Tones in Zhangzhou:

Pitch and Beyond

by

Yishan Huang

A thesis submitted for the degree of

Doctor of Philosophy of The Australian National University

April 2018

This thesis is my own original work except where otherwise indicated.

Yishan Huang

Acknowledgments

The completion of this thesis would not have been possible without inspiration and support from a number of wonderful people over the years. I wish to offer my most heartfelt gratitude to them.

I would like to express a deep sense of gratitude to my informants who contributed their utterances for this study and/or for the local vocabulary documentation for future studies. They are Ruizhu Lin, Yayuan Lu, Maoyong Liu, Meiling Yang, Sumei Chen, Bingren Yan, Zhibin Huang, Yuanfa Wang, Tonglong Huang, Yueshen Huang, Jianguo Wang, Huixue Huang, Muchi Huang, Rongxia Lin, Baoqing Zhang, Jinhong Huang, Zhupen Lin, Chunhua Lin, Qiumei Huang, Yonghui Huang, Shuixuan Huang, Ronghua Huang, and Hairong Yan. In the same vein, I would like to extend my sincere gratitude to the local communities, especially to Xiadianmei and Huaqiao University, whose utmost effort and commitment contributed to a very productive, successful and enjoyable field trip in Zhangzhou in 2015.

I am forever indebted to my academic supervisors throughout the process. I owe my deepest gratitude to Dr. Paul Sidwell, who has been a great mentor for both my professional and personal life. None of my achievements thus far would have been possible without his inspirational supervision, infinite patience, and immense knowledge. He has commented thoroughly on every single draft and on every aspect, ranging from theoretical interpretation, empirical data discussion to academic writing instruction. My gratitude to him is beyond words.

I am thankful to Dr. Mark Donohue for giving access to and valuable guidance on both the abstract world (phonology) and the real world (phonetics) in which the IPA symbols and diacritics were frequently not sufficient to transcribe the sounds that may not be produced by humans but rather by his loving dog Flash. I also thank Mark for his insightful comments, constructive criticisms, marvellous presentation skills, and also for the hard questions that have immensely influenced this research and my thinking as a linguist candidate.

I am especially grateful to Prof. Phil Rose for his brilliant hand-in-hand instructions with respect to how to write R codes to extract and plot tonally relevant information in acoustic signals and how to construct the results in a coherent and informative way. His exceptional expertise in tonal studies, valuable feedbacks, and intelligent advices were a significant influence in sharpening many of the concepts presented in this thesis, and enlightening my pursuit to be a phonetician.

I would like to thank my colleague Dr. Siva Kalyan greatly for providing me with substantial technical assistance on speech processing, especially on the statistical testing using exhausted pairwise *t*-test comparison by effect sizes. Without his help, the extensive statistical testing undertaken in this thesis would not have been possible. I would like to extend the gratitude to his parents for their very warm concern and encouragement shown over the years.

I would like to express my comprehensive appreciation to my other panel members Prof. Wayan Arka, Prof. Shunichi Ishihara, Dr. Julia Miller, Dr. Gwendolyn Hyslop, and Dr. Cathryn Donohue for their professional instructions and supportive assistances at different stages of my candidature. I appreciate Shun in particular for accepting me as a PhD student at ANU, and for his very inspirational quote "You can do it!" I forever appreciate Gwen for welcoming me at the airport and showing me around the campus and Canberra upon my arrival in Australia for the first time. I forever cherish the powerful hugs and encouragements from Julia no matter whether I was too excited, too stressed, or just in a normal state. I know Cathryn is always there for me with her sparkling thoughts, charming smile and unconditional support. I am grateful to Wayan for taking over the chairmanship at the last stage. His enoumous calmness, kindness, and patience impressed me very much. His course in Syntax is one of the most informative, engaging, and fruitful courses I have ever taken. I owe Wayan a piece of work on Zhangzhou Grammar.

I am also deeply grateful to many other academic staff members, including Prof. Nick Evans, Dr. Mark Ellison, Dr. Beth Evans, Dr. Jennifer Hendricks, Prof. Jane Simpson, Dr. Cynthia Allen, and Dr. Zhengdao Ye, for their very enlightening and enjoyable lectures, discussions, comments, and/or seminars, which have extensively broadened my knowledge beyond phonetics, and beyond the Sinitic language family. I appreciate Nick in particular for broadening my horizons on one of the most linguistically diverse areas in the world—the area where Papuan languages are spoken—in his course, and training me to investigate the language systems from the typological point of views. I owe Nick a piece of typological work on tones.

I would like to thank Ms. Joanne Bushby in particular for her very inspirational guidance and assistance on the administrative matters and her heart-warming concern for my personal life since the first day of my studies at ANU. I extend my particular gratitude to Ms. Chintana Sandilands, Dr. Xiaoli Zhou, Dr. McComas Taylor, and Mr. Mark Jones for their wise counsel, caring concern, and sympathetic ear during my stay in BPB building where I wrote up the first draft and completed the submission of the thesis. I would like to extend my sincere gratitude to Prof. Chris Ballard, Prof. Paula Newitt, Prof. Ann Kumar, Prof. Peter Kanowski, Dr. Peter Friedlander, Prof. Yongxian Luo, Mr. Michael Zaja, and Mr. Tony Karrys for all their concern, encouragemnt and assistance during the past years.

I am also especially grateful to the ANU security team, especially to John, Bob and Oscar, who frequently escorted me back to Unihouse after the midnight, and encouraged me massively during the final year. Many nice people in the CAP IT team and CHL admin, such as Celin, Shahandra, Sandhi, Ethan, Anthony, and Christina Gee, also deserve my particular appreciation. Special appreciation should also go to the loving neighbourhood in Coombs, BPB, Unihouse and the phonetic group. Without their marvellous assistance and fabulous fellowship, the writing journey for this thesis would not have been enjoyable.

I would like very much to extend my sincere gratitude to my MA supervisors/advisors/lecturers in China, especially to Prof. Feng Shi, Prof. Yingwei Guan, Prof. Xiaonong Zhu, Prof. Xiangdong Shi, and Prof. Xiaoyu Zeng, who guided me into a world full of linguistically intriguing sounds, and motivated me to explore what the sounds are, and why they exist. My sincere thanks also go to Prof. Larry Hyman, Dr. Susan Lin, Prof. Eric Bakovic and Prof. Aijun Sun who immensely extended my knowledge of phonetics and phonology at the 2015 LSA Summer Institute. I would also like to appreciate my three anonymous examiners for their valuable comments and suggestions for this thesis and for future studies.

I owe a big "thank you!" to lots of loving and caring friends inside and outside ANU for the wonderful times we shared, the beautiful friendships we made, and for the warmest concern and encouragement I received. They are Charlotte van Tongeren, Katerina Naitoro, Hedvig Skirgard, Darja Hoenigman, Fitrilailah Mokui, Gita Nasution, Penelope Judd, Donna Holland, Saowapha Viravong, Daniel Forest, Teena Saulo, Benny Tong, Nguyen Bui, Myra Abubakar, Christ Weedall, Kwang-Ju Cho, Owen Edwards, Matthew Carroll, Fanny Cottet, John Mansfield, Eri Kashima, Tina Gregor, Yoko Yonezawa, Thoms Honeyman, Sydney Kingstone, Yuri Takhashi, Jeong Yoon Ku, Yu-Hua Chen, Evi Eliyanah, Bryce Kositz, Chris Park, Weichen Huang, Xueqing Norah Zhong and her family, Yu Sang and her family, Xiaofei Tang and her family, Yuan Jin and her family, Jinghong Zhang and her family, Yuewu Huang and his family, Yujie Zhu, Xuan Dong, Geng Li, Liang Chen, Naijing Liu, Juncheng Guo, Lijun Gao, Wendi Xue, Bo Liu, Yang Hu, Ran Li, Ding Yuan, Kaiyun Cai, Shuang Zhang, Linliang Qian, Meixi Zhuang, Ian Zhang, Dan Liu, Julia Yang, Yuqi, Karen Tu, Zibo Miao, Ge He, Luni Xu, Xiang Lei, Changxin Cong, Lyn Ning, Ning Zhao, Lisa among so many others.

Most important of all, my deepest and sincerest gratitude go to my grandparents, parents, brother, and many loving and respectful relatives. This thesis would not have been possible without their unparalleled love, support, encouragement, and blessing since the day I enrolled as a primary student. I am forever indebted to my beloved parents—Jinhui Huang and Caizhu Huang—for giving me the opportunities and experiences that have made me who I am. They selflessly encouraged me to explore the directions that I am interested in and supported me to realise my pursuits. I will never be able to pay them back for their constant unconditional love, care, affection, pain and sacrifice. I dedicate this milestone to them, especially to my father, who is my ultimate role model.

.

Abstract

This study draws on various approaches—field linguistics; auditory and acoustic phonetics; and statistics—to explore and explain the nature of Zhangzhou tones, an under-described Southern Min variety. Several original findings emerged from the analyses of the data from 21 speakers. The realisations of Zhangzhou tones are multidimensional. The single parameter of pitch/F0 is not sufficient to characterise tonal contrasts in either monosyllabic or polysyllabic settings in Zhangzhou. Instead, various parameters, including pitch/F0, duration, vowel quality, voice quality, and syllable coda type, interact in a complicated but consistent way to code tonal distinctions.

Zhangzhou has eight tones rather than seven tones as proposed in previous studies. This finding resulted from examining the realisations of diverse parameters across three different contexts—isolation, phrase-initial, and phrase-final—, rather than classifying tones in citation and in terms of the preservation of Middle Chinese tonal categories. Tonal contrasts in Zhangzhou can be neutralised across different linguistic contexts. Identifying the number of tonal contrasts based simply on tonal realisations in the citation environment is not sufficient. Instead, examining tonal realisations across different linguistic contexts beyond monosyllables is imperative for understanding the nature of tone.

Tone sandhi in Zhangzhou is syntactically relevant. The tone sandhi domain is not phonologically determined but rather is aligned with a syntactic phrase XP. Within a given XP, the realisations of the tones at non-phrase-final positions undergo alternation phonologically and phonetically. Nevertheless, the alterations are sensitive only to the phrase boundaries and are not affected by the internal structure of syntactic phrases.

Tone sandhi in Zhangzhou is phonologically inert but phonetically sensitive. The realisations of Zhangzhou tones in disyllabic phrases are not categorically affected by their surrounding tones but are phonetically sensitive to surrounding environments. For instance, the pitch/F0 onsets of phrase-final tones are largely sensitive to pitch/F0 offsets of preceding tones and appear to have diverse variants.

The mappings between Zhangzhou citation and disyllabic tones are morphologically conditioned. Phrase-initial tones are largely not related to the citation tones at either the phonological or the phonetic level while phrase-final tones are categorically related to the citation tones but phonetically are not quite the same because of predictable sensitivity to surrounding environments. Each tone in Zhangzhou can be regarded as a single morpheme having two alternating allomorphs (tonemes), one for non-phrase-final variants and one for variants in citation and phrase-final contexts, both of which are listed in the mental lexicon of native Zhangzhou speakers but are phonetically distant on the surface.

In summary, the realisations of Zhangzhou tones are multidimensional, involving a variety of segmental and suprasegmental parameters. The interactions of Zhangzhou tones are complicated, involving phonetics, phonology, syntax, and morphology. Neutralisation of Zhangzhou tonal contrasts occurs across different contexts, including citation, phrase-final, and non-phrase-final. Thus, researchers must go beyond pitch to understand tone thoroughly as a phenomenon in Southern Min.

Table of Contents

List of Tables

List of Figures

Part A: Preliminaries

Chapter 1: Introduction

This chapter provides an overall introduction of this thesis. It is organised as follows. Section 1.1 introduces the research goals of this thesis. Section 1.2 discusses the research opportunities contributing to the formulation of this thesis. Section 1.3 reviews tonological studies. Section 1.4 reviews previous studies on tonetics, especially how tonal pitch interacts with other phonetic parameters. Section 1.5 reviews the studies on Sinitic tone sandhi. Section 1.6 discusses previous studies on Southern Min tone sandhi. Section 1.7 reviews previous studies on Zhangzhou. Section 1.8 introduces the structure of this thesis.

1.1. Research Goals

This thesis is designed to explore the nature of tones in Zhangzhou, an under-described Southern Min variety. Five major research questions guided this study:

(1) How many tones are there in Zhangzhou?
(2) What are tones essentially realised in Zhangzhou?
(3) How do tones interact in constructions beyond monosyllables?
(4) How are disyllabic tones related to citation tones phonologically?
(5) How are various forms of Zhangzhou tones structured abstractly?

The first three questions are addressed using interacting multi-methodological approaches, including linguistic fieldwork, auditory and acoustic phonetics, and statistics. A multidimensional framework is proposed to investigate what Zhangzhou tones are essentially realised in utterances. Three linguistic contexts—citation, phrase-initial and phrase-final—are examined to reveal how tones interact with each other; and how many tones are there in this language. Linguistic theories are drawn upon to explain how tones are related to each other across different contexts, how native speakers structure different forms of tonal realisations in their mental grammar, and how different linguistic levels interact to shape Zhangzhou tone sandhi.

