INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

Same F0, Different Tones: A Multidimensional Investigation of Zhangzhou Tones

Yishan Huang

The University of Sydney
yishan.huang@sydney.edu.au

Abstract

This paper explores tonal encoding in Zhangzhou Southern Min, a Sinitic dialect spoken in South Fujian province of China. Results show that tones sharing an identical F0 contour can differ considerably in duration, vowel quality, voice quality, syllable coda type, and sandhi behavior. It is proper to consider tones are multidimensionally distinguished and treat each tone as a complex of phonetic features. The discussion stretches and deepens our understanding of the phonetic nature of tone, while questions the conventional definition of tone as a contrastive use of pitch. It contributes vital empirical data to the typology of tone as an important phenomenon in world's languages, and sheds important light on how tones should be defined.

Index Terms: tone, F0, multidimension, encoding, Zhangzhou

1. Introduction

Approximately 50% [1] or as much as 60-70% [2] of the world's spoken languages are tonal, which are largely distributed in Sub-Saharan Africa, South-Central Mexico, and East and Southeast Asia [1-4], and partly exist in Amazonia and New Guinea [1] and [5]. Tone is conventionally characterized with the perceptual dimension of pitch (as indicated in various tonological models and definitions languages [1-9], which corresponds to the fundamental frequency (F0) in acoustics and reflects the rate of vibration of vocal cords [10-12]. Thus, the issue of tonal studies is generally considered to be a business of investing how pitch varies in utterances to convey lexical and/or grammatical information. For example, the four tones in Mandarin possess four different pitch contour shapes: level (/ma55/ 'mother'), rising (/ma35/ 'numb'), dipping (/ma214/ 'horse', and falling (/ma51/ 'scold').

This conventional impression of tone is being challenged by increasing numbers of studies of unrelated languages, which assert that, beyond the difference in pitch, tonal contrasts can involve systematic difference (interaction) in (with) other parameters. For example, in Southeast Asia, those so-called register languages are reported to redundantly use pitch, voice quality, vowel quality, and durational difference to distinguish contrastive categories of words [13-19]. Such as, in Mon, an Austroasiatic language spoken throughout Myanmar and Thailand, the high-registered vowels are articulated with high pitch, shorter duration, and modal voicing, whereas chest-registered vowels with a lower pitch, longer duration, and breathy voice [15]. Each tone in Burmese is characterized by a particular phonation, such as a low, modal-voiced tone; a high, breathy tone; a high, creaky tone; and a very high, tense-voiced tone [17]. Languages in Otomanguean family are found to use pitch and phonation independently, which can multiple the number of tonal contrasts [19-22]. For example, San Felipe Jalapa de Díaz (Jalapa) Mazatec, an Otomanguean language of the Popolocan branch, contrasts three phonation types (breathy, modal, and laryngealised) and three-level tones (low, mid, and high) independently, leading to a total of nine contrasts [22]. These question the conventional definition of tones as the lexical phonemicisation of pitch distinctions.

In Zhangzhou Southern Min, a Sinitic dialect spoken in Southern Fujian of China, the single dimension of pitch appears not to be sufficient to characterize tonal contrasts, instead, various segmental and suprasegmental parameters interact to shape its tonal system. For example, tone 4 and tone 6 in this dialect both present a mid-high [41] falling contour in citation, but they behave considerably different in terms of duration, vowel quality, voice quality, syllable coda type, and tone sandhi, as summarized in Table 1. Given their substantial differences in the parameters other than pitch, should they be treated as two independent tones or a single one?

