinese dialects, such as Chengdu Mandarin [5] and Shaoxing Wu [6], thereby not serving as a universal cue of metrical prominence.

Intensity can also indicate the prosodic strength at the word level, realizing metrical prominence. In languages with stress, the trochaic structure shows prominence mainly with intensity contrasts [25]. And a weak-strong pattern can be found in the right-dominant metrical structure, while a strong-weak pattern can be found in the left-dominant one, such as Rebkong Tibetan [26].

Pitch realization is the key parameter of metrical prominence among Chinese dialects. Specifically, the pitch correlates of prosodic strength in prior studies include pitch height, pitch range, and pitch manifestation of tonal patterns [4, 5, 7]. The greater pitch range is associated with the realization of stress [27]. For pitch manifestation of tonal patterns among most of Chinese dialects, the citation tone is usually retained by the metrically strong unit in the disyllabic word, while the sandhi tone is presented in the metrically weak unit [5, 9].

Three critical and progressively related issues remain unresolved, regarding the realization of metrical prominence in tone languages. First, prior studies have not reached a

 10.21437/Interspeech.2024-666

consensus on which acoustic cues are the most robust to realize metrical prominence. Second, the pitch realization, a key parameter for metrical prominence, calls for detailed exploration. While studies concerning the pitch realization of metrical prominence usually focus on sandhi patterns [4, 5, 9, 15], there is a mismatch between sandhi prominence and metrical prominence in tone languages [6, 7]. Other studies tend to emphasize roles of pitch range and pitch height in realizing prominence [26], neglecting the importance of the role of pitch dynamics. Given its significance in classifying tone languages, pitch realization deserves a more detailed investigation according to the realization of metrical prominence. Third, most importantly, the comparison of tone languages based on metrical prominence require acoustic cues effective for both within- and cross-language analyses. Such cues facilitate a cross-language measurement of acoustic correlates of metrical prominence, enhancing classification and identification of tone languages.

This study aims to identify the effective acoustic cues for cross-language analysis based on the realization of metrical prominence, in the case of three Chinese dialects with either left- or right-dominant metrical prominence. First, the within-dialect analysis, with a detailed exploration in pitch realization, explores how metrical prominence is realized by acoustic cues in one metrically left-dominant dialect (Chongqing Mandarin) and two metrically right-dominant dialects (Kunming Mandarin and Xiamen Min). Second, the cross-dialect analysis, with machine-learning approaches, conducts a cross-dialect comparison based on the acoustic realization of metrical prominence, to identify the effective cues cross dialects.

2. Methods

2.1. Materials

The selected Chinese dialects include one metrically left-dominant dialect (Chongqing Mandarin) and two metrically right-dominant dialects (Kunming Mandarin and Xiamen Min). Both Chongqing Mandarin and Kunming Mandarin are variants of the Southwestern Mandarin group. Xiamen Min is a variant of the Min dialect group. The tone inventories of three dialects are listed in Table 1. The rationale for selecting these dialects is that Kunming Mandarin and Chongqing Mandarin are phonologically closer, while Kunming Mandarin and Xiamen Min share similarities in realizing prominence. Such a selection can verify the effectiveness of cross-dialect analyses and further enrich our understanding of the classification of tone languages.

The materials of production experiments are disyllabic prosodic words in the selected dialects. For each tonal combination in each dialect, three disyllabic words in closely connected structures with high frequency were chosen. The checked tones are excluded as the duration of the checked tone is shorter compared to other tones. To control the effects of sandhi domains across different dialects, the carrier sentence is not used. All materials were read two times by each speaker, producing a total of 2,736 tokens.

Table 1: *Tone inventories of the dialects* [28, 29, 30].

Metrical Structure	Dialect	T1	T2	T3	T4	T5	T6	T7
left-dominant	Chongqing Mandarin	55	21	42	214	—	—	—
right-dominant	Kunming Mandarin	441	42	53	31	—	—	—
right-dominant	Xiamen Min	44	24	53	21	22	32	44

2.2. Participants

A total of 24 middle-aged participants were recruited, including eight native speakers of Chongqing Mandarin, Kunming Mandarin, and Xiamen Min, respectively. The detailed information is shown in Table 2. Strict standards were guaranteed during recruitment. All participants were born and raised in their native language-speaking regions, frequently engaged in dialectal communication in their daily interactions, with both their parents and grandparents being native speakers of the dialect. And participants never left their native language-speaking regions for more than three months.