The phonetic study, supplemented with statistics, is essential for observing, understanding, and generalising the multidimensional realisations and complicated interactions of Zhangzhou tones, using graphic, quantified, and testable patterns. The theoretical analysis benefits from the descriptive results of phonetics and is expected to achieve a higher level of prediction of how those cognitive aspects of tone exist in the mental system, and to explain the ways in which the cognitive aspects are mapped onto their phonetic parametric spaces and vice versa.

Four specific research goals are expected to be achieved in this study. First, this study aims to fill research gaps in tonal studies of Zhangzhou using modern linguistic methodologies and theories.

Second, it is designed to broaden existing knowledge of Southern Min tone sandhi based on a systematic and sophisticated phonetic investigation with a new set of data. Third, it is intended to improve our understanding of the nature of tone as an important linguistic category. Fourth, it will hopefully expand the empirical domain for typological surveys of tone in the East and Southeast Asia.

1.2. Research Opportunities

Four major research opportunities contributed to the formulation of this project. First, Zhangzhou is tonally interesting but is under-described. The language makes good use of pitch contour to distinguish lexical meanings. For example, as shown in Table 1-1, the syllabic nasal can bear five different pitch shapes to differentiate five different lexical items. Nevertheless, the single dimension of pitch appears inadequate to characterise all tonal contrasts, but rather, a variety of segmental and suprasegmental parameters function together to construct the tonal distinctions. For example, the citation tones 4 and 6 appear not distinguishable in terms of pitch, but they differ considerably in a variety of other parameters, as indicated in Table 1-2, in which the symbol * signifies no tokens were found.

Table 1-1. Contrastive use of pitch in the lexical distinctions in Zhangzhou

ʔm̩35 ‘to drink’	ʔi35 ‘he/she’
ʔm̩22 ‘flower bud’	ʔi22 ‘move’
ʔm̩51 ‘aunt’	ʔi51 ‘chair’
ʔm̩41 ‘affirmative’	ʔi41 ‘intention’
ʔm̩33 ‘negative’	ʔi33 ‘play’

Table 1-2. Comparison of multidimensional realisations of citation tones 4 and 6 in Zhangzhou

Parameter	Tone 4	Tone 6
Pitch	[41]	[41]
Length	medium	short
Vowel quality	no diphthongisation	high vowel diphthongisation
High vowel	breathy	creaky
Mid vowel	modal/creaky	*
Low vowel	creaky	creaky
Syllable coda	sonorant	not realised obstruent

In addition, tonal interactions in multisyllables reflect complicated interfaces among different linguistic levels—phonetics, phonology, syntax and morphology. For example, as Table 1-3 shows, tone 3 has a falling pitch [51] in citation, but the pitch is alternated to [25] at the non-rightmost position of a syntaxtic phrase (XP) and to [55] before the nominal and/or diminutive morpheme /ʔɐ/.

Table 1-3. Examples of pitch alternations of Zhangzhou tones across different linguistic contexts

Context	Tone 3		Tone 7	
Citation	[51]	kɐw51 'dog'	[22]	ʥjɔk22 'jade'
Non-XP-final	[25]	kɐw25.ɓɵ51 'bitch'	[32]	ʥjɔk32.tsu35 'jade bead'
Before /ʔɐ/	[55]	kɐw55.ʔɐ51 'little puppy'	[35]	ʥjɔk35.ʔɐ51 'little jade'

Nevertheless, this language is under-described with respect to what tones are essentially realised in utterances, how tones interact in multisyllables, and how tones are structured abstractly. Several impressionistic studies have been conducted on its citation tones, but they largely provide only a list of pitch contrasts, and the results are inconsistent (see Section 4.3). Few acoustic studies can be found in the literature, but they all exhibited problematic aspects in terms of research design and analysis (see Section 5.4). No empirical and theoretical study has systematically explored the nature of tones beyond monosyllables (see Section 6.4). The interesting tonal phenonmeona, along with their under-description, provide a valuable opportunity in this study to explore and explain the nature of Zhangzhou tones, and to fill in the research gaps.

Second, the existing interpretations of Southern Min (SM) tone sandhi are indeterminate and controversial. No theoretical work has been conducted on Zhangzhou tone sandhi, but substantial studies have been done on the sandhi phenomenon, known as *SM tone circle*, in other SM varieties of Xiamen (Amoy) and Taiwan (see Section 1.6). Nevertheless, little agreement has been reached concerning several fundamental issues, in particular,

- What is the directionality of Southern Min tone sandhi?
- How can Southern Min tone sandhi domain be specified?
- How can the nature of Southern Min tone sandhi be interpreted theoretically?

The indeterminacy and challenge in the explanation of the nature of Southern Min tone sandhi require a systematic phonetic study, using a new set of data, to explore what SM tones are realised in utterances, and to examine the existing competing proposals. Zhangzhou, as a typical Southern Min variety, provides a fertile ground to extend the understanding of the nature of tone sandhi in Southern Min.

Third, a sophisticated phonetic study is imperative for a better understanding of how tones should be represented abstractly. Various phonological models have been proposed to represent tones (see Section 1.3); however, a substantial amount of disagreement exists with respect to several basic issues. For example, what is tone? What are the primitive tonal features? What is the tone bearing unit? How are tonal features defined? How are tonal features structured underlyingly?

These problems call for a crucial investigation of what tones are essentially realised in phonetics for a better understanding of the nature of tone as an important linguistic category.

Last but not the least, being a native-speaking linguist candidate offers several advantages. As a native speaker of Zhangzhou, I have personal interest in investigating my mother tongue and have the internalised competence and performance of using this language. As a native-speaking linguist candidate, I have the intuition and sensitivity to capture subtle details of linguistic significance that have not yet been found or documented properly. In addition, the linguistic training received helps reveal and explain the mechanism behind the utterances in a systematic way.

Considering the four specific research opportunities, the study of Zhangzhou tones is expected to be linguistically significantly interesting. It is hopped it will factor out the problems of inconsistency and inadequacy of the previous studies of Zhangzhou tones using modern linguistic methodologies and theories, but also will deepen the understanding of the nature of Southern Min tone sandhi, and the nature of tone as an important linguistic category.

Sections 1.3-1.7 provide relatively comprehensive reviews on tonology, tonetics, Sinitic tone sandhi, Southern Min tone sandhi, and previous studies of Zhangzhou. It is hoped that the reviews will establish a solid research foundation for this thesis to build on, and also offer insights about what has been done and what needs to be improved for the understanding of the nature of tone in general.

1.3. Tonology

Approximately 50% (Hyman, 2011, p. 198) or as much as 60-70% (Yip, 2002, p. 1) of the world's spoken languages are tonal, which are largely distributed in Sub-Saharan Africa, south-central Mexico, and East and Southeast Asia (Pike, 1948; Wang, 1967; Yip, 2002; Hyman, 2011), and partly exist in Amazonia and New Guinea (Donohue, 2003; Hyman, 2006; 2011). The definition of what a tonal language is varies among linguists (e.g., Pike, 1948; Gandour, 1978; Yip, 2000; Hyman, 2006, 2011; Maddieson, 2013). For example:

> A tone language may be defined as a language having lexically significant, contrastive, but relative pitch on each syllable (Pike, 1948, p. 3).
>
> A tone language is language in which pitch is used to contrast individual lexical items or words (Gandour, 1978, p. 41).
>
> A language is a "tone language" if the pitch of word can change the meaning of the word (Yip, 2000, p. 1).
>
> A language with tone is one in which an indication of pitch enters into the lexical realisation of at least some morphemes (Hyman, 2006, p. 229).
>
> Tone is the term used to describe the use of pitch patterns to distinguish individual words or the grammatical forms of word (Maddieson, 2013).

Similarly, linguists have devoted themselves to addressing the issue of how tones can be represented formally, as shown in Table 1-4; however, consensus has not yet been reached on some essential issues. For example,

- What are the nature of the primitive features of tones?
- What is the tone-bearing unit?
- Are tonal features binary or privative?
- How are different features arranged?

Table 1-4. Comparisons of different models for tonal feature representation

Model	Number of feature	Number of level	Number of register	Register feature	Pitch feature
Wang, 1967	7	5	1	[high], [central], [mid], [contour], [rising], [falling], [convert]	
Woo, 1969	3	5	1	[high], [low], [modify]	
Maran, 1971	2	3	1	[raised F0], [lowered F0]	
Halle & Stevens, 1971	2	3	1	[stiff], [slack]	
Yip, 1980	2	4	2	[upper]	[high]
Clements, 1981	2	4	2	[high]	[low]
Yip, 1989	2	4	2	[upper]	[raised]
Bao, 1990	2	4	2	[stiff]	[slack]
Duanmu, 1990	4	9	3	[stiff],[slack]	[above],[below]
Chang, 1992	3	4/6	2	[stiff]	[constricted glottis], [spread glottis]
Hyman, 1993	1/2	9	3	[high], [low]	[high], [low]
Fu, 1995	2	5	3	[high], [low]	[high], [low]

1.3.1. What are the primitive features of tones?

Many attempts have been made to formulate a set of distinctive features for the underlying tonal representation of Asian tonal languages (e.g., Wang, 1967; Woo, 1969; Yip, 1980, 1989; Bao, 1990; Duanmu, 1990; Chang, 1992; Fu, 1995) and African tonal languages (e.g., Maran, 1971; Halle & Stevens, 1971; Clements, 1981; Hyman, 1993). Nevertheless, the questions remain as to what are the primitive features for tones; how many features are needed for an adequate account, and how many pitch levels and registers can be distinguished? For example, as indicated in Table 1-4, the various models can be essentially classified into one-register (Wang, 1967; Woo, 1969; Halle & Stevens, 1971; Maran, 1971), two-register (Yip, 1980; Clements, 1981; Yip, 1989; Bao, 1990; Chang, 1992), and three-register systems (Duanmu, 1990; Hyman, 1993; Fu, 1995).

In the one-register system, tonal features were arranged as part of the feature matrix of vowels while, in the multi-register systems, different features for tones were structured hierarchically, separated from their bearing units on a different tier, and separate feature(s) for register were introduced to encode the articulatory correlation between tonal pitch and segments. Consensus has not yet been reached on how many contrastive pitch levels are needed (four, five, or nine), and how many registers are distinguished (two or three), although the diversity of tonal systems across linguistic areas may also constrain the development of a unified interpretation.

1.3.2. How are tonal features defined?

The distinctive features have specific phonetic bases for their definitions (Hyman, 1975; Clark & Yallop, 1990); however, different proposals for tonal representations vary in how tonal features can be defined perceptually, acoustically, or by articulation, as discussed below.

1.3.2.1 Perceptual features

Wang (1967), Yip (1980), Duanmu (1990), Hyman (1993), and Fu (1995) adopted the perceptual dimension of pitch to define tonal features in their models; however, their terminologies varied considerably. For example, Yip (1980) used [upper] and [low] while Duanmu (1990) used [above] and [below]. Wang (1967) used [high], [mid], and [central] to account for five pitch levels but used [contour], [falling], [rising], and [convex] to distinguish different shapes of pitch contour.

1.3.2.2 Articulatory features

Articulatory features are used to encode the physiological correlations between tonal and segmental production (Halle & Stevens, 1971; Duanmu, 1990; Bao, 1990; Chang, 1992; Zhu, 2012). For example, Halle and Stevens (1971) used the laryngeal features [stiff] and [slack] to specify three pitch levels associated with three states of vocal cord tension. Nevertheless, no consensus exists on what motivates the physiological correlations between tones and segments. For example, Halle and Stevens (1971) related pitch production with the states of vocal cord tension, and Duanmu (1990) correlated them with the thickness of vocal cord tension. Bao (1990) suggested the muscle activities of vocal cords while Chang (1992) indicated a different status of the glottis. Thus, how tonal features can be defined from the articulatory point of view still remains an open question.

1.3.2.3 Acoustic features

Acoustic features are less frequently used to specify tonal features in the literature, presumably because of the highly variable property of fundamental frequency (F0), the acoustic correlate of pitch. It appears that only Maran (1971, p. 9) adopted [raised F0] and [lowered F0] to account for tones in African languages.

1.3.3. Are tonal features binary or privative?

The distinctive features in phonology are often assigned values to indicate their presence or absence in specific entities. Two different ways have been observed with respect to how tonal features are valued. Most models (e.g., Wang, 1967; Woo, 1969; Yip, 1980, 1989; Clements, 1981; Hyman, 1993; Bao, 1990; Duanmu, 1990; Chang, 1992) considered tonal features as binary, given both values “+” and “-” could be equally active in making tonal contrasts. For example, the opposition of high and low tones was represented as [+high] or [-high] (Yip, 1980; Hyman, 1993) or [-slack] or [+slack] (Halle & Stevens, 1971; Bao, 1990). Nevertheless, this binary treatment is claimed to be over-generative and inter-determinate (Clements, 1983; Hyman, 1978; Fu, 1995). For example, Duanmu (1995)'s model is regarded able to generate as many as nine contrastive pitch levels, overpredicting the possible maximum of five levels observed cross-linguistically (Wang, 1967).

