Duration in Zhangzhou Southern Min: Variation, Correlation, and Constraint

Yishan Huang

University of Sydney; Australian National University

yishan.huang@sydney.edu.au

Abstract

This study explores the status of duration in Zhangzhou Southern Min with rich tonal contrasts across three different linguistic contexts. Several innovative findings have gone beyond conventional assumptions and advanced our knowledge of how suprasegmentals (F0 and duration), segments (coda types), and linguistic contexts (citation, phrase-initial, phrase-final) interact to shape dynamic realisations of duration in the tonal phonology and morphosyntax of Zhangzhou speech. It is hoped to contribute valuable empirical data to the typology of tonal studies in Asia and shed important light on how humans encode and decode duration in their cognitive grammar.

Index Terms: duration, F0, tone, contexts, Zhangzhou, Southern Min

1. Introduction

As a linguistic phenomenon, Duration has been conventionally viewed as an important property of speech segments to convey linguistic content and/or paralinguistic information, but languages use duration differently. Some employ duration for segment intelligibility. For example, in English, the five tense vowels /i, e, ɑ, o, u/ are found longer than the six tense vowels /ɪ, ɛ, æ, ɔ, ʊ, ʌ/ by 38 ms [1]; segment types can be ranked hierarchically in terms of duration: vowels (135ms) > fricative (88ms) > nasals (87ms) > glides (70ms) > liquids (69ms) > stops (30ms), as cited in [1]. Contrastively, some other languages, though the proportion is barely about 3% of the world's languages, use duration for phonemicity, whereby different meanings of lexical items or grammatical structures can be distinguished [2]. For example, in Niger-Congo languages, a phonemic contrast exists between short and long vowels, along with many geminates indicating longer consonants in their segmental inventory [1].

Beyond the scope of segments, a great deal of attention has also been given to exploring the duration in the domain of the suprasegmental system, especially in those Asian languages with rich tonal contrasts. For example, the intrinsic correlation between duration and pitch has been studied cross-linguistically [3], [4], [5], and [6]. In general, vowels with low-pitched tones are considered to have a longer duration than with high-pitched tones, and vowels on rising tones appear longer than on falling or level tones [5] and [6]. In other words, pitch/F0 height and duration are correlated negatively: The higher the pitch/F0, the shorter the duration; the lower the pitch/F0, the longer the duration. However, counterexamples are also found to show a positive association between pitch/F0 and duration, such as in Taiwanese [3] and [4], Cantonese [6], and Yucatec Maya [6].

This study, built upon field linguistics, phonetics, phonology, and statistics, explores the duration status in the rich tonal contrasts of the southern min dialect of Zhangzhou in three different contexts: citation tone, phrase initial tone, and phrase-final tone. It is shown that tones can be classified into different groups in terms of duration, and the realisations can be alternated in accordance with the change of linguistic contexts, and so are the intrinsic relations between F0 and duration. It is guided by the following specific research questions: (a) What function does duration serve in the tonal system of Zhangzhou? (b) How are tones realised in duration across different linguistic contexts? and how many lengths are statistically different across contexts? (c) How are tonal duration and F0 related to each other? and how does their correlation change in accordance with the changing linguistic contexts? (d) Are tonal duration realisations affected by surrounding environments? If so, to what extent are they affected, and what conditions the variations? (e) how are durational realisations in the sandhi position related to their corresponding forms in the citation?

It is hoped to reveal how the local residents employ the parameter of duration as part of the suprasegmental system for communication purposes. It aims to advance our understanding of how suprasegmentals (F0 and duration), segments (coda types), and linguistic contexts (citation and disyllabic phrase) interact to shape dynamic realisations of duration in the tonal phonology and morphosyntax of Zhangzhou Southern Min. It is also hoped to contribute empirical data to the typology of tonal studies in Asia and shed important light on how humans encode duration in their cognitive grammar.

2. Research material and design

Zhangzhou is a prefecture-level city in the south of Fujian province of Southern China with a registered population of about 5.10 million. The colloquial language spoken by native Zhangzhou people is predominantly Southern Min, which is mutually intelligible with other Southern Min varieties (e.g., Taiwanese, Xiamen, and Quanzhou) but is entirely unintelligible with other Sinitic dialects (e.g., Mandarin, Hakka, Cantonese, Wu, Xiang, and Gan). The data used in this study were collected by the author in the urban districts of Longwen and Xiangcheng from 21 native speakers (9 males and 12 females) who were selected based on a set of strict criteria with an average age of 56.5 for males, and 50 for females [7] and[8].