Table 2: *Information of participants.*

Dialect	Participants	Mean Age
Chongqing Mandarin	4 females, 4 males	51.5 (SD=5.9)
Kunming Mandarin	4 females, 4 males	53.6 (SD=7.5)
Xiamen Min	4 females, 4 males	54.4 (SD=4.0)

2.3. Procedure

This study was approved by the ethical committee. The experiments were conducted in Chongqing, Kunming, and Xiamen city, respectively. Each participant read materials presented on a screen at a natural and fluent speed in a quiet room. The recording equipment was a Plantronics Blackwire 3220 Series headset microphone connected to a computer. The audio samples were recorded with Praat [31] at a sampling rate of 44.1 kHz, 16-bit resolution, and were stored as mono tracks.

2.4. Measurements for robust acoustic cues

2.4.1. For within-dialect analysis

The within-dialect analysis is to find out the dominant acoustic cues for realizing metrical prominence in metrically left- and right-dominant dialects, respectively.

For the tone-bearing unit of each syllable in disyllabic words, the duration, intensity, and f_0 values of ten points were extracted by Praat [31]. With the z-score method [32], the raw data was normalized for each participant.

For analyses of the duration and intensity, to investigate whether there is a significant difference in the duration distribution and intensity distribution of the two adjacent syllables within the word, a Mann-Whitney test was conducted on the normalized data, followed by Bonferroni correction.

For analysis of the pitch manifestation, the normalized data were smoothed using a linear time registration, using B-spline interpolation in R [33] for pitch contour refinement. Four acoustic parameters concerning the pitch realization of each syllable will be calculated, including pitch height (mean f_0), pitch range (maximal f_0 minus minimal f_0), pitch slope (f_0 range divided by duration), and cumulative pitch dynamics (sum of absolute f_0 change rates across sequential time points). The first three parameters have been focused on in prior studies, whereas the last parameter is often overlooked. Since this parameter measures the overall dynamics of the pitch change, this study also takes it into account. To reduce the tone-irrelevant variation, 10% of the extreme f_0 values at both ends of the pitch contour were omitted from calculations. The Mann-Whitney tests were conducted to explore the differences of the two adjacent syllables in pitch height, pitch range, pitch slope, and cumulative pitch dynamics.

2.4.2. *For cross-dialect analysis*

The cross-dialect analysis is to conduct a cross-dialect comparison based on the acoustic cues of metrical prominence. To effectively measure the acoustic realization of the prominence, the differences between two adjacent syllables in the six acoustic parameters—duration, intensity, pitch height, pitch range, pitch slope, and cumulative pitch dynamics—are measured by their differential normalized values (σ1 minus σ2), generating six features for cross-dialect analysis.

First, feature selection for cross-dialect comparison was conducted with the Boruta algorithm in the R package [34], a method based on the Random Forest classification. The Boruta algorithm identifies statistically significant features by generating "shadow features", which are randomized permutations of the original features in the dataset, to establish a benchmark for evaluating the significance of actual features.

Second, Linear Discriminant Analysis (LDA), a statistical method for dimensionality reduction and classification, was trained and run with the R package MASS [35] to reduce features to two dimensions for cross-dialect visualization. Features evaluated as important by the Boruta algorithm were used for cross-dialect comparison, based on their acoustic realization of metrical prominence. Even in the potential presence of correlations among features (six acoustic parameters), LDA can still be used for dimensionality reduction. It aims to maximize the ratio of between-class variance to within-class variance, thereby seeking the most effective separation of dialects. And five-fold cross-validation is conducted to ensure the robustness of model evaluation.