Alternatively, some other scholars (e.g., Clements, 1985; Hyman, 1978; Fu, 1995) claimed the privative features for a simple and economic explanation, because they considered only the positive value was able to fulfil the linguistic function while the negative value was phonologically inert. For example, Fu (1995) eliminated extraneous negative values in his model. Thus, high tone/register was specified as [high], and low tone/register was specified as [low] in his model while no feature was specified for the neutral tone or register.

The privative approach appears to have an advantage in eliminating redundancy toward an economical representation. Nevertheless, whether features should be doubled-valued or single-valued depends on several factors, for example, whether the tonal system investigated is African or Asian. Further factors include what the inventory of tonal contrasts is and how the feature geometry is constructed.

1.3.4. What is the tone-bearing unit?

The literature has offered different answers to the question of what is the tone-bearing unit (TBU), including the mora (Duanmu, 1990; Hyman, 1993; Chang, 1992; Fu, 1995), the syllabic segment (Woo, 1969; Maran, 1971; Halle & Steven, 1971; Bao, 1990), and the syllable (Wang, 1967; Yip, 1980). The question of what domain a tonal feature is aligned with at the underlying level is not a simple one. It is a problem that may trigger a series of theoretical constraints for an economic and adequate account of tonal representations, in particular,

- How many tones can a TBU take?
- What is the tone-bearing ability of a TBU?
- Why is the existence of contour tones greatly constrained?
- How does a TBU accommodate tonal contour complexity?
- What are the relations among rhyme structure, tonal contour complexity, and the TBU?

For example, the moraic theory was proposed to accommodate the relation between rhyme structures and tonal contour complexity (Duanmu, 1990; Hyman, 1993; Chang, 1992; Fu, 1995), which reflected the tone-bearing ability of syllable rhymes, as Duanmu (1990, p.125) asserted,

> When the tone-bearing ability is increased, the rime must be lengthened; when the rime is shortened, the tone-bearing ability is reduced.

1.3.5. How are different features arranged?

In the theory of feature geometry, distinctive features are assumed to be structured hierarchically, rather than stored in a simple matrix (McCarthy, 1988; Clements, 1985). In previous tonal models, the features for register and tone are treated in different ways, giving rise to differences in the internal structures of the geometry trees, as shown in Figure 1-1.

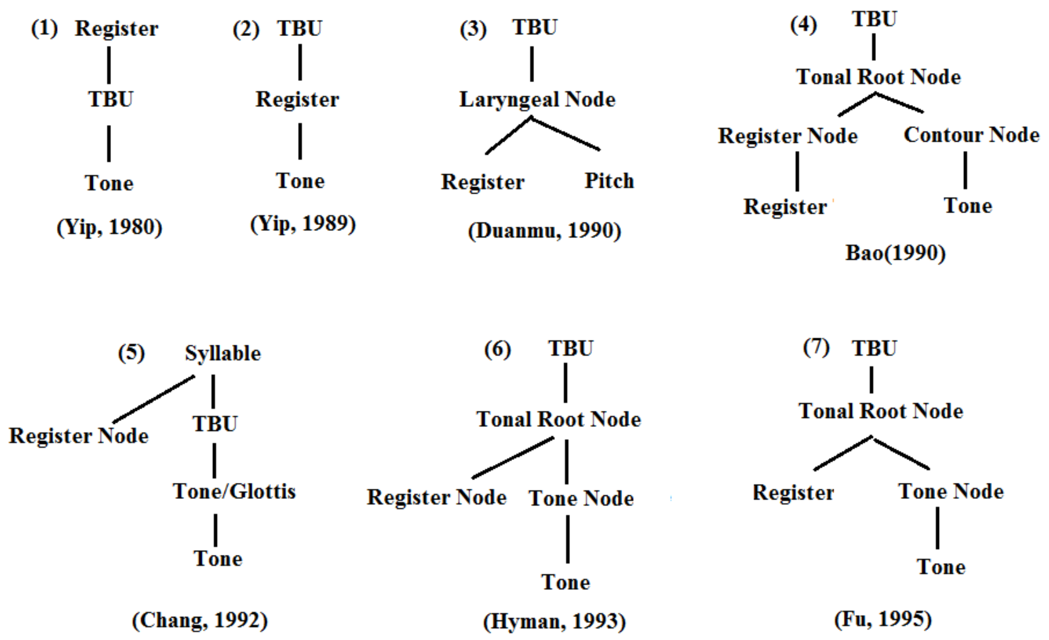


Figure 1-1. Comparison of the internal structures of different tonal models.

As indicated in Figure 1-1, the register and tone tend to be formulated either in a parallel or a dominant correlation. Some proposals (Yip, 1980; Bao, 1990; Duanmu, 1990; Fu, 1995) indicated tone and register are parallel, thus the existence of tone is independent of the existence of register. For example, Yip (1980) asserted that register and tone are independent autosegmentally, so the register feature may remain stable while the tone feature changes in a sandhi process of deletion.

Alternatively, some proposals (Yip, 1989; Hyman, 1993) treated tone and register as hierarchical, with the register dominating tone. This dominant relation is proposed based on the assumption that register and tone may simultaneously undergo a process, such as spreading. Because multiple instances of spreading are generally prohibited in feature geometry (Clements, 1985), it seems appropriate to require register to dominate tone, and moving a register node means the features for tone have to be moved simultaneously.

As discussed, considerable disagreements can be identified with respect to how tones should be represented. The representability of most existing models may also be challenged by the polydimensional reality of tonal category, as increasingly revealed in cross-linguistic data (see Section 1.4, and this thesis). Thus, conducting a sophisticated phonetic study is imperative for a better understanding of how tones should be better represented abstractly, which is one of the major aims to be achieved in this thesis.

1.4. Tonetics

As indicated in various tonological models, tonal contrasts are largely characterised with the perceptual dimension of pitch corresponding to the acoustic dimension of fundamental frequency (F0), reflecting the rate of vibration of vocal cords (Rose, 1982, 1993). Thus, the differences in tones are often assumed to be a matter of differences in pitch or F0 (Kong, 1987). Nevertheless, increasingly, studies of unrelated tonal languages have been revealing that, beyond the differences in pitch/F0, tonal contrasts often involve systematic interaction with other phonetic parameters, for example, duration, aspiration, phonation, vowel quality, and manner of initial consonants (e.g., Schuh, 1978; Rose, 1982; 1988; Whalen & Levitt, 1995; Fox, 2000; Connell, 2002; Faytak & Yu, 2011; Zhu, 2012). This section reviews how tonal pitch/F0 interacts with other phonetic parameters to create tonal realisations and distinctions.

1.4.1. Pitch/F0 and vowel quality

Research on the interaction between pitch/F0 and vowels can be categorised into two areas. One concerns the intrinsic correlation between pitch/F0 and vowel height (Hombert, 1978; Connell, 2002; Fox, 2000; Whalen & Levitt, 1995; Hoole & Hu, 2004) while the other considers the occurrence restriction of vowels with respect to different tones (Yip, 1980, 2002; Ping, 1995; Chan, 1985; Myers & Tsay, 2003; Donohue, 2007, 2013; Mortensen, 2013).

First, it has been attested cross linguistically that high vowels tend to have a higher fundamental frequency than do low vowels (Connell, 2002; Fox, 2000; Whalen & Levitt, 1995). For instance, Whalen & Levitt (1995) conducted a survey on 31 languages from 11 language families to investigate the intrinsic interaction between F0 and vowel quality. Results justified the general assumption of a positive correlation between vowel height and F0 height, and also indicated the intrinsic F0 varies as a function of vowel height and having no significant correlation with the dimension of vowel backness.

Second, vowel quality has been reported undergoing alternation in accordance with specific tonal environments, for example, in the Mang dialect of Shuijingping (Mortensen, 2013) and Fuzhou (Yip, 1980, 2002; Ping, 1995; Chan, 1985; Myers & Tsay, 2003; Donohue, 2007; 2013). Specifically, a vowel-raising process was found in Fuzhou with respect to the higher pitched tones, as shown in Table 1-5 (Myers & Tsay, 2003, p. 119).

Table 1-5. Vowel alternations with respect to tones in Fuzhou

Tone	Vowel alternation								
LM, LML	æ	ai	au	ay	ei	ou	œy	ieu	uoi
H, HL, M	ɛ	ei	ou	oy	i	u	y	iu	ui

1.4.2. Pitch/F0 and duration

The correlation between vowel duration and pitch has been studied cross-linguistically (e.g., Zee, 1978; Kong, 1987; Yu, 2010; Faytak & Yu, 2011). In general, vowels on low tones are found to have a longer duration than they are on high tones, and vowels on rising tones appear longer than on falling or level tones (Yu, 2010; Faytak & Yu, 2011). 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.

Counter-examples are also found showing a positive association between pitch and duration, such as in Taiwanese (Zee, 1978; Kong, 1987), Cantonese (Faytak & Yu, 2011), and Yucatec Maya (Faytak & Yu, 2011). Specifically, the mid-level tone in Cantonese is observed longer than the mid-low level tone (Faytak & Yu, 2011).

1.4.3. Pitch/F0 and consonant

The intrinsic correlation between pitch/F0 and (prevocalic) consonants has been widely reported in the literature (e.g., Halle & Stevens, 1971; Hyman & Schuh, 1974; Hombert, 1976, 1978; Abramson & Erickson, 1978; Ohala, 1978; Rose; 1996; Thurgood, 2002; Yip, 1980, 1989, 2002; Xu & Xu, 2003; Lai & Jongman, 2005; Francis et al., 2006; Lai et al., 2009; Ladefoged & Disner, 2012). For example, Hyman & Schuh (1974, p. 110) proposed a phonetic hierarchy showing continuous consonant-induced perturbative effects ontonal pitch as shown in Figure 1-2.

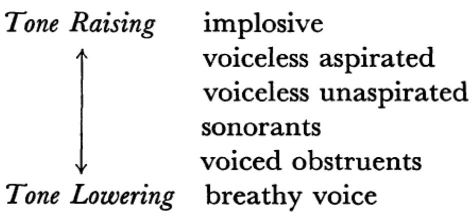


Figure 1-2. Phonetic hierarchy of consonant-induced perturbative effects on tonal pitch/F0 realisation.

As indicated in this figure, the prevocalic voiced obstruents are treated to have a depressing effect on the pitch/F0 onset of following vowels while their voiceless counterparts can raise the onset (Hombert, 1978; Abramson & Erickson, 1978; Yip, 2002; Lai & Jongman, 2005). The voicing-induced perturbation effect on pitch/F0 generally has been considered a main motivator of tonal development in Chinese (Yip, 2002), Thai (Abramson & Erickson, 1978), Vietnamese (Thurgood, 2002), and many languages in Tibetan-Burman, Austronesian, and Hottentot languages in South Africa (Hombert, 1978, p. 79 and reference therein). Specifically, the concept of Yin and Yang register in Sinitic languages reflects the association between tonal pitch and voicing status of syllable onsets in historical times, and is still commonly used to classify spontaneous tones.

In addition, aspiration (positive lag of voice onset timing) is generally considered to have a raising perturbation effect on the pitch/F0 of following vowels in languages regardless of whether they are tonal or not (Hombert, 1976, 1978; Yip, 2002; Xu & Xu, 2003; Lai & Jongman, 2005; Francis et al., 2006; Lai et al., 2009). For example, Lai et al. (2009) found the F0 is statistically significantly higher in the aspirated than in the unaspirated environment in Taiwanese.

Nevertheless, some languages are reported as showing conflicting results (Hombert, 1976, 1978). For example, the aspiration-induced perturbation effect on the F0 onset is found not systematic in American English and French (Hombert, 1976), and Hindi has a higher F0 after voiceless, unaspirated obstruents rather than after aspirated obstruents (Hombert, 1978).

1.4.4. Pitch/F0 and phonation

The function of phonation in tonal systems has received substantial attention (e.g., Ladefoged, 1971; Laver, 1980; Huffman, 1987; Andruski & Ratliff, 2000; Gordon & Ladefoged, 2001; Pan, 2005; Garellek & Keating, 2011; Esposito, 2010, 2012; DiCanio, 2009, 2012; Zhu, 2012; Kuang, 2013; Mortensen, 2013). Ladefoged (1971) proposed a continuum of phonation types with respect to the openness of glottal status, as shown in Figure 1-3 (Gordon & Ladefoged, 2001, p. 1).

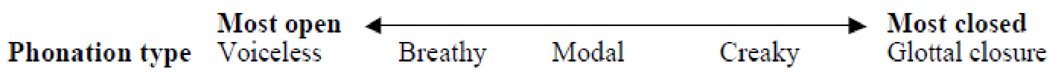


Figure 1-3. Continuum of phonation types.