Two corpora were used in this study. One is about 160 monosyllabic tokens for investigating citation tones with an average of 20 tokens for each tone. The other one is about 588 tokens for investigating tone sandhi behaviour across 64 (=8 tones * 8 tones) tonal combinations in disyllabic constructions, with an average of 12 tokens for each combination, but some combinations, in particular with tone 8, had less than 12 tokens to be processed, because the tone was less productive in this dialect. Tokens were elicited by individual speakers in Praat via a professional cardioid condenser microphone at a sampling frequency of 44100 Hz. Tonally relevant duration for each token incorporated those elements that excluded the syllable onset. Acoustically, the durational onset was set at the glottal

pulse where the amplitude of air pressure fluctuation began to increase; the periodicity of speech wave vibration appeared regular in the waveform, and the formant patterns in the spectrogram were clearly stable and identifiable. The offset was set at the point where periodicity and formant patterns cease to be visible. F0 and duration values were extracted using a script at ten equidistant sampling points in Praat.

Because acoustic signals are highly variable, the process of normalisation was applied to abstract away the variable content from the invariable linguistic content in this study, with the formula (1) z-score normalisation approach for F0 and (2) the absolute approach for the duration [7], [8] and [9]. For example, each tonal duration was expressed as a percentage of the average duration of all tones from the speaker being considered.

$$Z_i=(X_i-m)/s \quad (1)$$

$$D_{norm}=(D/D_{mean})*100 \quad (2)$$

Because this study involved 8 citation tones and 64 tonal combinations, the technique of pairwise t-test comparison by effect size [10] was employed to determine whether the variables (e.g., duration) among a set of tones (e.g., citation tones) differ from each other in a statistically significant way. For example, there were 28 (=8*7/2) paired differences to be tested in the citation context, as illustrated in Figure 1.

```
      tone1   tone2   tone3   tone4   tone5   tone6   tone7
tone2 8.6e-10 -       -       -       -       -       -
tone3 < 2e-16 < 2e-16 -       -       -       -       -
tone4 < 2e-16 < 2e-16 8.0e-06 -       -       -       -
tone5 2.3e-06 1.000   < 2e-16 < 2e-16 -       -       -
tone6 < 2e-16 < 2e-16 < 2e-16 < 2e-16 < 2e-16 -       -
tone7 < 2e-16 < 2e-16 < 2e-16 < 2e-16 < 2e-16 < 2e-16 -
tone8 1.1e-10 1.000   < 2e-16 < 2e-16 0.041   < 2e-16 < 2e-16
```

Figure 1: *Example of pairwise t-tests in citation context.*

The Bonferroni correction was applied to control for the Type I Error and achieve a significance. The corrected alpha was calculated by dividing the critical P value by the number of comparisons. For example, the corrected alpha was 0.00186 (= 0.05/28) in the citation. If the calculated t value was less than the corrected alpha, the paired difference was considered statistically significant, and vice versa. The testing result was visualised using the hierarchical clustering algorithm to help assess how many groups the sets of data can be clustered into from a scientific perspective.

3. Duration in Citation Form

Zhangzhou presents eight tones, although to fully appreciate this finding, one needs to recognise multidimensional characteristics of tonal realisations across different linguistic contexts [7] and [8]. Figure 2 shows the normalised F0 system of Zhangzhou citation tones, in which all F0 contours are expressed as a function of their corresponding normalised duration values from 21 speakers, representing the central tendency of this dialect as an independent variety. As indicated, tones in this dialect exhibit variation in both F0 contour shape and F0 height, which, on the one hand, involves four contour shapes of rising, level, mid-low level with a final fall, and falling contours, while, on the other hand, involving four contour heights of high, mid-high, mid, and low.

Like the F0, tones vary considerably in duration in this citation context. Figure 3 shows the acoustically normalised duration system of Zhangzhou in which each bar is expressed as a percentage of the average duration value of all tones from 21 speakers. Exhaustive pair-wise t-tests were conducted to compare 28 (8*7/2) paired differences under the assumption that all putative duration levels are independently and identically distributed. The testing results were visualised hierarchically, as shown at the bottom of figure 3. The threshold selected at 1 clusters the eight citation tones into four classes, with 1 representing the longest and 4 the shortest. Thus, a duration system of four levels can be achieved, which are noted as extra-long [V::], long [V:], medium [V], and short [V̌], as summarised in Table 1.