3. Results and Discussion

3.1. Within-dialect analysis

Figure 1 shows the distribution of six acoustic parameters—normalized duration, mean intensity, pitch height (mean f_0), pitch range, pitch slope, and cumulative pitch dynamics—for disyllabic words in Chongqing Mandarin, Kunming Mandarin, and Xiamen Min. Following sections analyze robust acoustic cues in left- and right-dominant prominence, respectively.

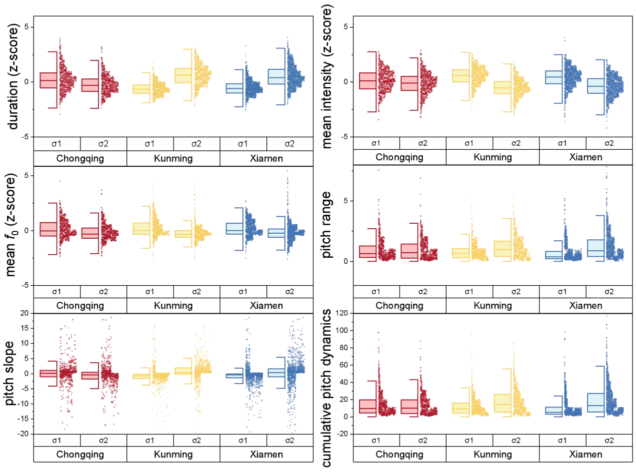

Figure 1: *Distributions of the six acoustic parameters across the three dialects.*

Mann-Whitney tests were conducted to test the differences between the two adjacent syllables in the six acoustic cues across the dialects, with results listed in Table 3.

Table 3: *Results of Mann-Whitney tests.*

Acoustic cues	**Left-dominant dialect**	**Right-dominant dialects**	
	Chongqing Mandarin	Kunming Mandarin	Xiamen Min
duration	U=178769.000 p<.001	U=81258.500 p<.001	U=275589.000 p<.001
intensity	U=209043.500 p<.001	U=107123.500 p<.001	U=383781.500 p<.001
mean f0	U=187116.500 p<.001	U=185485.500 p<.001	U=508883.000 p<.001
pitch range	U=240804.500 p=0.925	U=234077.000 p<.001	U=459293.500 p<.001
pitch slope	U=196293.000 p<.001	U=159229.000 p<.001	U=466677.000 p<.001
cumulative pitch dynamics	U=233090.500 p=0.260	U=230821.500 p<.001	U=443901.000 p<.001

3.1.1. *Realization of left-dominant metrical prominence*

In the dialect of the left-dominant metrical prominence (Chongqing Mandarin), there are significant differences between σ1 and σ2 for the duration, mean intensity, mean f_0, and pitch slope, and there is no significant difference between σ1 and σ2 for pitch range and cumulative pitch dynamics.

The metrically strong unit in the left-dominant metrical structure is related to longer duration, stronger intensity, higher mean f_0, and greater pitch slope. Results of durational realization reach a consensus with findings in [5]. The results of the intensity realization may not be due to metrical realization but to the first syllable at the utterance, where intensity naturally increases. Results of pitch realization complement prior studies by refining several parameters of f_0 realization, indicating that the metrically strong unit in the left-dominant structure is more related to the direction of pitch change, but not to its overall dynamic.

3.1.2. *Realization of right-dominant metrical prominence*

Dialects of the right-dominant metrical prominence include Kunming Mandarin and Xiamen Min in this study. The results from these two dialects are in agreement with each other. There are significant differences between σ1 and σ2 for the duration, mean intensity, mean f_0, pitch range, pitch slope, and cumulative pitch dynamics.

The metrically strong unit in the right-dominant metrical structure is related to longer duration, lower mean intensity, lower mean f_0, greater pitch range, greater pitch slope, and greater cumulative pitch dynamics. Results of durational realization reach a consensus with previous studies [5, 17]. Results of the intensity reveal a new observation: despite the higher intensity of the initial syllable caused by the utterance start, the initial syllable in the right-dominant structure shows greater intensity than that in the left-dominant one under the same conditions, indicating that intensity is not always a robust acoustic cue for realizing prominence. Results of pitch realization illustrate that the metrically strong unit in the right-dominant structure is related to the overall dynamic, direction, and the overall range of pitch change. The cumulative pitch dynamics is rarely explored in prior research, but it is a robust cue for realizing the right-dominant structure in tone languages.