Cross-linguistic studies have revealed three different functions that phonation generally plays in tonal languages (Andruski & Ratliff, 2000; Gordon & Ladefoged, 2001; Kuang, 2013; Garellek & Keating, 2011; DiCanio, 2012; Zhu, 2012; Mortensen, 2013).

First, non-modal phonation occurs as an allophonic variant of modal phonation, depending on a specific tone, for example, in White Hmong (Esposito, 2012), White and Blue Hmong (Huffman, 1987), Green Hmong (Andruski & Ratliff, 2000), and Itunyoso Trique (DiCanio, 2012). For example, each Burmese tone is characterised by a particular type of phonation, including a low, modal-voiced tone; a high, breathy tone; a high, creaky tone; and a very high, tense-voiced tone (Mortensen, 2013, p. 204).

Second, non-modal phonation intersects independently with the tonal system, possibly multiplying the number of tonal contrasts (Andruski & Ratliff, 2000; Gordon & Ladefoged, 2001; DiCanio, 2009; Garellek & Keating, 2011; Kuang, 2013; Mortensen, 2013). For example, San Felipe Jalapa de Díaz (Jalapa) Mazatec contrasts three phonation types (breathy, modal, and laryngealised) and three level tones (low, mid, and high) independently (Garellek & Keating, 2011). Another possibility is that non-modal phonation is occasionally associated with a certain tone among some speakers and for some occasions; for example, the third tone in Mandarin is produced occasionally with creaky phonation (Kuang, 2013).

1.4.5. Pitch/F0 and syllable coda

The presence of syllable obstruent codas is generally considered able to constrain the number of tonal contrasts and determine which tone should appear in many tonal languages in East and Southeast Asia (e.g., Benedict, 1948; Pike, 1948; Leiste, 1976; Gandour, 1978; Ballard, 1988; Chen, 2002; Zhang, 2007; Ratliff, 2015). For example, Hyman (2012, p. 191) summarised the limiting effect of obstruent codas on tonal contrasts in several Sino-Tibetan languages, as shown in Figure 1-4.

Language	*classification*	*"smooth" syllables*	*stopped*	*ratio*	*codas*
Bola	Lolo-Burmese	H, L, HL	H, HL	3:2	p t k ʔ m n ŋ
Maru	Lolo-Burmese	H, M, L	H, L	3:2	p t k ʔ m n ŋ
Tangkhul	Naga	H, M, L	M, L	3:2	p t k m n ŋ r w y
Trung	Nungish	H, L, HL, LH	H	4:1	p t k ʔ m n ŋ l r
Jingpho	Jingpho	H, M, L, HL	H, L	4:2	p t k ʔ m n ŋ
Karen (Pa'o)	Karenic	H, M, L, HL	M, L	4:2	p t k ʔ m n ŋ
Xiamen	Chinese	44, 24, 22, 21, 53	4, 32	5:2	p t k ʔ m n ŋ w y
Cantonese	Chinese	53~55, 33, 22, 21, 35, 23	5, 4, 3	6:3	p t k m n ŋ w y

Figure 1-4. Limiting effect of obstruent codas on tonal contrasts in Sino-Tibetan languages.

As indicated in Figure 1-4, a full inventory of tonal contrasts is preserved in "smooth" syllables that end in sonorant codas while the number of tonal contrasts is reduced greatly in "stopped" syllables that end in obstruent codas at the underlying level (Hyman, 2012, p. 191). In addition, the tonal contrasts found in stopped syllables tend to be a subset of those shown in smooth syllables. Thus, tones appear asymmetrically distributed with respect to the syllable coda types. Nevertheless, how the asymmetry of tonal distribution can be accounted for formally can give in the literature. For example, the tonemic approach considers the stopped tones and smooth tones as two distinct categories (Sagart, 1998; Chen, 2000) while the allotonic approach treats the stopped tones as an allotonic subset of the smooth tones (Benedict, 1948; Hyman, 2012).

As discussed, the close interactions between tonal pitch and other phonetic parameters have been increasingly reported; however, it is not common to see empirical studies that have incorporated various segmental and suprasegmental parameters to investigate the nature of tonal realisations, especially in Sinitic languages. This thesis is devoted to advancing tonal studies by proposing a multidimensional framework to examine what tones are essentially realised across different linguistic contexts, and how various parameters interact to shape tonal distinctions.

1.5. Sinitic Tone Sandhi

The realisation of tones can be alternated when tones interact with each other in utterances beyond monosyllables. The process of contextually triggered tonal alternation generally is referred to as *tone sandhi* in the linguistic literature (Benedict, 1948; Pike, 1948; Leiste, 1976; Gandour, 1978; Ballard, 1988; Chen, 2002; Zhang, 2007; Ratliff, 2015). For example,

> Under the circumstances appropriate to a given language, a syllable may have its normal toneme removed and a different one substituted for it. (Pike, 1948, p. 22)

> [I]ts phonetic realisation may be influenced by the presence or type of tone on an adjacent syllable (or word). This phenomenon is referred to in linguistic literature under the name of tone sandhi. (Leiste, 1976, p. 231)

> [T]ones change their shapes due to the effect of neighbouring tones upon one another. (Gandour, 1978, p. 55)

> Tone sandhi describes phonetically conditioned morphotonemic alternations at the junction of words or morpheme. (Chen, 2000, p. F37)

> Tone sandhi refers to tonal alternations conditioned by adjacent tones or by prosodic or morphosyntactic position in which the tone occurs. (Zhang, 2007, p. 259)

A tone is referred to generally as a *citation tone* when its monosyllabic-bearing unit is produced in isolation or in a prominent position in a tone group, and it is called a sandhi tone when its monosyllabic-bearing unit is used in a non-prominent position (Chen, 2002). In this study, citation tones are specifically referred to those tones produced in isolation. Surveys of cross-linguistic data have revealed instances of tone sandhi are typologically diverse and dynamic, which is generally reflected in the motivation of tonal alternation, the dominancy of the sandhi pattern, and the selection of sandhi domain in which various processes of tonal alternation operate.

1.5.1. Motivation of tone sandhi

Tone sandhi can be motivated by different linguistic factors across different languages (Ballard, 1988; Chen, 1987, 2000; Yip, 1980, 1995, 2002; Duanmu, 2005; Zhang, 2010, 2014), mainly including the following:

- Phonological environment of an adjacent tone, and
- The morphosyntactic contexts where the tone occurs

First, the tone sandhi can be phonologically conditioned by a specific tonal category. For example, the context triggering tonal alternation of Mandarin tone 3 is phonological in nature (Chen, 2000). The pitch realisation of tone 3 is alternated to that of tone 2 when preceding another tone 3, as illustrated in (1; Chen, 2000, p. 21).

(1)	a.	mai3	ma3	‘to buy a horse’	b.	mai2	ma3	‘to bury a horse’
		214	214	base form		35	214	base form
		35	214	sandhi form		35	214	sandhi form

Alternatively, it can be morphosyntactically determined (Ballard, 1988; Chen, 1987, 2000; Duanmu, 2005; Zhang, 2010, 2014). For example, Wenzhou Wu exhibits a special sandhi pattern for exhortatives and vocatives and another pattern for diminutives (Ballard, 1988). Puxian (Min) shows separate sandhi patterns for reduplicated and triplicated adjective besides the general sandhi paradigm (Ballard, 1988).

The sandhi realisations conditioned by different morphosyntactic categories are generally not predictable and present irregular patterns (Yip, 2002; Chen, 2000). Nevertheless, the irregularity and unpredictability of such sandhi paradigms offer much to deepen understanding of the overall organisation of the grammar in general and the interface between morphosyntax and phonology in particular (Chen, 2000).

1.5.2. Dominancy of tone sandhi

A considerable number of studies have investigated how sandhi processes are implemented with respect to dominancy (Wright, 1983; Shih, 1986; Ballard, 1988; Rose, 1994, 1995, 2004, 2011, 2016; Duanmu, 2005; Chen, 2000; Zhang, 2007). Generally, the tone sandhi system in Sinitic languages can be classified as either left-dominant or right-dominant, depending on the position that syllables retain the forms of citation tones while keeping the full range of tonal contrasts within a sandhi domain.

In the left-dominant sandhi system, tonal realisations on the initial syllables are retained and syntagmatically extended rightwards over the entire sandhi domain, regardless of how many syllables are there in the domain and how those syllables are structured, prosodically or morphosyntactically (Ballard, 1998; Chen, 2000; Duanmu, 2005; Zhang, 2007). For example, in Shanghai of Wu dialect, the pitch of the initial syllables determines the pitch contour of the entire domain, as shown in (2; Duanmu, 2005, p. 3).

(2)	a.	se	pe	‘three cups’	b.	sz	pe	‘four cups’
		HL	HL	base form		LH	HL	base form
		H	L	sandhi form		L	H	sandhi form

In the right-dominant pattern, tonal realisations on the final syllables are considered preserved, while those of non-final tones are paradigmatically replaced by their corresponding sandhi forms (Wright, 1983; Shih, 1986; Ballard, 1988; Chen, 2000; Zhang, 2007; Rose, 2016). The right-dominant sandhi pattern is commonly found in Southern Wu and most Min in China (Zhang, 2007; Rose, 2016). For example, in Wuyi of Southern Wu, those tonal contrasts are preserved on the prominent-final position but are significantly reduced to only two simple high and low levels on the first syllables, as shown in Figure 1-5 (Zhang, 2007, p. 262).

σ_1 \ σ_2	24	213	53	31	55	13
24	$55\text{-}T_{\sigma 2}$					
213						
53						
31	$11\text{-}T_{\sigma 2}$					
55						
13						

Figure 1-5. Example of the right-dominant tone sandhi in Wuyi.

Some Sinitic languages (e.g., Wuxi, Wenzhou, Zhenhai, Shainghai, Lianyungang, etc.) exhibit bidirectional sandhi patterns (Ballard, 1988; Rose, 1995; Zhang, 2007). In general, sandhi patterns fall under either left-dominant or right-dominant, but a special sandhi pattern with a different direction can also be observed, which tends to be morphosyntactically conditioned (Ballard, 1988; Rose, 1995; Zhang, 2007). For example, in Shanghainese, the modifier-noun phrase (NP) has a left-dominant sandhi pattern while the verb-object phrase (VP) exhibits a right-dominant sandhi pattern, as shown in (3, Duanmu, 2005, p. 6).

(3)	a.	tsho	ve	'fried rice' (NP)	b.	tsho	ve	'to fry rice' (VP)
		LH	LH	base form		LH	LH	base form
		L	H	sandhi form		L	LH	sandhi form

1.5.3. Selection of tone sandhi domain

The process of tone sandhi is characteristically domain sensitive; however, the selection of tone domain appears language dependent. Each prosodic unit, such as metrical foot, prosodic word, or phase, has the potential to be an operational domain for sandhi processes (Wright, 1983; Shih, 1986; Ballard, 1988; Chen, 1987, 1992, 2000; Duanmu, 1990; Yip, 1993, 2002). Many Wu dialects (e.g., Wuxi, Danyang, and Shanghai) and some Min dialects (e.g., Fuzhou) are reported having a prototypical stress-sensitive tonal system (Ballard, 1988; Chen, 2000; Duanmu, 1990; Shih, 1986; Wright, 1983). Within a given stress foot, all unstressed syllables lose their original tonal contrasts while the stressed syllables preserve their lexical tones and extend them to the rest of the metrical unit (Wright, 1983; Shih, 1986; Chen, 2000). The sandhi domain in Southern Min is generally taken to be a phonological phrase in the prosodic hierarchy but is syntactically determined, reflecting an interface between phonology and syntax (see Section 1.6.2).

As reviewed, conducting a survey of the tone sandhi of Sinitic language family not only entails examining what tones are realised, but also revealing and explaining how the sandhi domain is defined; what its dominancy is; and what motivates tone sandhi to occur in utterances.

1.6. Southern Min Tone Sandhi

Southern Min, as a major branch of the Min language in the Sino-Tibetan language family, has long been asserted exhibiting an intricate tone sandhi phenomenon known as the *Southern Min tone circle*, and has attracted substantial studies from different perspectives, including

- Generative phonology (e.g., Wang, 1967; Cheng, 1968)
- Autosegmental phonology (Yip, 1980; Shih, 1986)
- Optimality theory (e.g., Yip, 2002; Mortensen, 2002; Hsieh, 2005; Barrie, 2006; Thoms, 2008)
- Lexicalised phrasal phonology (Tsay & Myers, 1996)
- Psychological reality and production experiment (e.g., Hsieh, 1976; Wang, 1992; Tsay & Myers, 1996; Zhang et al., 2006; Zhang, et al., 2007; Chen et al., 2010)

- Interface between phonology and syntax (Chen, 1987; Lin, 1994; Zhang, 1993; Chen, 2000; Duanmu, 2005).