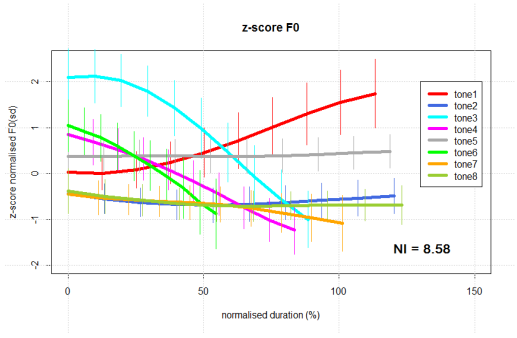

Figure 2: *Normalised F0 system of Zhangzhou citation tones.*

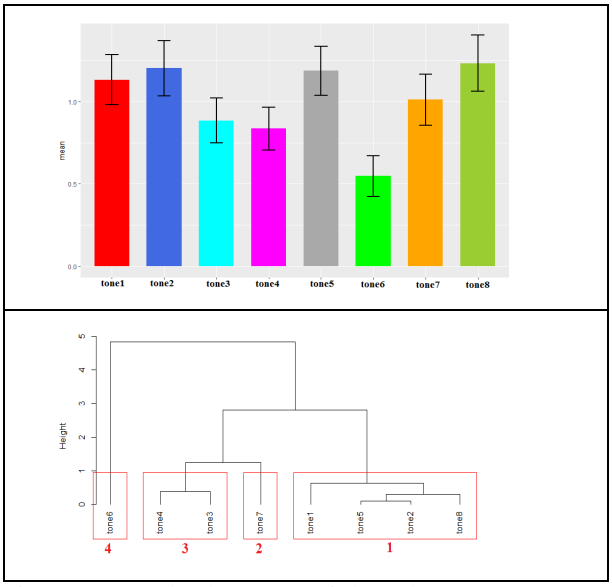

Figure 3: *Normalised Duration of Zhangzhou citation tones.*

Table 1. *F0 and duration of Zhangzhou citation tones.*

Tone	F0	Duration	Notation	Ranking
1 2 5 8	[35] [22] [33] [22]	extra long	[V::]	1
7 stopped	[221]	long	[V:]	2
3 4	[51] [41]	medium	[V]	3
6 stopped	[41]	short	[V̌]	4

Table 1 summaries the F0 and duration of these citation tones. Several aspects of linguistic significance can be observed with respect to the interaction between F0 and duration. (a) No statistically significant differences exist among the level (tones

2, 5, and 8), and among the unstopped falling tones (tones 3 and 4). In other words, the realisation of duration is not categorically affected by the changing F0 height, failing to support the across-linguistic assumption of a negative correlation between pitch height and duration. (b) The rising contour (tone 1) is grouped with the level contours (tones 2, 5, and 8), signifying no statistical hierarchy existing between rising and level-pitched tones in terms of duration. However, it is longer than any other falling contours, supporting the universal assumption that an upward F0 has a marked tendency to take a longer time than a downward change [11] and [12]. (c) The stopped tones are statistically significantly shorter than their corresponding unstopped tones that share a similar F0 realisation. For example, the stopped tone 6 shares a similar mid-high falling contour with the unstopped tone 4, but its duration is shorter. This indicates that the difference in syllable type can constrain the durational realisation because the stopped tones are associated with obstruent-ending syllables, while the unstopped tones with sonorant-ending syllables at the underlying level. (d) The stopped tones are not always the shortest as conventionally assumed in Sinitic studies [13], [14] and [15]. Instead, they can be longer than many other tones. For example, in citation, the stopped tone 7 is statistically significantly longer than tones of falling contours (tones 3, 4, and 6). (e) Tones 2 and 8 neutralise their contrasts in the citation context because of sharing common F0 and duration values. The reason why they are proposed as two different tones is because of their various realisations in other linguistic contexts, such as in the phrase-initial setting.

4. Duration in the Phrase-initial context

Zhangzhou presents a right-dominant sandhi system [7] and [8]. Tones at the non-rightmost position undergo categorical alternation but preserve their categories as their citation forms in the rightmost context. Its sandhi system is found not to be affected by the categories of surrounding tones; whereby individual tones are realised categorically the same regardless of which phonological environment they occupy; however, they may show a certain degree of phonetic variation due to co-articulation and position effects [7], and [8]. Figure 4 plots the acoustically normalised F0 system of phrase-initial tones comprising mid-level, rising and falling. Tones 2 and 3 have two statistically different variants in their offset values, which can be ascribed to the effect of regressive assimilation on the onset of their following tones. Tones 2 and 8 that share an identical F0 contour in citation can be realised differently in this context: tone 2 presents a mid-level, while tone 8 shows a mid-falling contour. As well as this, the whole F0 range is raised phrase-initially, with the average normalised values situating between -0.83 and 2.33, compared with the values between -1.44 and 2.1 in the citation context.