Table 1: *Phonetics of Zhangzhou tone 4 and tone 6*

Parameters	Tone 4	Tone 6
pitch/F0	[41]	[41]
duration	medium	short
phonation (high vowel)	breathy	creaky
vowel quality	monophthongisation	diphthongisation
coda type	sonorant	obstruent
sandhi pitch	[63]	[65] (extra short)

This study conducts a sophisticated investigation into how tones are encoded in Zhangzhou Southern. It aims to address three specific research questions. (1) How are tone 4 and tone 6 realised in this dialect; (2) how do segments (vowel quality, syllable coda type), suprasegments (F0, duration, voice quality), and linguistic contexts (citation, sandhi) interact to shape tonal distinctions, and (3) how can the multidimensional realizations of Zhangzhou tones be justified in acoustics? Built upon the acoustically noramlised patterns from 21 native speakers, this study is hoped to stretch and deepen the knowledge of tone in this Southern Min variety; contributing valuable empirical data to the typology of tone as an important phenomenon in human languages, while shedding an important light as to how tone should be defined.

2. Zhangzhou Speech

2.1. Zhangzhou City

Zhangzhou is a prefecture-level city situated in the south Fujian province of Southern China, covering an area of approximately 12,600 square kilometers and about 5.10 million population [23-26]. The colloquial language spoken by native speakers is predominantly Southern Min, known as Hokkien, which is mutually intelligible with Southern Min varieties of Xiamen, Quanzhou, and Taiwan but is unintelligible with other Chinese

10.21437/Interspeech.2023-2534

dialects, such as Mandarin, Hakka, Wu, and Cantonese. Certain regional variation can be observed in the sound system of its eleven administrative areas [27] and [28]. This study restricts the locality to the urban area of Longwen and Xiangcheng districts, which is considered to be historically-socially-culturally-linguistically representative of Zhangzhou [27].

2.2. Zhangzhou Syllables

A template of C(G)V(X) structure can be generalized for Zhangzhou speech in which onset (C) and nucleus (V) are compulsory while glide (G) and coda (X) are optional to occur [23-25]. Oral vowel, nasalized vowel and syllabic nasal can occur as nucleus, while nasal consonant, glide, and obstruent consonant can function as coda. Table 2 presents the segmental inventory for individual syllable components in this dialect.

Table 2: Phoneme inventory for Zhangzhou syllables.

Component		Phoneme
Onset	C	p, pʰ, ɓ, t, tʰ, ɗ, k, kʰ, ɠ, ts, tsʰ, s, z, ħ, ʔ
Glide	G	j, w
Nucleus	V	i, e, ɛ, ɐ, ɔ, ɵ, u, ĩ, ɛ̃, ɐ̃, ɔ̃, m̩, ŋ
Coda	X	j, w, m, n, ŋ, p, t, k

2.3. Zhangzhou Tones

All work prior to Huang [23] and [24]'s initiatives documented a seven-way tonal contrast [27-35] for Zhangzhou citation tones, with the pitch values differing among themselves but also from Huang's studies. For example, tone 2 that corresponds to Yangping in terms of Middle Chinese tonal category has been transcribed as [212], [13], [12], [23], and [22], covering contour shapes of dipping, rising, and level. Huang [23] proposes an eight-way tonal system based on two major facts. (a) Tones sharing a similar pitch contour can differ considerably in other parameters, such as duration, syllable type and phonation. (b) Tones having an identical realization in citation can differ in other linguistic contexts. Table 3 illustrates the eight citation tones with their pitch and duration values, in which tonal pitch is transcribed using Chao [36]'s notation system, with number 5 representing the highest pitch level of the speaker's pitch range and number 1 the lowest.

Table 3: Examples of Zhangzhou citation tones.

T	Pitch	Duration	Example 1	Example 2
1	[35]	extra-long	/tɐŋ/ 'east'	/kɔ/ 'aunty'
2	[22]	extra-long	/tɐŋ/ 'copper'	/kɔ/ 'glue'
3	[51]	medium	/tɐŋ/ 'to wait.'	/kɔ/ 'drum'
4	[41]	medium	/tɐŋ/ 'frozen'	/kɔ/ 'oversee'
5	[33]	extra-long	/tɐŋ/ 'heavy'	/ħɔ/ 'rain'
6	[41]	short	/tɐp/ 'answer'	/kɔk/ 'country'
7	[221]	long	/tsɐp/ 'ten'	/tɔk/ 'poison'
8	[22]	extra-long	/tsi/ 'tongue'	/kɔ̃ / 'snore'

3. Research Design

3.1. Data Elicitation

The data used in this study were obtained by the author on a formal linguistic field trip in 2015 [23]. Twenty-one native speakers that included 9 males and 12 females from the urban area of Longwen and Xiangcheng districts were recorded their utterances in an acoustically absorbent room at Zhangzhou Hotel. The average age for males was 56.5 and 50 for females.