Published studies largely addressed the following four aspects,

- Directionality of Southern Min tone sandhi (e.g., Tsay, 1991, 1994; Hashimoto, 1982)
- Definition of Southern Min sandhi domain (e.g., Chen, 1987; Lin, 1994; Zhang, 1993)
- Rule- and Constraint-based interpretation of Southern Min tone circle (e.g., Wang, 1967; Yip, 1980; Shih, 1986; Yip, 2002; Mortensen, 2002; Hsieh, 2005; Barrie, 2006; Thoms, 2008)
- Lexicon-based interpretation of Southern Min tone sandhi (e.g., Hsieh, 1976; Wang, 1992; Tsay & Myers, 1996; Zhang et al., 2006; Zhang et al., 2007; Chen et al., 2010).

Thus, numerous attempts have been made to account for the nature of Southern Min tone sandhi, primarily focusing on the variety of either Xiamen or Taiwan; however, little agreement has been reached with regard to each aspect. The following section reviews previous work to build a solid foundation for this thesis.

1.6.1. Domain of SM tone sandhi

The specification of tone sandhi domain in Southern Min reflects a close interaction between phonology and syntax. The SM sandhi domain is generally considered as syntactically related (Chiu, 1931; Cheng, 1968; Chen, 1987; Lin, 1994; Zhang, 1993), but conflicting opinions also exist. Chiu (1931, p. 24, as cited in Chen, 1987, p. 114) assumed SM sandhi domain (also known as tone group) to be in a one-to-one correspondence to syntactic constituents, and the citation tones marked the end of morphosyntactic units. Chen (1987, p. 147) claimed the prosodic domain for Southern Min "is circumscribed by, but not necessarily isomorphic to, surface syntactic structure". Correspondingly, he characterised the SM sandhi domain as, "Mark the right edge of every XP with #, except where XP is an adjunct c-commanding its head" (Chen, 1987, p. 131).

Chen's definition was revised further by Zhang (1992, as cited in Zhang, 1993, p. 302): "Mark the right edge of every XP with #, except where XP is an adjunct m-commanding either its head or the head of XP on the right except infl". Lin (1994, p. 248) defined the prosodic domain for SM tone sandhi in terms of lexical government: " $]_{Xmax}$, X^{max} not lexically governed".

As mentioned, the tone sandhi domain in Southern Min is generally considered as aligned with syntactic phrases (Chiu, 1931; Chen, 1987; Zhang, 1993). The right-hand boundary of a prosodic phrase is assumed to occur at the right-hand boundary of a syntactic phrase XP; however, not every XP is considered being able to mark a prosodic group within which a sandhi process occurs. For example, Chen (1987) considered the adjuncts within a projection of lexical categories, such as N or V, tend not to form separate tonal domains but are subsumed under the larger XP group categorised as NP or VP.

1.6.2. Directionality of SM tone sandhi

Two competing proposals exist with respect to the directionality of Southern Min tone sandhi. First, the direction of tone sandhi is generally viewed as being from the juncture (domain-final) to context (the non-final) position (e.g., Wang, 1967; Cheng, 1968; Chen, 1987; Yip, 1980; Shih, 1986; Zhang; 1993; Yip, 2002; Mortensen, 2002; Hsieh, 2005; Barrie, 2006; Thoms, 2008). In other words, these proposals consider that the citation/juncture form is the basis. The mapping between citation and sandhi forms is thus generally schematised as follows (Chen, 1987, p. 113):

$$T \rightarrow T'/\_T$$

(where T is the citation tone and T' is the sandhi tone).

This sandhi rule indicates that one tone's citation form is the sandhi form of another tone. The whole tone sandhi process of sonorant-ending syllables—also known as *unstopped*, *smooth*, or *free syllables*—are considered to be circular, while the tones of obstruent-ending syllables—also known as *stopped*, or *checked syllables*—form a subsystem of their own, as shown in Figure 1-6 (Chen, 1987, pp. 111-112).

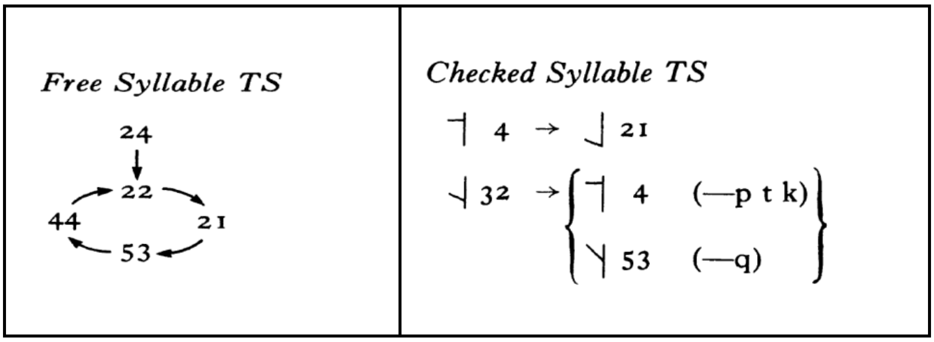


Figure 1-6. Southern Min tone circle in Xiamen (juncture to context).

Alternatively, some scholars (e.g., Hashimoto, 1982; Tsay, 1991; 1994) asserted the direction of alternation is from context form to prejuncture form. In other words, the juncture is considered as the conditioning environment, in which the prejuncture from is derived from the context. The derivation process is formulated as below (Tsay, 1991, p. 83).

$$T_{con} \rightarrow T_{pre}/\_\ \#$$

(where T_{con} for the context tone, T_{pre} for the prejunction tone, # for juncture)

Similarly, the whole tone sandhi process of sonorant-ending syllables is also found forming a circular change (Tsay, 1991, p. 83), as shown in Figure 1-7. Thus, as Tsay & Myers (1996, p. 402) stated, the direction of the Southern Min tone sandhi appears to be indeterminate, and there seems "no compelling reason for choosing one of the alternating tones as underlying and the other as derived" (Tsay & Myers, 1996, p. 402).

The tone alternation is given in (16).

(16) Context tone (basic) --> prejuncture tone (derived)

	T_{con}		T_{pre}
Tone I&II	Ø	-->	H
Tone III	H	-->	HL
Tone IV	HL	-->	L
Tone V	L	-->	Ø

The South Min Tone Circle now appears as in (17).

(17) The South Min Tone Circle

```
Ø -----> H
↑        |
|        v
L <----- HL
```

Figure 1-7. Southern Min tone circle in Taiwan (context to prejuncture).

In this thesis, the citation form specifically refers to the form when the related monosyllabic morpheme is produced in isolation; while the sandhi form refers to the form when the related morpheme is produced in the non-phrase-final context. However, determining which form is basic and which form is derived is difficult, because tonal contrast neutralisation is observed occurring across linguistic contexts, and both forms are considered being independently stored in the mental grammar of native speakers (see Chapter 9).

1.6.3. Theoretical interpretation of SM tone sandhi

Tremendous attempts have been made to account for the nature of tone sandhi in Southern Min (e.g., Wang, 1967; Cheng, 1968; Yip, 1980; Shih, 1986; Mortensen, 2002; Thoms, 2008; Hsieh, 2005; Barrie, 2006; Thoms, 2008, Hsieh, 1976; Schuh, 1978; Wang, 1992; Tsay & Myers, 1996; Chen, 2000; Chen et al., 2010). Nevertheless, it is claimed theoretically challenging to provide a unified, adequate and natural interpretation using the theories of both generative and optimality, indicating the need for other considerations from other perspectives beyond phonology.

1.6.3.1 Rule-based interpretation

In the paradigm of generative phonology, researchers (Wang, 1967; Cheng, 1968; Yip, 1980; Shih, 1986) have endeavoured to formulate the Southern Min sandhi process in terms of rules but have achieved little agreement with respect to how many and what rules are needed. For example, Wang (1967) characterised the tonal alternations occurring among five unstopped tones in Amoy (Xiamen), using a single alpha-switching rule of [αhigh, βfall] → [βhigh, -αfall]. Each alternating process involved only one change in the specification of either the high or falling feature, as shown in Figure 1-8 (Wang, 1967, p. 104).

$$\begin{bmatrix} +\text{HIGH} \\ -\text{FALLING} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{HIGH} \\ -\text{FALLING} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{HIGH} \\ +\text{FALLING} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{HIGH} \\ +\text{FALLING} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{HIGH} \\ -\text{FALLING} \end{bmatrix}$$

Figure 1-8. Feature switching rule for Xiamen tone sandhi in Wang (1967)'s model.

Yip (1980) proposed a set of three separate rules, including register switch, recessive deletion, and dissimilation; however, phonetically similar tones were represented differently at the underlying level in Yip's model. For example, citation form [33] of tone 7 was represented as [+upper; LL], while the sandhi form [33] of tone 1 was represented as [-upper; HH], as shown in Figure 1-9 (Yip, 1980, p. 325). The divergent representations for the same phonetic realisation can create a certain degree of comfusion in the analyses.

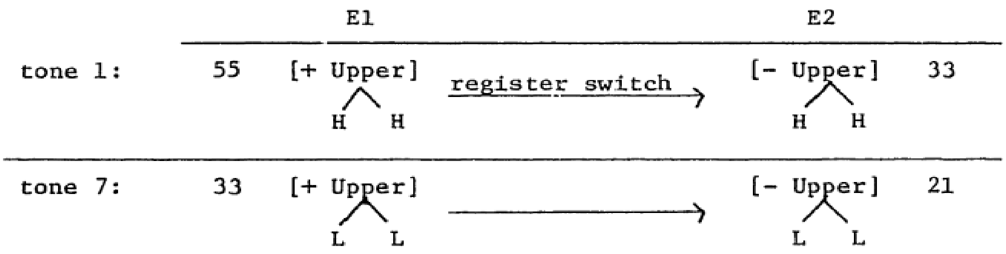


Figure 1-9. Example of similar phonetic pitch but different underlying representations in Yip's (1989) model.

Schih (1986, p. 14) formulated four rules to account for the tone circle in Xiamen, as shown in Figure 1-10, but this model essentially fails to characterise the naturalness of tonal alternation.

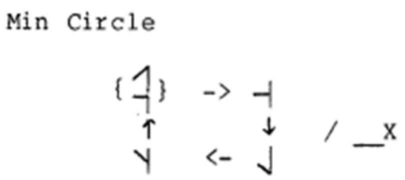


Figure 1-10. Tone sandhi rules in Shih's (1986) model.

Tsay (1991) proposed one principle (contrastive preservation principle) and two contrasts (rule argument constraint and rule exchange constraint) for a unified explanation of Southern Min tone sandhi. Given these requirements, she asserted the tonal alternation rule could be formulated simply as "Insert H" (Tsay, 1991, p. 85). However, these rule-based models appear unsatisfactory to capture the naturalness of phonological processes and environments, as well as the mechanisms motivating tones to alternate in a circular way in Southern Min. For example, Anderson (1978, p. 157) stated,

> [T]he rule defining a tone 'circle' is obviously ingenious, but it is not so obvious that it expresses a linguistically significant generalization.

Alternatively, he considered the sandhi processes as arbitrary substitutions. Chen (2000, p. 44-45) asserted the Xiamen tone circle was an historical accident, because other Southern Min variety, such as Longxi (present-day Zhangzhou), also presented a similar tone circle that could be stated in terms of Middle Chinese tonal categories. Thus, Chen considered the sandhi phenomenon in Southern Min area as "the operative tonal categories hold constant across dialects, but are filled by phonetic values that vary from one dialect to the other".

1.6.3.2 Optimality theory-based interpretation

While the circular tonal alternations have been claimed to be problematic for the rule-based approach, phonologists (e.g., Yip, 2002; Mortensen, 2002; Hsieh, 2005; Barrie, 2006; Thoms, 2008) in recent decades have resorted to the optimality theory (OT) to explore the phonological nature of Southern Min tone sandhi. The classical OT paradigm is viewed as non-computable and inadmissible in formalizing the circular tone change in which A and B are assumed simultaneously to be more marked than the others (Mortensen, 2002; Thoms, 2008). Thus, various extensions of classical OT have been invented, but they differ in constraints and the ways in which constraints are ranked.

Hsieh (2005) proposed a phonetically driven constraint (Max (Tone) longer-syllable) to explain, for example, why a rising tone was prohibited in the sandhi position and an anti-merger tone constraint (*Merger (tone)) to prevent tones from being neutralised during the process. The constraints were ranked hierarchically in a strict order as follows (Hsieh, 2005, p.1 09):

Max (Tone) longer-syllable (*Rise) >> *Merger (L) >> *Merger (H) >>
*Merge (Tone) >> Ident-OO-(H-Register)

Nevertheless, Hsieh's analysis does not account theoretically for some facts, for example, why tones must be alternated to other forms, why higher register tones are not preferred in non-prominent contexts, and why tones alternate in a circular fashion.

Barrie (2006) asserted the circular tonal alternation was a phonological activity rather than a morphological process (Hsieh, 1970, 1976; Wang, 1992; Tsay & Myers, 1996) or a paradigmatic replacement (Schuh, 1978). He claimed the tonal circle could be characterised by expanding constraints of preservation contrasts, as proposed by Lubowicz (2003, as cited in Barrie, 2006). The constraints that Barrie proposed can be ranked hierarchically in a strict order as follows (Barrie, 2006, p. 138):

*Rise >> *PCout(Register) >> *PCout(Pitch) >> *Contour >> *High.