Figure 5 plots the normalised duration of the eight tones at the phrase-initial context across 64 (=8*8) tonal combinations, among which only tone 5 has two variants, with the value before tones 3 and 6 being statistically shorter. The nine phonetic duration levels form 36 (=9*8/2) paired differences to be tested by effective size. The result is visualised and shown at the bottom of figure 5. Thus, a system of two lengths can be achieved for the tones in this context, which are noted as medium [V] and extra-short [V̆] in Table 2. Additionally, several aspects deserve further discussion.

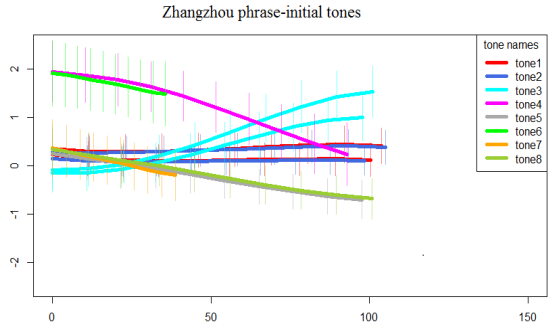

Figure 4: *F0 system of Zhangzhou phrase-initial tones.*

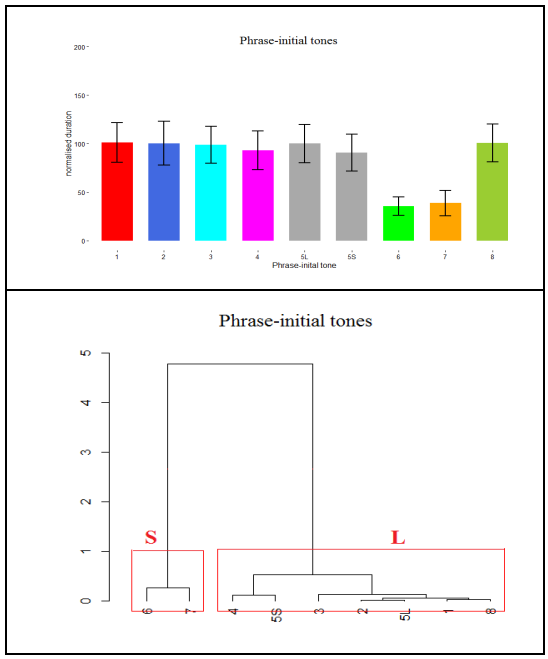

Figure 5: *Durations of Zhangzhou phrase-initial tones.*

Table 2. *F0 and duration of Zhangzhou phrase-initial tones*

Tone	F0	Duration	Notation
1	[33]	medium	[V]
2	[33]		
3	[24]/[35]		
4	[63]		
5	[32]		
8	[32]		
6 stopped	[65]	extra short	[V̆]
7 stopped	[32]		

(a) The number of lengths is considerably reduced phrase-initially. Only two lengths are tested statistically significantly different compared to four in the citation. (b) Phrase-initial tones are subject to neutralisation processes severely. The durational contrasts of the unstopped tones are neutralised to be medium. In contrast, that of stopped tones are neutralised to be extra short. (c) No significant difference exists between duration and F0. In other words, the durational realisation in this context is not affected by the F0 contour shapes and heights of surrounding tones at both phonological and phonetic levels. This can be seen from those six unstopped tones with variable

F0 contours of mid-level, low rising, extra-high falling, and mid falling have an identical medium duration. This manifestation not only falsifies the cross-linguistic assumption of a negative relation between F0 height and duration but also denies the universal tendency that considers a rising contour having a longer duration than falling one. (d) The conditioning factor on the two-length distinction appears to be solely associated with syllable structure. The medium length occurs on sonorant-ending syllables, regardless of tonal phonology and phonetics, whereas the extra short length occurs on the syllables ending in obstruent codas. (e) In this context, the parameter of duration can be seen as having a similar function as the syllable coda type that classifies syllables/tones into either stopped or unstopped categories. (f) A new duration value of an extra-short property emerges phrase initially, which can be regarded as a coupling effect of position (phrase-initial) and syllable type (obstruent-ending).