The corpus study incorporated about 160 monosyllabic morphemes with an average of 20 tokens for each tone, and about 588 disyllabic phrases. Tokens were chosen across different syllable types; onsets of different manners and places of articulation and vowels of varying height and frontness. This design was to maximally balance the intrinsic perturbation effects on tonal F0 from tautosyllabic segments. Tokens to be elicited were all written in simplified Chinese characters and presented via PowerPoint, with one token in one slide, which substantially ensured a balanced intensity and well-controlled speech rate over the recording [23]. All recordings were digitized at a sampling frequency of 44100 Hz in Praat.

3.2. Acoustic Processing

The acoustic processing of obtained field data was conducted in Praat. Tonally relevant duration was identified incorporating all elements except syllable onsets [23]. Praat scripts were run to automatically extract tonal duration and F0 and the first three formant values. F0 were extracted at 10 equidistant sampling points, and the first three formant values at 9 equidistant sampling points. The amplitude values of the first three harmonics, labelled as H1, H2, and H3, were manually extracted from the last 10% portion of related tokens shown in spectral slices in Praat for voice quality examination.

Because of the highly variable acoustic signals that carried both linguistic and extralinguistic information [37-40], all extracted acoustic data were normalized to reduce the indexical variances. Formular (1) was the z-score approach for F0 [39-40]; (2) was the absolute approach for duration [40]; (3) was the Lobanov [41]'s approach for formant frequency normalization, while (4) was the z-score approach for voice quality normalization:

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

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

$$F_i= (X_i-m)/s \quad (3)$$

$$A_i = (X_i-m)/s \quad (4)$$

In these formulars, the parameters *m* and *s*, separately, stand for the raw mean value and the standard deviation estimated from all sampling values over all tokens from a given speaker. *Xi* is an observed value, while Z_i, F_i, and A_i refer to normalised F0, formant, and amplitude value, respectively. D is an observed duration value, and D_{norm} is its corresponding normalised value expressed as a percentage of the average duration of all tones from the speaker being considered.

4. Results

4.1. F0

Figure 1 presents the normalized F0 pattern of Zhangzhou citation tones as a function of their corresponding normalized duration, which represents the central tendency of Zhangzhou as an independent variety. As indicated, the system includes one rising (tone 1), three level (tones 2, 5, and 8), one mid-low level with a final fall (tone 7), and three falling contours (tones 3, 4, and 6). Tone 4 and tone 6 both present a mid-high [41] falling trend. Similarly, both tones 2 and 8 show a mid-low-level contour [22]. Given their similar F0 realization, it seems plausible to classify them into the same tonal category. The reason why they are treated as contrastive is primarily because of their considerable differences in other parameters.

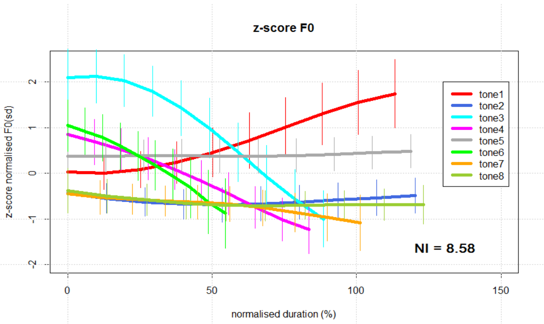

Figure 1: *Normalized F0 system of Zhangzhou citation tones.*

4.2. Duration

Zhangzhou tones vary considerably in duration. Figure 2 about the normalised duration of Zhangzhou citation tones. Tone 6 is the shortest that only occupies about 50% of the average duration of all tones over 21 speakers. The three level contours (tones 2, 5, and 8) are shown to be the longest, followed by the rising contour (tone 1). Tone 4, reaching 80% of the average value, is substantially longer than tone 6. Thus, the two tones differ considerably in their length realization.