Nevertheless, these constraints account for only a small number of scenarios of the tonal alternations in Xiamen. The testability and predictability of his model needs to be verified further.

Thoms (2008) proposed six different constraints to explain the mechanisms underlying the Xiamen tone circle. The constraints were proposed based on two fundamental assumptions: (1) Tonal structures were preserved during the process, and (2) the evaluation of candidates was systemic. These constraints were ordered strictly as follows (Thoms, 2008, p. 427):

DIFFER, *RISE >> *MERGE >> DRAT >> Ident(Pitch) >> Ident(Shape)

Thoms's model seems can predict and explain all attested chain shifts in Xiamen; however, whether this model can make cross-linguistic predictions for languages with similar circular tone changes must be tested further.

In summary, OT grammarians have endeavoured to construct sets of constraints to argue for the phonological reality of circular tone shifts. Various extensions of classical OT have been invented, but they also encountered theoretical difficulties in providing an adequate and unified explanation, as in the rule-based approaches.

1.6.3.3 Lexicon-based interpretation

While it appears phonologically challenging to offer a natural and unified explanation for the motivation of Southern Min tone sandhi, a group of researchers conducted experiments to test the psychological reality of SM tone sandhi rule (e.g., Hsieh; 1976; Wang, 1992; Tsay & Myers, 1996; Zhang et al., 2006; Chen et al., 2010), the results of which largely falsify the full productivity of tone sandhi rule in Taiwan, and favour a lexicon-based interpretation for the relation between sandhi and citation tones.

Hsieh (1976) asserted that both sandhi and non-sandhi forms are listed in the lexicon and the choice between the two alternatives is optional, because the speaker failed to demonstrate any knowledge of her bi-dialectal tone sandhi rules. However, the explanatory power of his assertion is thought to be limited, because Hsieh's research corpus only involved one bi-dialectal female speaker of coastal and inland Taiwanese, and only focused on two tones of low-rising [13] and high-falling [53] s at the underlying level

Wang (1992) also claimed that both isolation and sandhi tones are listed in the lexicon based on his long-term experiment for seven tones of 22 native speakers of Hsin-chu Taiwanese. Contrary to Hsieh (1976)'s finding, Wang believed that "some sandhi rules are real to a certain degree, although perhaps not to the degree of automatically generative as assumed in SPE" (p. 125), to which he assumed that the sandhi rules are organisational rather than derivational. However, the difference between Wang's lexicon-based proposal and previously rule-based models appears to be how generative the rules are, which may to a certain degree undermine its plausibility.

Tsay & Myers (1996) also proposed an allomorph selection hypothesis for Taiwanese tone sandhi. They asserted that morphemes are stored in the lexicon with an abstract diacritic indicating tonal category, and no lexical rules are applied to derive the allomorphs, as shown in Figure 1-11 (p. 403). Nevertheless, the claim they made is largely not built on systematic empirical investigations on their own, but rather by summarising the results of proposed studies from multiple sources, which calls into question the plausibility of the model, and calls for substantial empirical studies to test the hypothesis.

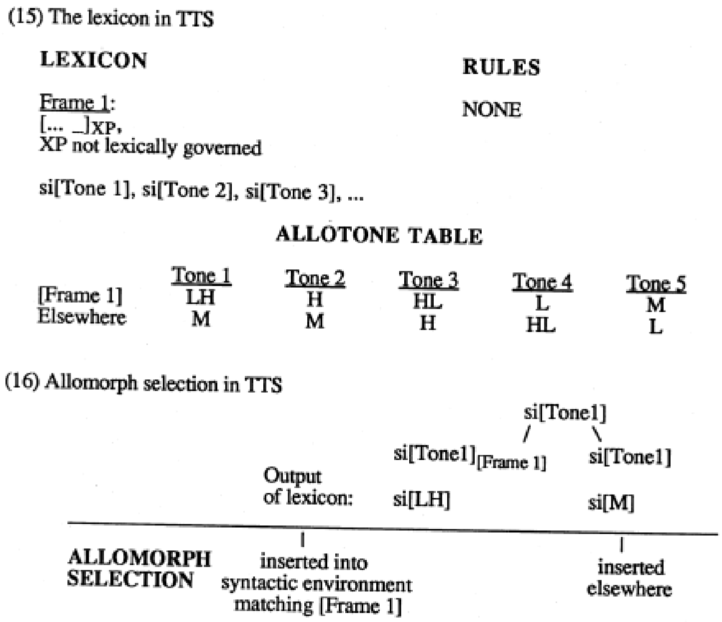


Figure 1-11. Allomorph selection in Taiwanese tone sandhi in the model of Tsay & Myers (1996).

Chen et al. (2010) asserted that Taiwanese tone sandhi is not grammar-governed but rather an allomorph selection on the basis of the results of their production experiment on five non-checked from two groups of 12 Taiwanese speakers. Results also showed that the performance of speakers' tone sandhi production is significantly affected by age group, allomorph frequency, positions of the morpheme besides the lexical experience. Similarly, Zhang et al. (2006) asserted that the Taiwanese tone sandhi productivity is considerably affected by phonological opacity, durational property and lexical frequency of the sandhi. Thus, examining the productivity of Southern Min tone sandhi must take into account various factors for a plausible and adequate interpretation.

As discussed, linguists have devoted themselves to exploring the nature of Southern Min tone sandhi. Either rule-based or OT-based approaches have been asserted failing to provide a unified and adequate account to capture and interpret the naturalness and mechanism of tone sandhi in Southern Min. The unsatisfactory efforts have led linguists to reconsider the fundamental question of whether tones change in a circle and to look for other analytical solutions beyond phonology. For example, a considerable number of production experiments have been conducted to question the psychological reality of tone sandhi rule in Taiwanese, giving rise to a lexicon-based interpretation of allomorph selection.

This thesis asserts a morphological nature for Zhangzhou tones, and a morphophonemic relation between sandhi and citation tones through mapping the relatedness of citation, phrase-initial and phrase-final tones in accordance to their multidimensional realisations in phonetics (see Chapter 9). It is hoped that it will improve our understanding of how various forms of tonal realisations in Southern Min are structured and related in the mental grammar of native speakers.

1.7. Studies of Zhangzhou

Zhangzhou belongs to Southern Min of the Sinitic language family. This dialect has received a considerable amount of linguistic attention since the early 19th century, and studies can be categorised into four main types:

- Documentation of homophone dictionary of Zhangzhou (Xie, 1818; Medhurst, 1838; Dyer, 1838; Schlegel, 1886; Dong, 1959; Lin, 1992; Ma, 1993; ZCCEC, 1999; Gao, 1999; Zhou, 2006; among others).
- Diachronic comparisons between the sound system of present-day Zhangzhou and that of Archaic Chinese (Huang, 1999), that of "Guangxun" of Middle Chinese (Chen, 1990; Ma, 1997a, 1997b, 1999; Yang, 1999), and that of Zhangzhou in the 19th century (Chen, 1990; Guo, 1999; Du, 2013; Huang, 2014).
- Synchronic comparisons between the sound system of Zhangzhou and other varieties of Southern Min, for example, Xiamen, Quanzhou, Taiwan, Chaozhou, Longyan, and so on (Chen, 1998; Li &Yao, 2008; Ma, 2002, 2008; Liang, 2014; Wu, 2012; Ma, 2013; Huang, 2014; Ding, 2016; among others).
- Synchronic description of the sound system of Zhangzhou, including vowels (Huang, 2009; 2010) and citation tones (Yang, 2008; 2014; Yin, 2009; Huang et al., 2016).

These descriptive materials—lists of segmental and tonal inventories, as well as the rhyme table—indicate many interesting details with respect to the sound systems of Zhangzhou and local vocabularies; however, there are significant limits to the insights they can offer, both empirically and theoretically. No studies have addressed the basic issues with respect to syllables and segments (see Section 3.6), for example, what are the individual syllable components realised phonetically, and how can they be treated phonologically? How do segmental and suprasegmental parameters interact to shape tonal distinctions?

Several studies have provided a list of pitch contrasts for Zhangzhou citation tones, but the transcriptions differ not only from each other but also largely from the result of this study (see Section 4.3). Four pieces of acoustic work of citation tones are seen in the literature, but they all present problematic aspects in terms of research design and analysis (see Section 5.4). No empirical and theoretical analysis has been conducted on Zhangzhou tone sandhi (see Sections 6.4, 7.3, 8.3, and 9.4). Therefore, Zhangzhou can be considered as a hitherto under-described variety with respect to both segmental and tonal studies.

In this thesis, each body chapter (3-8) will contain a separate section to review pertinent previous studies of Zhangzhou with respect to specific research issues being addressed in that chapter. Constrained by the under-descriptions of previous studies, the review section will be consistently conducted after the presentation of description and discussion of the findings of this study. This arrangement ensures the original results of this study are better discussed and presented, and also ensures the research gaps and the contributions that the study made are better identified by comparisons with previous work.

1.8. Structure of the Thesis

This thesis is divided into four parts. Part A introduces the preliminaries before formal descriptions and discussions of Zhangzhou tones.

- Chapter 1 has introduced the research background of this study, including tonology, tonetics, Sinitic tone sandhi, Southern Min tone sandhi, and previous studies of Zhangzhou. It also provided an introduction to the research goals and research opportunities of this study.

- Chapter 2 describes the multi-methodological approaches used in this study. It largely addresses two issues: (1) How the field data were elicited in the field site of Zhangzhou, and (2) how the post-field data were processed auditorily, acoustically, and statistically.

- Chapter 3 discusses Zhangzhou syllables and segments. It addresses three primary issues. (1) How are syllables structured in Zhangzhou? (2) What are the syllable components? (3) What are the individual syllable components realised phonetically? and how are they considered phonologically?

Part B addresses the realisations of Zhangzhou citation tones.

- Chapter 4 discusses the auditory properties of Zhangzhou citation tones by addressing two issues. (1) What are the eight tones realised with respect to pitch in monosyllabic settings? (2) How do segmental and suprasegmental parameters interact to make tonal distinctions?

- Chapter 5 addresses the acoustic properties of Zhangzhou citation tones. It applies the pairwise *t* tests and the hierarchical clustering algorithms to determine how many normalised F0 and duration levels are contrastive among Zhangzhou citation tones. It uses the SplitsTree software to address how Zhangzhou citation tones are related to each other from the phonological perspective.

Part C includes discussions of Zhangzhou disyllabic tones.

- Chapter 6 describes the auditory properties of Zhangzhou disyllabic tones by exploring how individual tones interact in disyllabic constructions and how various parameters contribute to make tonal contrasts. It also introduces the general properties of Zhangzhou tone sandhi, for example, the dominancy and specification of Zhangzhou tone sandhi domain.

- Chapter 7 explores the acoustic properties of Zhangzhou phrase-initial tones, supplemented with statistical testing and articulatory explanation. It examines how the phrase-initial tones are realised acoustically with respect to multidimensional parameters, and whether the realiations may be affected by following tones.

- Chapter 8 discusses the acoustic properties of Zhangzhou phrase-final tones, supplemented with statistical testing and articulatory explanation. It addresses how the phrase-final tones are realised acoustically in terms of various parameters and examines whether their realisations may be affected by preceding tones.

Part D addresses the mapping between Zhangzhou citation and disyllabic tones.

- Chapter 9 explains how the disyllabic tones are mapped onto their citation tones, and how native speakers structure the different mappings in their mental grammars.

Part E (Chapter 10) summarises the findings emerged from the analysis of the data from 21 speakers, and discusses directions for future studies of Zhangzhou tones.

Chapter 2: Methodologies

This study draws upon interacting multi-methodological approaches—linguistic fieldwork; auditory and acoustic phonetics; and statistical testing—to explore the nature of Zhangzhou tones. Methodologically, it collects primary data from interactions with native-speaking informants in the field via a carefully controlled elicitation procedure. It adopts the pitch notation system of Chao (1930) and IPA to present transparently detailed auditory properties of Zhangzhou tones and segments. It uses a standardised quantification and normalisation procedure to extract tonally relevant F0 and duration in acoustic signals. It applies the technique of pairwise *t*-test comparisons to determine the statistical relation among different normalised F0 and duration levels. It also uses the software of SplitsTree to generate the phylogenetic networks for the mapping of the relatedness among different categories of tone. It combines articulatory and acoustic phonetics to explain how the acoustic manifestations of parameters change with respect to the adjustment of articulatory configurations for different tonal productions.

This chapter largely addresses how the field data were elicited in Zhangzhou, as well as how the post-field data were auditorily, acoustically, and statistically processed. The research protocol was reviewed and approved by the Australian National University Human Research Ethics Committee as proposal 2015/03 prior to the field trip. No problematic ethical issues arose during the research.

2.1. Pre-data Elicitation

The selection of research area and participants is purposeful and predominantly constrained by the research design of a given linguistic study (Samarin, 1967; Ladefoged, 1999; Dörnyei, 2007; Bowern, 2008; Chelliah & de Reuse, 2010). This section primarily describes how the research area and informants were determined and selected in the field site of Zhangzhou, during April-May 2015, to address the research questions stated in Chapter 1.