3.2. Cross-dialect analysis

3.2.1. Selection of acoustic cues

The cross-dialect analysis aims to compare and visualize dialects based on the acoustic realization of metrical prominence. To identify the most relevant acoustic cues, we employed the Boruta algorithm, a feature selection method based on Random Forest classification. Plotted in Figure 2, the Boruta algorithm identified all the six acoustic parameters—duration, intensity, pitch height, pitch range, pitch slope, and cumulative pitch dynamics—as significantly important, with all attributes surpassing the shadowMax threshold. And it also suggests a strong effect of durational realization.

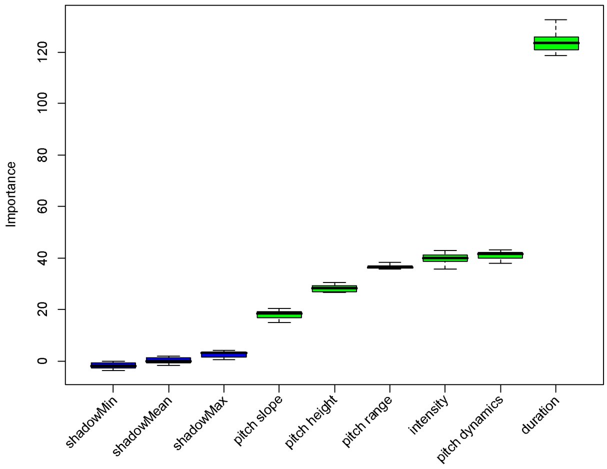

Figure 2: *Importance of acoustic cues estimated by the Boruta algorithm.*

3.2.2. Dimensionality reduction and visualization

Linear Discriminant Analysis (LDA) was applied to project the six statistically significant acoustic features to a lower-dimensional space with optimal class separability, making it possible for cross-dialect visualization based on the realization of metrical prominence.

LDA calculates the means for each acoustic cue in dialect group, indicating the centroids of the dialects in the transformed space, suggesting the cumulative pitch dynamics as the most significant acoustic cue across dialects. Specifically, the metrically left-dominant dialect (with a group mean of 1.304617 in Chongqing Mandarin) contrasts sharply with the metrically right-dominant dialects (with group means of -6.297154 in Kunming Mandarin and -9.306130 in Xiamen Min) in the cumulative pitch dynamics. The pitch realization has been considered as a robust indicator of the metrically binary contrast previously [5], and the current study goes further into the cumulative pitch dynamics.

In terms of the coefficients of the linear discriminants, the first linear discriminant (LD1) accounts for 98.12% of the variance, largely influenced by the feature of duration. The second linear discriminant (LD2), although it explains a minor portion of the variance (1.88%), is primarily influenced by intensity, further distinguishing dialects when combined with LD1.

The LDA results is shown in Figure 3. The metrically left-dominant dialect (Chongqing) and metrically right-dominant dialects (Kunming and Xiamen) are roughly separated along the LD1 axis. Notably, the overlapping data points of Kunming and Xiamen dialects underscores their shared characteristics in right-metrical prominence. This observation validates this approach to exploring metrical prominence. In detail, despite Chongqing and Kunming Mandarin both being classified within the Southwestern Mandarin group, which is different from the Min dialect group of Xiamen Min, the LDA model groups Kunming Mandarin with Xiamen Min. This outcome suggests that this LDA model, which primarily focuses on metrical prominence, provides an alternative perspective to traditional phonological origin-based classifications for tone languages.

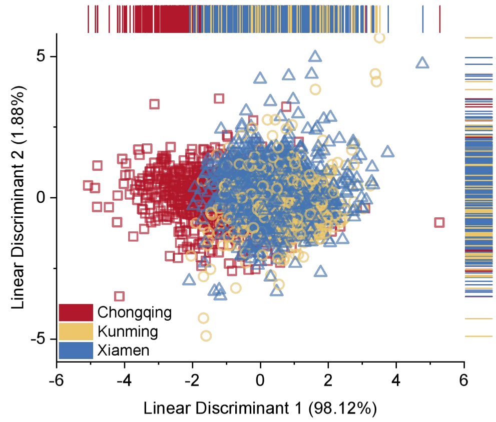

Figure 3: *Results of LDA analysis.*

4. Conclusions

This study identifies key acoustic cues for cross-language analysis based on the realization of metrical prominence in tone languages, with the quantitative investigation of one metrically left-dominant dialect (Chongqing Mandarin) and two metrically right-dominant dialects (Kunming Mandarin and Xiamen Min).