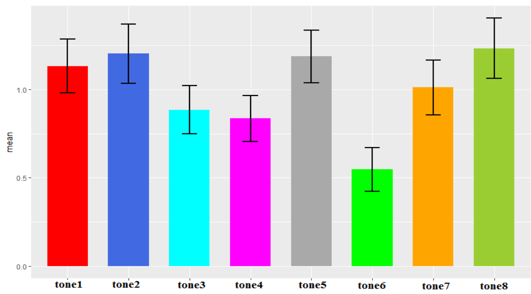

Figure 2: *Normalized duration of Zhangzhou citation tones*

4.3. Voice Quality

Vowels can have different phonation in tone 4 and tone 6. In tone 4, the high vowel is found to be breathy, however, in tone 6, the phonation is alternated to be creaky. Figure 3 shows the normalised spectral tilt patterns of monosyllabic morphemes that contain the high vowel /i/ in tone 4 and tone 6. As seen, in tone 4, the normalised H1-H2, H2-H3 and H1-H3 are all steeply positive, signifying a dramatic drop in the amplitude as the frequency increases over time. These reflect an increased open quotient and a less abrupt glottal closing gesture, causing a persistent leakage of airflow and energy to be distributed dominantly in the fundamentals and to decrease in the higher frequency regions [42-45]. These suggest a breathy phonation of the high vowel in tone 4.

The lower figure plots the phonation realization of high vowel /i/ across different obstruent codas in tone 6. As seen, the intervals of glottal pulses become irregular and less frequent during the last 10% of the tokens, indicating that the vocal folds vibrate at a relatively low and less frequent rate [42-45]. The normalised H1-H2, H2-H3 and H1-H3 are all negative, with the H1-H2 value approaching zero (-0.1 s.d.). These two aspects indicate a typical characteristic of creaky phonation in this dialect [23]. The reason why the high vowel is pronounced with a creaky voice resides in the fact that the high vowel /i/ is found undergoing diphthongization and becoming [iɛ̯] with a target ending in a mid-low position of the mouth, which creates an articulatory basis for the creaky voice to occur.

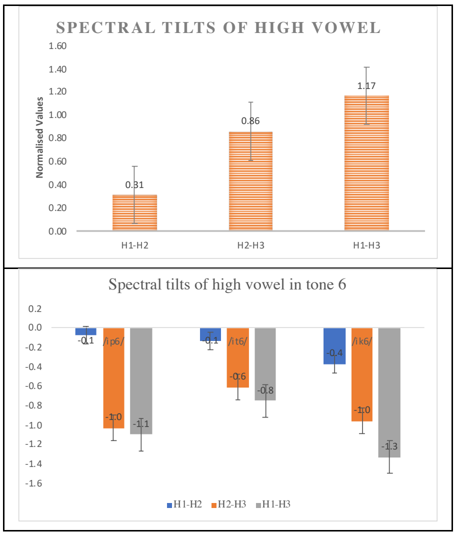

Figure 3: *Waveforms and spectral tilt patterns of /i/ vowel*

4.4. Vowel Quality

Vowels can be realised differently in tones 4 and 6. The two high vowels (/i/ and /u/) undergo an active alternation between monophthongs ([i] and [u]) in tone 4 and diphthongs ([iɛ̯] and [uɤ̯]) in tone 6. The diphthongisation indicates two articulatory trajectories; while suggests the acoustic correlates of vowels—mainly the first two formants change as a function of time [47-48]. Figure 4 plots the normalised formant patterns of the high vowels in tones 4 and 6.