5. Duration in the Phrase-final context

The tones in the phrase-final context are presumed to maintain their categories and forms as their citation counterparts under the conventional assumption of a right-dominant sandhi system [16], [17], [18], [19], [20] and [21]. However, in Zhangzhou dialect, the right-most tonal realisations are highly sensitive to the phonetic environment of their preceding tones and presenting variation. This can be seen in figure 6, which plots the normalised F0 system of phrase-final tones across 64 tonal combinations. All tones except tone 3 have two variants, with the onset value statistically significantly higher after tones with a non-low F0 offset. Similarly, as seen in figure 7 about the normalised duration system, tones 5 and 7 are found to have two variants with one marginally longer than the other, resulting in 10 phonetic variants to be pairwise t-tested by an effect size of 45 (=10*9/2) paired differences, the result of which is visualised using the clustering algorithm. As seen, the phrase-final tones are clustered into three classes which can be labelled as extra-long [V::], long [V:], and medium [V], as summarised in Table 3. Tone 7 has a medium duration across most phrase-initial tones but has a long variant that only occurs after tone 8, which can thus be seen as a marked form.

Additionally, there also are several interesting aspects being further noted. (a) No statistically significant difference exists between rising and level contours, but both are longer than falling contours. (b) No statistically significant difference exists between low falling and high falling tones, but both are longer than the non-stopped tone 4 of a mid-high falling contour. (c) Tone 7 presents a shorter duration than tone 2, which shares similar F0 contours. (d) The number of duration differences is reduced to three (extra-long, long, and medium) from four (extra-long, long, medium, and short) in the citation. (e) The duration of most phrase-final tones is not the same as their corresponding citation form. Tones 2 and 8 have shorter durations than their citation forms, while tone 6 becomes longer in the phrase-final position. However, tones 1 and 5 maintain their extra-long property, while tone 4 keeps a medium level across citation and phrase-final contexts. This manifestation may be ascribed to the factor from the position effect and phonetic adjustment to preceding tones. (f) The durational manifestation of the phrase-final tones can also question the conventional assumption regarding the property of right-dominant tone sandhi because the forms of right-most tones do not necessarily preserve their corresponding citation forms without change.

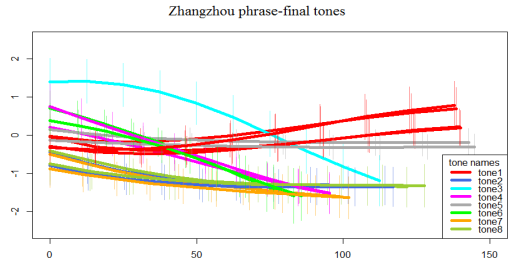

Figure 6: *F0 system of Zhangzhou phrase-final tones.*

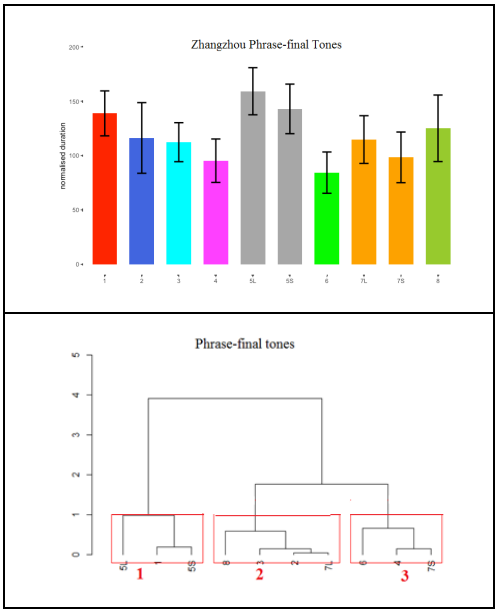

Figure 7: *Durations of Zhangzhou phrase-final tones*.

Table 3. *F0 and duration of Zhangzhou phrase-final tones.*

Tone	F0	Duration	Notation
1 5	[34]/[35] [33]/[43]	extra long	[V::]
2 3 8	[211]/[311] [52] [211]/[311]	long	[V:]
4 6 stopped 7 stopped	[41]/[51] [41]/[51] [211]/[311]	medium (marked long)	[V]

6. Conclusion

As discussed, the duration can function as F0 to distinguish Zhangzhou tones into different categories. The durational realisations are seen changing in accordance with the changing linguistic contexts and showing variation, which are also affected by the phonetics of surrounding tones and syllable types. This study fills in our knowledge of the Zhangzhou dialect but also upgrades our understanding of the relations between F0 and duration and the phonetic characteristics of a right-dominant tone sandhi system while falsifying several conventional assumptions. It contributes valuable empirical data to the typology of tonal studies in Asia and sheds important light on how humans encode and decode duration in their cognitive grammar.