Within-dialect analyses reveal the differences in the realization of metrically strong units between left- and right-dominant structures. Left-dominant metrical prominence is realized by longer duration, higher mean f_0, and greater pitch slope. Right-dominant one is realized by longer duration, lower intensity, lower mean f_0, greater pitch range, greater pitch slope, and greater cumulative pitch dynamics.

The cross-dialect analysis suggests that all the six acoustic cues—duration, intensity, pitch height, pitch range, pitch slope, and cumulative pitch dynamics—are statistically significant for the classification of dialects based on the realization of metrical prominence, estimated by the Boruta algorithm. And the LDA model emphasizes the duration, intensity, and cumulative pitch dynamics as key discriminators across dialects. These findings will facilitate improvements of the classification and identification of tone languages.

The study highlights an often-overlooked acoustic cue, cumulative pitch dynamics, for the realization of metrical prominence. Furthermore, the attempt to explore metrical prominence using machine-learning approaches provides a new perspective of the comparison of tone languages based on acoustic correlates of metrical prominence. There are possible limitations related to uncontrolled variation in vowel quality, which will be further explored with more materials. The future work is to enrich database with more tone languages, enhancing our understanding of metrical prominence with a comprehensive perspective.

5. Acknowledgements

This work was supported by Key Laboratory of Linguistics, Chinese Academy of Social Sciences (Project #2024SYZH001). Gratitude is extended to Yuqin Zhou and Siqi Lu for their assistance in data collection in Chongqing and Kunming, China.

6. References

[1] A. M. Liberman and P. Alan, "On Stress and Linguistic Rhythm," *Linguistic Inquiry*, vol. 8, no. 2, pp. 249–336, 1977.

[2] B. Hayes, "Iambic and trochaic rhythm in stress rules," *Annual Meeting of the Berkeley Linguistics Society*, pp. 429–446, 1985.

[3] L. M. Hyman, "Word-prosodic Typology," *UC Berkeley Phonology Lab Annual Reports*, vol. 1, 2005.

[4] C. Guo, "Types of Prosodic Structure and Tone Sandhi of Disyllabic Words in Chinese Dialects," Doctoral dissertation, Shanghai Normal University, Shanghai, 2020.

[5] C. Guo and F. Chen, "Phonetic Realizations of Metrical Structure in Tone Languages: Evidence From Chinese Dialects," *Front. Psychol.*, vol. 13, p. 945–973, 2022.

[6] X. Wen, Y. Chen, and L. L.-S. Cheng, "Disyllabic Tone Sandhi Patterns in Shaoxing Wu Chinese," in *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS 2023)*, 2023, pp. 1989–1993.

[7] T. Huang and Y. Do, "Directionality of Disyllabic Tone Sandhi across Chinese Dialects is Conditioned by Phonetically-Grounded Structural Simplicity," in *Proceedings of the Annual Meetings on Phonology*, 2021.

[8] A. O. Yue-Hashimoto, "Tone Sandhi across Chinese Dialects," *Wang Li memorial volumes*, vol. English volume, pp. 445–474, 1987.

[9] J. Zhang, "A directional asymmetry in Chinese tone sandhi systems," *J East Asian Linguist*, vol. 16, no. 4, pp. 259–302, 2007.

[10] H. S. Lin, "Chengdu Reduplication: An Optimality Theoretic Analysis," *Language and Linguistics*, vol. 16, no. 6, pp. 843–872, 2015.

[11] M. Y. Chen, "The syntax of Xiamen tone sandhi," *Phonol. yearb.*, vol. 4, no. 1, pp. 109–149, 1987.