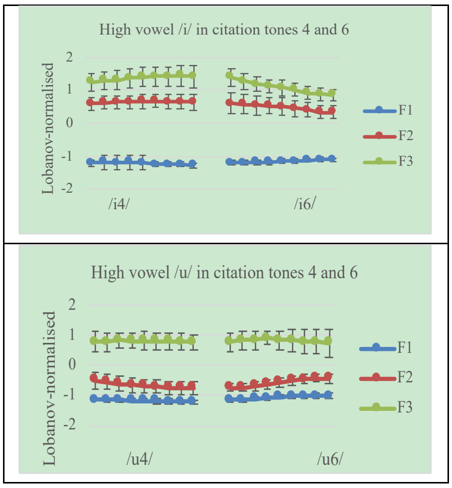

Figure 4: *Normalized formant patterns of high vowels.*

In tone 4, the F1 and F2 curves are relatively steady without significant changes across the whole duration in both vowels /i/ and /u/, indicating that the tongue remains in the same position during the articulation. Thus, the two high vowels are realised as monophthongs. In tone 6, the formant patterns are changed considerably. In /i6/, the F1 is raised up from slightly below -1 to above, while the F2 is lowered down from slightly below 1 to above 0. The increasing F1 and decreasing F2 values imply that the tongue moves to a lower and more central position, so that this /i/ is realised as a diphthong [iɛ̯]. In /u6/, both F1 and F2 curves show an upward change during their second halves, indicating a downward but also frontal tongue movement, indicating two articulatory trajectories. Thus, this high back vowel /u/ is realised as a diphthong [uɤ̯] in tone 6.

4.5. Syllable codas

Syllable coda type serves as an important parameter to classify Sinitic tones into stopped and unstopped categories. As summarized in Table 4, the unstopped tones are associated with sonorant-ending syllables, such as tone 4, whereas the stopped tones are related to those obstruent-ending syllables, such as tone 6. Thus, tones 4 and 6 can be distinguished in terms of the syllable coda type that they have at the underlying level.

Table 4: Classification of Sinitic tones with respect to syllable coda type.

Tones	Glide	Nasal	Obstruent
unstopped	+	+	-
stopped	-	-	+

What needs a special attention is that, the obstruent codas in this dialect turn out not to be realised at the uttance-final context, but are identifiable at the non-utterance-final setting. This can be seen in Figure 5. In the upper figure, the normalised formant patterns are consistently stable without significant changes over the major portions of syllables across the three different rhymes (/ɐp7/, /ɐt7/, and /ɐk7/) in the citation context, indicating the non-realization of obstruent coda. In contrast, in the lower figure, the formant patterns are considerably different over the last 10% of the duration across different rhyme types in the phrase-initial morphemes, indicating the articulation of obstruent codas of different places of articulation.