2.1.1. Zhangzhou city

Zhangzhou 漳州, romanised differently as *Chiang Chiu* or *Changchow*, is a prefecture-level city situated in the southern Fujian 福建 province of Mainland China. The latitude and longitude coordinates for Zhangzhou city are 24.5130° N, 117.6471° E. Zhangzhou faces the Taiwan Strait to its east but borders Fujian cities of Xiamen 厦门, Quanzhou 泉州, and Longyan 龙岩 on its east, northwest, and west, respectively, while its southwest region borders the Chaozhou 潮州 city of Guangzhou 广东 province, as indicated in Figure 2-1.

Figure 2-1. The location of Zhangzhou in China (Dedering, 2010).

Zhangzhou covers an area of approximately 12,600 square kilometers with a registered population of 5.05 millions in 2016, according to the official Zhangzhou government website (http://www.zhangzhou.gov.cn/cms/html/zzsrmzf/2017-02-17/1119782440.html). The whole city currently comprises eleven administrative regions, including two urban districts (Xiangcheng 芗城 and Longwen 龙文), one county-level city (Longhai 龙海), and eight counties (Changtai 长泰, Hua'an 华安, Nanjing 南靖, Pinghe 平和, Yunxiao 云霄, Zhao'an 诏安, Dongshan 东山, and Zhangpu 漳浦), as shown in Figure 2-2.

The colloquial language spoken by native Zhangzhou people is predominantly Southern Min (better known as Hokkien), which is referred to as Zhangzhou in this thesis. Zhangzhou speech is mutually intelligible with other Southern Min varieties in Fujian (e.g., Xiamen and Quanzhou), which are collectively considered as the mainstream division of Southern Min (known as Quan-Zhang division 泉漳片). The speech is also accessible in the local television programming, radio programming, popular songs, and traditional opera (known as Xiangju 芗剧). It has a certain degree of mutual intelligibility with other Southern Min divisions in Mainland China (e.g., Teochew and Leizhou); however, it is completely mutually unintelligible with other Chinese dialects (e.g., Mandarin, Hakka, Cantonese, Wu, Xiang, and Gan).

Mandarin, as the national language of China, is a compulsory subject in both primary and secondary schools of Zhangzhou, and is used as a medium of instruction in the university and public occasions. Hakka is also found but only spoken by a relatively small number of people living in mountainous areas of western Zhangzhou, for example, Hua'an, Nanjing, Pinghe, and Zhao'an counties, bordering a major Hakka-speaking city of Longyan.

As a consequence of geographic advantages and long-standing maritime trade, the speech of Zhangzhou, together with other varieties of Southern Min (Hokkien), has spread to many regions in Southeast Asia since the early 12th century and Southern Min has historically served as the lingua franca among overseas Chinese communities, such as in Taiwan, Singapore, Malaysia, Indonesia, and many other areas (Ma, 1994). Figure 2-3 shows the distribution of Southern Min (Hokkien) in present-day Fujian and Taiwan, in which the Zhangzhou variety (Chiang Chiu) are shown in blue and purple (Liaon98, 2016). Therefore, the local dialect of Zhangzhou must be clarified when conducting a rigorous linguistic study of Zhangzhou.

Figure 2-2. The eleven administrative regions in Zhangzhou (Chk2011, 2012).

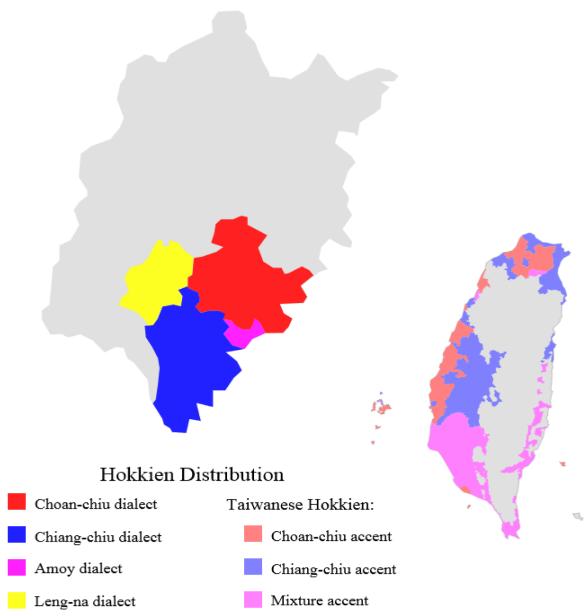


Figure 2-3. The distribution of Southern Min (Hokkien) in Fujian and Taiwan (Liaon98, 2016).

2.1.2. Research area selection

The research area investigated in this project is the urban districts of Xiangcheng and Longwen in Zhangzhou city of the Fujian province in Mainland China. Four major reasons are given for this research area selection.

First, using this area is research design driven. This project primarily investigates what tone, as a linguistic category, is realised and how tone behaves in multisyllables in the speech of one single site of Zhangzhou city, rather than conducting a typological or socio-phonetic survey of how tone in the Zhangzhou variety is realised and interacts within and outside Zhangzhou city.

Second, using this research area helps in minimising effects of regional variation. The present-day Zhangzhou city administers eleven regions in Fujian; however, regional variation has often been reported with respect to the sound system (e.g., Ma, 1994; Yang, 2008; Guo, 2014). For example, Yang (2008, p. 36) summarised pitch realisations across Zhangzhou administrative regions as shown in Table 2-1.

Table 2-1. Comparisons of tonal pitch realisation in different administrative regions of Zhangzhou

Region	Tone 1	Tone 2	Tone 3	Tone 4	Tone 5	Tone 6	Tone 7
Xiangcheng	34	13	53	21	22	32	121
Longwen	34	13	53	21	22	32	121
Longhai	34	13	53	21	22	32	4
Zhangpu	44	312	53	22	33	32	121
Yunxiao	44	21	53	22	33	42	121
Dongshan	44	213	51	22	33	32	13
Zhao'an	44	312	53	21	33	32	13
Hua'an	44	213	453	21	33	32	131
Changtai	44	214	53	21	31	32	33
Pinghe	34	212	53	21	22	42	121
Nanjing	34	212	53	21	33	22	121

As Table 2-1 indicates, regionally-correlated differences vary to some degree in pitch values, but Xiangcheng and Longwen present consistently-similar pitch for each tone. Therefore, recognising the regional variation is important before the area for data collection is confirmed.

Third, this research area allows the present research results to align with previous studies. Most studies of Zhangzhou since the 19th century have been based on the language data of the urban districts, predominantly Xiangcheng (e.g., Dong, 1959; Lin, 1992; Ma, 1994; FCCEC, 1998; ZCCEC, 1999; Gao, 1999; Zhou, 2006; Chen, 2007; Yang, 2008; Guo, 2014). Maintaining consistency with previous research concerning the research area is important to improve understanding of Zhangzhou and to make this study comparable with published work, both synchronically and diachronically.

The fourth reason for selecting this research area is self-motivation. I was born and reared in the Longwen district of Zhangzhou. My personal interest and motivation lie in investigating my mother tongue. Also, it is convenient for me to recruit qualified informants and collect data of high quality for further processing and analysis.

In light of these four factors, the research area was narrowed to the urban districts of Xiangcheng and Longwen of Zhangzhou. Nevertheless, as reported by my informants and also in the literature (Yang, 2008; Guo, 2014), some towns have slight differences in both accent and vocabulary from the mainstream, for example, Tianbao 天宝, Punan 浦南, and Shiting 石亭 in Xiangcheng, and the villages near Yundong Mountain 云洞岩 and Jiangdong Bridge 江东桥 in Longwen. Therefore, the research area was further restricted to the most inner urban communities, including Buwen 步文, Lantian 蓝田, Xiangkou 巷口, and Xiqiao 西桥, and excluding the problematic places mentioned previously.

2.1.3. Research participant selection

Not every speaker is suitable to be a language informant (Samarin, 1967, p. 25). It is important to select those informants who are reasonably representative of the speech community being studied and who are thoroughly experienced in using their mother tongue (Samarin, 1967, p. 25). This subsection discusses three issues concerning the participants for the data elicitation in Zhangzhou: (1) How are informants selected? (2) How are informants recruited? (3) How many speakers are needed?

2.1.3.1 How are informants selected?

A set of criteria was developed on the basis of suggestions by some linguists (Samarin, 1967; Ladefoged, 1999; Chelliah & de Reuse, 2010; Rasinger, 2013) for the informant selection in the field site for this study. These criteria range from a speaker's intellectual curiosity to his or her physical condition.

(1) Intellectual curiosity. Speakers should exhibit a positive attitude towards their native language and be enthusiastic about participating in the data collection. They should be confident of their utterances and should be able to talk freely and naturally on a wide range of subjects relevant to Zhangzhou culture.

(2) Physical condition. Because this is phonetic fieldwork, speakers are required to have all their own teeth and good hearing. They should not have speech impediments, such as a lisp or a stutter. Their receptors (eyes, ears, and brain) and effectors (brain, tongue, and mouth) should all operate normally and healthily (Chelliah & de Reuse, 2010).

(3) Linguistic environment. Zhangzhou is a bilingual city in which Southern Min is predominantly used as a vernacular dialect and Mandarin is used in official situations.

Informants must have been raised and must currently live in the inner city of Zhangzhou. Further, their parents and spouses are preferred to be native speakers. This criterion ensures informants have spoken this dialect since their childhood and use it in their daily lives so that the elicited data are thoroughly representative of Zhangzhou.

(4) Age. Speakers should be middle-aged. A young age is excluded because Mandarin seems more prestigious than Zhangzhou to younger generations, and they tend to have incomplete knowledge of Zhangzhou speech. Speakers of an advanced age, on the other hand, are more likely to have some physical defects, such as deafness, poor healthy, lenis articulation, insufficient teeth, and other problems.

(5) Education. Speakers must be literate and must have received secondary education or above. This criterion ensures they are able to recognise and read the corpus in Chinese characters and are able to provide extended responses and explanations in Mandarin in case they have difficulty expressing themselves in Zhangzhou.

2.1.3.2 How are informants recruited?

The recruitment of informants was accomplished with recommendations from the local community and my parents. Before the recording was officially conducted, potential informants were invited to my home or to the park for a relaxing and friendly gathering. This event proved to be a positive and inspiring interaction between the potential informants and myself as the investigator.

(1) The local community learned about this field trip, and most of them showed positive interest and curiosity in the data collection. They contributed a wealth of input to the design of the corpus and expressed pride in their contributions to documenting their own culture and language.

(2) I had the chance to meet different people and evaluated their strengths with respect to whether they would be good informants for the acoustic experimental purpose or purely for the vocabulary documentation purpose.

(3) I was given the chance to hear the language from different speakers and reduced the chances of biasing the description of the language towards the idiolect of a single speaker with whom I had been working before this field trip.

With the assistance of my parents and the local community, I had access to nearly 50 people, but only slightly more than half of them were recorded: 21 for the controlled corpus and 6 for the vocabulary documentation. Some people did not participate in the recording largely because they did not meet the selection criteria or because of other reasons, for example, catching a cold before the day of recording.

2.1.3.3 How many speakers are needed?

Another important issue with respect to the informant selection is how many speakers are needed for a quantitative phonetic study to achieve an accurate representation of a particular variety of speech. This question does not seem to have received much attention in the literature. For example, in his book *Phonetic Data Analysis*, Ladefoged (2003, p. 14) wrote:

> Ideally you want about half a dozen speakers of each sex. . . . If you can eventually find 12 or even twenty members of each sex, so much the better.

It is a little difficult to see how, if half a dozen speakers is ideal, then more speakers are better. This is perhaps why Ladefoged's figures were later quoted in the *Handbook of Descriptive Linguistic Fieldwork* (Chelliah & de Reuse, 2010, p. 254): ". . . from 24 to 40 speakers". An ideal database of 30 speakers was recommended by Zhu (2004) from a statistical perspective, perhaps because 30 is about the number where a *t* distribution becomes normal. Huang et al. (2016) used a preliminary Monte Carlo investigation of how the normalisation index changes with the number of speakers used to normalise.

This study was originally intended to recruit 16 speakers, selecting 12 of them with the best sound quality for the quantitative analysis. This number satisfied the ideal number of 12 proposed by Ladefoged (1999). Nevertheless, with the assistance of the local community, nearly 50 people came for the pre-recording interviews, and 27 speakers were ultimately recorded. Twenty-one speakers, including 9 males and 12 females, took part in the recording of the controlled stimuli for the phonetic experiments. Another group of 6 speakers, including 4 males and 2 females, participated in the local vocabulary documentation of 1366 lexical items. All speakers met the selection criteria and were enthusiastic about contributing to the data collection. The average age of male speakers was 56.6 and that of female speakers was 50.

The demographic information of all speakers appears in Tables 2-2 and 2-3, with seven independent variables concerning their sociocultural backgrounds: (1) sex, (2) age, (3) current residence, (4) birthplace, (5) education, (6) occupation, and (7) other language acquisition.