[12] J. Zhang, "Tones, tonal phonology, and tone sandhi," in *The handbook of Chinese linguistics*, First Edition., John Wiley & Sons, 2014, pp. 443–464.

[13] H. S. Lin, "Indirect tone-prominence interaction in Kunming tone sandhi," *Concentric*, vol. 45, no. 1, pp. 44–81, 2019.

[14] J. Zhang and J. Liu, "The productivity of variable disyllabic tone sandhi in Tianjin Chinese," *J East Asian Linguist*, vol. 25, no. 1, pp. 1–35, Feb. 2016.

[15] X. Zeng and S. Niu, "The disyllabic tone sandhi and its typological explain of Liujia dialect in Sanjiang county in Guangxi," *Fangyan*, vol. 4, pp. 290–308, 2006.

[16] A. M. C. Sluijter and V. J. Van Heuven, "Spectral balance as an acoustic correlate of linguistic stress," *The Journal of the Acoustical Society of America*, vol. 100, no. 4, pp. 2471–2485, 1996.

[17] M. Gordon and T. Roettger, "Acoustic correlates of word stress: A cross-linguistic survey," *Linguistics Vanguard*, vol. 3, no. 1, p. 1–11, 2017.

[18] C. Gussenhoven, *The Phonology of Tone and Intonation*, 1st ed. Cambridge University Press, 2004.

[19] L. Liang and F. Shi, "The energy ratio of bi-syllabic words in Standard Chinese," *Nankai linguistics*, vol. 16, no. 2, pp. 35–41, 2010.

[20] J. Guo and S. Cui, "A contrastive study on phonetic prominence between Mandarin Chinese and English trisyllabic words," *Chinese Journal of Phonetics*, vol. 18, no. 2, pp. 100–117, 2022.

[21] S. Potisuk, J. Gandour, and M. P. Harper, "Acoustic Correlates of Stress in Thai," *Phonetica*, vol. 53, no. 4, pp. 200–220, 1996.

[22] G. Hyslop, "Between Stress and Tone: Acoustic Evidence of Word Prominence in Kurtöp," *Language Documentation*, vol. 15, pp. 551–575, 2021.

[23] Y. Chen and Y. Xu, "Production of Weak Elements in Speech – Evidence from F_0 Patterns of Neutral Tone in Standard Chinese," *Phonetica*, vol. 63, no. 1, pp. 47–75, Mar. 2006.

[24] Y. Xu, "Timing and coordination in tone and intonation—An articulatory-functional perspective," *Lingua*, vol. 119, no. 6, pp. 906–927, 2009.

[25] B. Hayes, *Metrical stress theory: Principles and case studies*. University of Chicago Press, 1995.

[26] N. J. Caplow, "The role of stress in Tibetan tonogenesis: a study in historical comparative acoustics," Doctoral Dissertation, University of California, Santa Barbara, 2009.

[27] Y. Chao, *A grammar of spoken Chinese.* University of California Press, 1968.

[28] L. Liang and X. Meng, "A Sociophonetic Study on Tones of Chongqing Mandarin in Gender and Age Difference," *Proceedings of the International Congress of Phonetic Sciences (ICPhS 2011)*, pp. 1230–1233, 2011.

[29] X. Chen, "Study on the origin and change of the phonetic system of Yunnan dialect," Doctoral Dissertation, Nankai University, Tianjin, 2013.

[30] Q. Lin, "The Diachrony of Tone Sandhi: Evidence from Southern Min Chinese," Doctoral dissertation, Hong Kong University of Science and Technology, Hong Kong, 2017.

[31] P. Boersma, "Praat, A System for Doing Phonetics by Computer." *Glot Internationl*, pp. 341–345, 2001.

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

[33] R Core Team, "R: A language and environment for statistical computing." 2021.

[34] M. B. Kursa and W. R. Rudnicki, "Feature Selection with the Boruta Package," *J. Stat. Soft.*, vol. 36, no. 11, pp. 1–13, 2010.

[35] W. N. Venables, and B. D. Ripley, *Modern Applied Statisticswith S.* Springer Science & Business Media, New York, 2002.