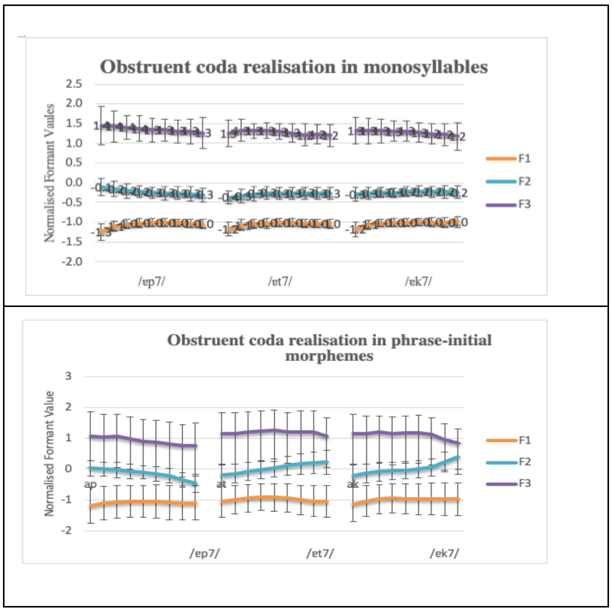

Figure 5: *Obstruent coda realisations in Zhangzhou.*

4.6. Tone Sandhi

Zhangzhou possesses a right-dominant tone sandhi system [23] and [24]. The rightmost boundary of the sandhi domain is aligned with the rightmost boundary of a syntactic phrase XP, where X is categorically variable, ranging over adjective (A), noun (N), verb (V), preposition (PP), and adverbial (Adv) among others. Tones at the non-right-most (sandhi) position of change their realisation phonologically and phonetically, while tones at the rightmost position maintain the realisation categorically similar to their citation forms but may subject to certain variation because of the co-articulation and position effects. Figure 6 plots the normalised F0 patterns of tones 4 and 6 as a function of their following tones.

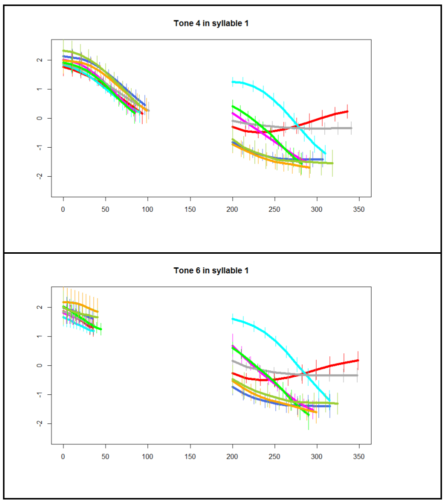

Figure 6: *Normalized F0 patterns of tone 4 and tone 6 in the sandhi position across eight following tones.*

Tone 4 presents an extra-high falling [63] contour that departs from about 2 s.d. to somewhere slightly over the midpoint across all following tones. Tone 6 shows an extra-short and extra-high-falling contour [65] between 2 s.d. and 1 s.d. above the midpoint. The F0 onset of both tones is considerably higher than the onset of its following tone 3, which is the highest among the tones at the rightmost position; thus, the number 6 is initiated to indicate the extra-high F0 level in this dialect [23] and [24]. As indicated, tone 4 and tone 6 differ from each other in two aspects: (a) tone 4 has a lower offset, which is nearly 1 s.d. lower than tone 6; and (b) the duration of tone 4 is nearly twice longer than that of tone 6. Therefore, the two tones can be distinguished in the sandhi context.

5. Discussion

As discussed, although having the same F0 realization, tone 4 and tone 6 in Zhangzhou can be distinguished in several other parameters including duration, vowel quality, voice quality, syllable coda type, and their sandhi behavior. This renders the conventional definition of tone as the lexical phonemicisation of pitch distinction inadequate for understanding their nature. Seen in this light, it appears to be more appropriate to consider Zhangzhou tones as being multidimensionally distinct. Each phonological tone can be seen as a complex of segmental and suprasegmental features that function together to encode tonal distinction. For example, tone 4 indicates a mid-high falling contour, medium duration, breathy high vowel, no high vowel diphthongization, and sonorant coda in citation, and an extra-high falling contour with a medium duration in the sandhi position. Tone 6 suggests a mid-high falling contour, short duration, creaky high vowel, diphthongization, obstruent coda, and an extra-high-and extra-short contour in sandhi.

This study provides a new profile as to the tonal encoding in Sinitic languages. It substantially stretches and deepens our understanding of the multidimensional nature of tone as an important language phenomenon in this Southern Min variety, while contributing vital empirical data to the typological study of tone in world's languages. The discussion in this study also questions the conventional definition of tone as a contrastive use of pitch contour, while shedding an important light on how tone should be defined for a fuller understanding of their nature.

6. References

[1] L. M. Hyman. "Tone: Is it different?" In J. A. Goldsmith, J. Riggle, & A. C. Yu (Eds.), *Handbook of phonological theory*, pp. 197-239). New York, NY: Wiley & Sons, 2011.

[2] M. Yip. *Tone*. Cambridge, England: Cambridge University Press, 2002.