7. Acknowledgements

I would like to express my sincere gratitude to Phil Rose, Mark Donohue, Paul Sidwell, and Shinichi Ishihara for their supervision on my tonal study at the ANU, to Siva Kalyan for his assistance in R, and to Michael Proctor for his valuable discussions on phonetics. I also appreciate my three anonymous reviewers for their time and very helpful comments, which I have tried to incorporate.

8. Reference

[1] Koffi, E., "A comprehensive review of the acoustic correlate of duration and its linguistic implications", Linguistic Portfolios, 10, 2021.

[2] Kubozono, H., "The phonetics and phonology of geminate consonants: Book review", Journal of the Acoustical Society of America, 142 (4), 1756, 2017.

[3] Zee, E., "Duration and intensity as correlates of F0", Journal of Phonetics, 6(3): 221-225, 1978.

[4] Kong, Q., " Influence of tones upon vowel duration in Cantonese", Language and Speech, 30: 387-399, 1987.

[5] Yu, A. C., "Tonal effects on perceived vowel duration", Laboratory Phonology,151-168, 2010.

[6] Faytak, M., & Yu, A. C., "A typological study of the interaction between level tones and duration", In E. Zee (Ed.), Proceedings of the International Congress of the Phonetic Sciences, XVII. International Congress of the Phonetic Sciences, 2011.

[7] Huang, Y., Tones in Zhangzhou: Pitch and beyond (PhD dissertation: Australian National University), 2018. Retrieved from https://openresearch-repository.anu.edu.au/handle/1885/144938, accessed on 19 June 2020.

[8] Huang, Y., Tones in Zhangzhou: Pitch and beyond, Cambridge, UK: Cambridge Scholar Publishing, 2020.

[9] Huang, Y., Donohue, M., Sidwell, P., & Rose, P., "Normalization of Zhangzhou citation tones", In C. Carignan, & M. Tyler (Eds.), Proceedings 16th Australasian International Conference on Speech Science & Technology, 217-220, Sydney, Australia: The Australian Speech Science & Technology Association,

[10] Levshina, N., How to do Linguistics with R: Data exploration and statistical analysis, Amsterdam, Netherlands: John Benjamins Publishing Company, 2015.

[11] Gandour, J., "On the interaction between tone and vowel length: Evidence from Thai dialects", Phonetica, 34: 54-65, 1977.

[12] Ohala, J. J., & Ewan, W. G., "Speed of pitch change", Journal of the Acoustical Society of America, 53, 345, 1972.

[13] Ma, C., Studies of Zhangzhou dialect 漳州方言研究, Hongkong: Zongheng Chubanshe 纵横出版社, 1994.

[14] Yang, X., Studies of tones and regional cultures of Zhangzhou dialect 漳州方言声调与地域文化研究, Beijing: Zhongguo Shehui Kexue Chubanshe 中国社会科学出版社, 2008.

]15] Guo, J., Zhangzhou Southern Min 漳州闽南方言. Zhangzhou: Zhangzhou Library 漳州图书馆, 2014.

[16] Wright, M. S., A metrical approach to tone sandhi in Chinese dialects (Doctoral dissertation, University of Massachusetts Amherst), 1983. Retrieved from http://scholarworks.umass.edu/dissertations/AAI8310348

[17] Shih, C.-L., The prosodic domain of tone sandhi in Chinese (Doctoral dissertation, University of California at San Diego), 1986. Retrieved from https://www.researchgate.net/publication/36071823_The_Prosodic_Domain_of_Tone_Sandhi_in_Chinese

[18] Ballard, W. L., "The history and development of tonal systems and tone alternations in South China (Vol. 22)", Study of Languages and Cultures of Asia and Africa: Monograph Series 22, 1988.

[19] Chen, M., Tone sandhi, Cambridge, England: Cambridge University Press, 2000.

[20] Zhang, J., "A directional asymmetry in Chinese tone sandhi systems", Journal of East Asian Linguistics, 16: 259-302, 2007.

[21] Rose, P., "Complexities of tonal realisation in a right-dominant Chinese Wu dialect—Disyllabic tone sandhi in a speaker from Wencheng", Journal of the South East Asian Linguistics Society, 9: 48-80, 2016.