[3] K. L. Pike. Tone languages. A technique for determining the number and type of pitch contrasts in language, with studies in tonemic substitution and fusion. Ann Arbor: University of Michigan Press, 1948.

[4] W. S.-Y. Wang. "Phonological features of tone". *International Journal of American Linguistics*, 33, pp. 93-105, 1967.

[5] M. Donohue. "Tone systems in New Guinea". *Linguistic Typology*, 1, pp. 347-368, 1997.

[6] Z. Bao. *The structure of tone*. New York, NY: Oxford University Press, 1999.

[7] S. Duanmu. *The Phonology of Standard Chinese*. Oxford, England: Oxford University Press, 2000.

[8] I. Maddieson. "The transcription of tone in the IPA". *Journal of the International Phonetic Association*, 20, pp. 28-32, 1990.

[9] M. Halle and K. N. Stevens. "A note on laryngeal features". In M. Halle (Ed.), In *Phonology and Phonetics [PP]: From memory to speech and back: Papers on phonetics and phonology* , PP.1954 - 2002. Berlin, Germany, de Gruyter, 1971.

[10] J. J. Ohala. "Producition of tone". In V. A. Fromkin (Ed.), *Tone: A linguistic survey*, pp. 5-32. New York, NY: Academic Press, 1978.

[11] P. Rose. "Acoustic characteristics of the Shanghai-Zhenhai syllable-types". In D. Bradley (Ed.), *Papers in South-East Asian linguistics*, No.8: Tonation, pp. 1-53. Canberra, Australia: Pacific Linguistics, 1982.

[12] P. Rose. "A linguistic phonetic acoustic analysis of Shanghai tones". *Australian Journal of Linguistics*, 13, pp.185-219, 1993.

[13] C. T. DiCanio. "The phonetics of register in Takhian Thong Chong". *Journal of the International Phonetic Association*, 39, pp.162-188, 2009.

[14] M. Brunelle. "Tone perception in Northern and Southern Vietnamese". *Journal of Phonetics*, 37, pp.79-96, 2009.

[15] M. Brunelle and J. Kirby. "Tone and phonation in Southeast Asian languages". *Language and Linguistics Compass*, 10 (4), pp.191-207, 2016.

[16] M. Brunelle and T. Tấn. "Register in languages of Mainland Southeast Asia: state of the art". In P. Sidwell & M. Jenny (Ed.), *The Languages and Linguistics of Mainland Southeast Asia: A comprehensive guide.* Berlin, Boston: De Gruyter Mouton, pp.683-706, 2021.

[17] D. R. Mortensen. "Tonally conditioned vowel raising in Shuijingping". *Journal of East Asian Linguististics*, 22, pp.189-216, 2013.

[18] J. E.Andruski and M. Ratliff. "Phonation types in production of phonological tone: The case of Green Mong". *Journal of the International Phonetic Association*, 30, pp.37-61, 2000.

[19] C. Esposito and S. Khan. "The cross-linguistic patterns of phonation types". *Language and Linguistics Compass*, 14 (12), pp.1-25, 2000.

[20] C. M. Esposito. "An acoustic and electroglottographic study of White Hmong tone and phonation". *Journal of Phonetics*, 40, pp.466-476, 2012.

[21] A.Kelterer and B. Schuppler. "Phonation type contrasts and tone in Chichimec". *The Journal of the Acoustical Society of America*, 147(4), pp.3043-3059, 2020.

[22] M. Garellek and P. Keating. "The acoustic consequences of phonation and tone interactions in Jalapa Mazatec". *Journal of the International Phonetic Association*, 41, pp.185-205, 2011.

[23] Y.Huang. *Tones in Zhangzhou: pitch and beyond*. Doctoral thesis: Australian National University, 2018.

[24] Y. Huang. *Tones in Zhangzhou: pitch and beyond*. Cambridge, UK: Cambridge Scholars Publishing, 2020.

[25] Y. Huang. Zhangzhou Southern Min: Syllables and Phonotactics. Lincom Europa, 2021.

[26] Y. Huang. "Pairwise T-Test: Identification of Tonal Relations". In *Preceedings of the International Conference on Asian Language Processing (IALP 2022),* 27-28 October 2022, Chinese University of Hong Kong (Shenzhen), China; the Chinese and Oriental Languages Information Processing Society, Singapore, 2022.

[27] C. Ma. *Studies of Zhangzhou dialect.* Hongkong: Zongheng Chubanshe, 1994.

[28] X. Yang. *Studies of tones and regional cultures of Zhangzhou dialect.* Beijing, China: Zhongguo Shehui Kexue Chubanshe, 2008.

[29] T. Dong. *Four Southern Min varieties.*Taipei: Zhongyang Yanjiuyuan, 1959.

[30] FJG. *Fujian Province Gazette-Dialect Volume.* Beijing, China: Fangzhi Chubanshe, 1998.

[31] ZZG. *Zhangzhou City Gazette-Dialect Volume (Vol. 49)*. Beijing, China: Zhongguo Shehui Kexue Chubanshe, 1999.

[32] B. Lin. "Zhangzhou vocabularies". *Fangyan*, pp.1-3, 1992.

[33] C. Zhou. *The great Southern Min dictionary*. Fuzhou, China: Fujian Renmin Chubanshe, 2006.

[34] J. Guo. *Zhangzhou Southern Min.* Zhangzhou, China: Zhangzhou Library, 2014.

[35] R. Gao. "Introduction to the sound system of Zhangzhou". In *Minnan dialect-studies of Zhangzhou variety, pp.*109-116. Beijing, China: Zhongguo Wenlian Chubanshe, 1999.

[36] Y. Chao. "A system of "tone letters"". *Le Maître Phonétique*, 45, pp.24-27, 1930.

[37] J. Harrington. "Acoustic phonetics". In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd ed), pp.81-129. Hoboken, NJ: Wiley-Blackwell, 2010.

[38] P. Foulkes, J. M. Scobbie, and D. Watt. "Sociophonetics". In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd ed), pp.703–754. Hoboken, NJ: Wiley-Blackwell, 2010.

[39] R. Rose. "The normalization of tone". Proceedings of the First Australian Conference on Speech Science and Technology, Canberra, pp.130-135, 1986.

[40] Y. Huang, M. Donohue, P. Sidwell, and P. Rose. "Normalization of Zhangzhou citation tones". In C. Carignan, & M. Tyler (Eds.), *Proceedings 16th Australasian International Conference on Speech Science & Technology*, pp.217-220. Sydney, Australia: The Australian Speech Science & Technology Association, 2016.

[41] B. M. Lobanov. "Classification of Russian vowels spoken by different speakers". *The Journal of the Acoustical Society of America*, *49(2B),* pp. 606-608, 1971.

[42] K. Stevens. "Physics of laryngeal behavior and larynx modes". *Phonetica, 34*, pp. 264-279, 1977.

[43] M. KHuffman. "Measures of phonation type in Hmong". *Journal of the Acoustical Society of America, 81*, pp. 495-504, 1987.

[44] E. B. Holmberg, R. E. Hillman, J. S. Perkell, P. C. Guiod, and S. L. Goldman. "Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice". *Journal of Speech, Language, and Hearing Research, 38*, pp. 1212-1223, 1995.

[45] J. E. Andruski and M. Ratliff. "Phonation types in production of phonological tone: The case of Green Mong". *Journal of the International Phonetic Association, 30*, pp. 37-61, 2000.

[46] J. Laver. *Principles of phonetics*. Cambridge, England: Cambridge University Press, 1994.

[47] J. Harrington. "Acoustic phonetics". In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences (2nd)*, pp. 81-129. Hoboken, NJ: Wiley-Blackwell, 2010.

[48] K. Johnson. *Acoustic and auditory phonetics (3rd)*. Hoboken, NJ: Wiley-Blackwell, 2011.