2.2. Data Elicitation

Data elicitation essentially reflects an intimate human relationship established between participants and the investigator (Samarin, 1967) and is predominantly constrained by the purpose of the research project and the techniques of the data collection and processing. The elicitation will be successful if it is properly planned, involving an appropriate corpus design and selection of elicitation equipment and environment, as well as a careful control of the elicitation procedure. This section describes how the procedure of elicitation was conducted in the field site of Zhangzhou.

Table 2-2. The personal profile of 21 speakers for controlled-data collection

No.	Name	Sex	Age	Current residence	Birthplace	Education	Occupation	Other language
1	LRZ	F	62	Longwen	Longwen	Secondary	Government	Mandarin
2	LYY	M	68	Xiangcheng	Longwen	B.A.	Education	Mandarin
3	LMY	M	44	Longwen	Xiangcheng	College	Business	Mandarin
4	YML	F	62	Xiangcheng	Xiangcheng	Secondary	Education	Mandarin
5	CSM	F	51	Longwen	Longwen	Secondary	Business	Mandarin
6	YBR	M	53	Xiangcheng	Longwen	B.A.	Government	Mandarin
7	HZB	M	42	Longwen	Longwen	Secondary	Government	Mandarin
8	WYF	M	59	Longwen	Longwen	College	Company	Mandarin
9	HTL	M	56	Xiangcheng	Xiangcheng	Secondary	Company	Mandarin
10	HYS	M	57	Longwen	Longwen	College	Business	Mandarin
11	WJG	M	56	Xiangcheng	Xiangcheng	College	Education	Mandarin
12	HHX	F	37	Longwen	Longwen	College	Education	Mandarin
13	HMC	M	49	Longwen	Longwen	Secondary	Company	Mandarin
14	LRX	F	55	Xiangcheng	Xiangcheng	Secondary	Company	Mandarin
15	ZBQ	F	40	Longwen	Nanjing	Secondary	Government	Mandarin
16	HJH	F	40	Longwen	Longwen	Secondary	Business	Mandarin
17	LZP	F	58	Longwen	Longwen	Secondary	Education	Mandarin
18	LCH	F	53	Xiangcheng	Longwen	Secondary	Business	Mandarin
19	HQM	F	36	Longwen	Longwen	College	Business	Mandarin
20	HYH	F	34	Longwen	Longwen	B.A.	Education	Mandarin
21	HSX	F	56	Longwen	Longwen	Secondary	Education	Mandarin

Table 2-3. The personal profile of six speakers for local vocabulary documentation

No.	Name	Sex	Age	Current residence	Birthplace	Education	Occupation	Other language
1	HRH	M	67	Longwen	Longwen	Secondary	Business	Mandarin
2	HMC	M	49	Longwen	Longwen	Secondary	Company	Mandarin
3	LYY	M	68	Xiangcheng	Longwen	B.A.	Education	Mandarin
4	YHR	M	68	Longwen	Longwen	Secondary	Education	Mandarin
5	YML	F	62	Xiangcheng	Xiangcheng	Secondary	Education	Mandarin
6	LZP	F	58	Longwen	Longwen	Secondary	Education	Mandarin

2.2.1. Corpus design

The preparation of a good data corpus is the key to linguistic phonetic fieldwork (Ladefoged, 1999). A good corpus is a body of data within which patterns, systems, or generalisations can be discovered, resulting in hypotheses proposed by induction (Samarin, 1967). In this study, two separate corpora resulted from the data collection: a controlled one and a general one.

(1) For the acoustic investigation, a controlled corpus was carefully designed under the collaboration between the local community and myself as a native-speaking linguist candidate. This controlled corpus incorporated two sets of word lists: one for investigating Zhangzhou citation tones and the other for investigating the multisyllabic tones. A supplementary word list was also designed for the segmental investigation. The principle of the controlled corpus design is described in the rest of this sub-section.

(2) For vocabulary documentation, a word list of 1366 words, categorised according to different semantic domains, was prepared; however, this corpus was not used for experimental phonetic purposes in this thesis. The list can be referred to in Appendix C (attached USB).

2.2.1.1. A word list for citation tone elicitation

The preliminary investigation identified seven monosyllabic citation tones in Zhangzhou (see Huang et al., 2016). In the word list for citation tone elicitation, there were usually 20 tokens for each tone, resulting in approximately 160 tokens to be recorded by 21 speakers in the field and to be processed acoustically and statistically in the lab. However, with the progress of the research in the post-field stage, eight tones rather than seven were posited in this study (see Chapter 4). This word list of 160 tokens was, thus, effectively shared by eight tones, with some tones (e.g., tones 7 and 8) having less than 20 tokens to be processed. While it appears impractical to return back to the field site and collect an equal number of tokens for tones 7 and 8 from the same speakers, the existing corpus from 21 speakers can to a large degree compensate the shortage and ensure the derived results represent Zhangzhou variety as a whole.

The word list included as many minimal or sub-minimal pairs as phonotactics allowed. Tokens were chosen across different syllable types and contained comparable numbers of syllable onsets with different manners and places of articulation as well as vowels of different height and backness, which maximally balanced the intrinsic perturbation effects on the realisation of tonal F0 from tautosyllabic segments. Table 2-4 shows examples from this word list, but more examples for each tone appear in Table A1 in Appendix A.

2.2.1.2. A word list for disyllabic tone elicitation

Based on the preliminary assumption of seven tones in Zhangzhou, logically, there would be 49 (= 7 * 7) disyllabic combinations. Twelve examples were thus selected for each combination, and 588 (= 12 * 49) tokens made up the word list to be elicited in the field. As stated above, eight tones rather than seven were ascertained in this thesis; therefore, these 588 tokens were effectively spread across 64 (= 8 * 8) tonal combinations, leading some combinations to have no tokens or less than 12 tokens for processing on the basis of this word list. Nevertheless, additional tokens were chosen from the supplemented word list, as introduced below, to make sure every combination had tokens to be processed auditorily, acoustically, and statistically.

The tokens for disyllabic tone elicitation were also designed under the principle of maximally balancing the tautosyllabic segments' intrinsic perturbation on the tonal F0 realisation. Another special consideration with respect to this corpus design was that the selected constructions largely contained one morpheme having a tone that can both precede and follow any tone in the inventory. For example, in Table 2-5, these disyllabic phrases all contain one common morpheme (/ɓɛ3/ 'horse') that can precede (A) and follow (B) any of the eight citation tones. Designing in this way can largely enable investigation of how the tone of a given morpheme is realised across different contexts and how surrounding tones influence its realisation. Additionally, the tokens are all disyllabic phrases, mainly incorporating noun phrases, verb phrases, adverbial phrases, and adjective phrases, which are commonly used in the daily life of the local community rather than being arbitrarily created for experimental purposes. All examples for disyllabic tone elicitation appear in Appendix A (Table A2), which has been reorganised in accordance to the tonal pattern.

Table 2-4. Examples of monosyllabic tokens for citation tone elicitation

Tone	Example 1	Example 2	Example 3	Example 4
1	/si/ 'poetry'	/kɔ/ 'mushroom'	/tɐŋ/ 'east'	/ʔi/ 'he/she'
2	/si/ 'time'	/kɔ/ 'glue'	/tɐŋ/ 'copper'	/ʔi/ 'move'
3	/si/ 'die'	/kɔ/ 'drum'	/tɐŋ/ 'wait'	/ʔi/ 'chair'
4	/si/ 'four'	/kɔ/ 'look after'	/tɐŋ/ 'chilly'	/ʔi/ 'intenson'
5	/si/ 'affirmative'	/ħɔ/ 'rain'	/tɐŋ/ 'heavy'	/ʔi/ 'play'
6	/sik/ 'colour'	/kɔk/ 'country'	/tɐk/ 'kick'	/ʔit/ 'one'
7	/sit/ 'solid'	/tɔk/ 'poison'	/tɐt/ 'stimulating (food)'	/ʔik/ 'bathe'
8	/tsi/ 'tongue'	/kɔ̃/ 'snore'	/tɐ/ 'step on'	/zi/ 'press'

Table 2-5. Examples of disyllabic tokens for disyllabic tone elicitation

Pattern A	Example	Pattern B	Example
3+1	/ɓɛ.kʰɐ/ 'horse feet'	1+3	/ʔɔ.ɓɛ/ 'black horse'
3+2	/ɓɛ.tɐ/ 'horse hoof'	2+3	/tɐŋ.ɓɛ/ 'copper horse'
3+3	/ɓɛ.sɐj/ 'horse shit'	3+3	/pɵ.ɓɛ/ 'precious horse, BMW'
3+4	/ɓɛ.tsʰwi/ 'horse mouth'	4+3	/tsjɐn.ɓɛ/ 'battle horse'
3+5	/ɓɛ.pɔ/ 'horse stance'	5+3	/tsʰi.ɓɛ/ 'to feed horse'
3+6	/ɓɛ.kut/ 'horse bone'	6+3	/pɐt.ɓɛ/ 'Eight Horse'
3+7	/ɓɛ.sut/ 'horsemanship'	7+3	/pɐt.ɓɛ/ 'to bind horse'
3+8	/ɓɛ.tsi/ 'horse tongue'	8+3	/pɛ.ɓɛ/ 'white horse'

2.2.1.3. A supplemented word list for multisyllabic tone elicitation

In addition to the two well-controlled word lists for citation and disyllabic tone elicitation, a supplementary word list was also prepared to further illustrate segmental contrasts and their allophonic distributions, as well as some multisyllabic data showing specific sandhi patterns from the mainstream. These word lists ensured obtaining sufficient language data for the descriptive and explanatory study of the sound system of Zhangzhou. Examples will be shown in relevant sections in this thesis.

2.2.2. Recording

Because this study is predominantly phonetic, acoustic in particular, it is important to regard the way that data were elicited, including the selection of appropriate recording environment and equipment, as well as control of the recording procedure.

2.2.2.1. Recording place

Physical environment is important in the recording of speech sound for phonetic analysis (Samarin, 1967; Ladefoged, 1999). Having the very supportive assistance of the local community, especially my parents and relatives, I was able to investigate several places for recording, including broadcasting and meteorological stations. Unfortunately, most of these places were not acoustically desirable for a variety of reasons, for example, external noise from upper floors and/or from surrounding building construction, dominant echoes and/or noisy machines inside the rooms, or time conflict between my recording schedule, the availability of informants, and the working time provided by the stations.

In the end, the Zhangzhou Hotel was selected as the recording place for several reasons.

(1) This hotel is surrounded by tall, leafy trees that, to a large extent, dampen external noise from vehicles and building construction.

(2) The recording room was in a very quiet corner and was acoustically absorbent. The carpet and the sound-proof windows and door largely minimised background noise and echo inside the room.

(3) This hotel is located in the innermost part of the city, making it convenient for my informants. The comfortable environment made them feel relaxed and happy to provide their utterances in their best voice, even though most of them had never been recorded in phonetic fieldwork before. The excellent psychological and voice status of informants largely guaranteed the language data quality for processing and analysis.

(4) I was able to manage the recording schedule flexibly within the pre-arranged period of field work without worrying about the constraints of the working place or the time needed to balance in the official recording stations.

(5) I was able to focus entirely on data recording without being interrupted, which is normally an issue for an investigator doing research in his or her hometown.

2.2.2.2. Recording equipment

Even if the corpus for data elicitation is carefully designed, the informants are well selected, and the recording environment is excellently sound-proof, the recordings will be bad if the recording equipment has a low quality (Samarin, 1967; Ladefoged, 1999; Bowern, 2008). Several microphones, kindly provided by the local broadcasting stations and universities, were tested and compared, but not every recorder could produce satisfactory auditory and acoustic results. One professional but expensive cardioid condenser microphone, Blue Baby Bottle SL, provided by Huaqiao University, was adopted for the recording.

This recorder features a new switchable 100Hz high-pass filter and -20 dB response, which can accurately capture the the human voice, regardless of vocal type, instruments, or environments. It also features a single-membrane large-diaphragm capsule mounted in a “lollipop” style enclosure for enhanced sensitivity. Its cardioid polar pattern is effective in minimising noise and ambiance at the off-axis sections of the capsule.

Another important aspect concerning the value of this microphone is that it worked comparably well with my MacBook Air computer and the sound recorder in Praat, version 5.3. The sounds recorded and reproduced via this microphone had very high quality with clear waveforms and spectrograms shown in Praat. During the recording, this microphone was placed about 8 inches from the speaker’s mouth and about 4-6 inches down, as suggested by Samarin (1967, p.100), to avoid recording strong aspirations such as puffs or pops.

2.2.2.3. Recording procedure

As Samarin (1967, p.99) stated, “The highest quality of recording equipment is no substitute for poor recording procedure. Any one of a number of details can ruin what could have been an extremely good and valuable recording”. The recording procedure of this phonetic fieldwork was also very carefully scheduled in terms of the time management and corpus presentation.

2.2.2.3.1 Time management

Each informant was assigned an individual time slot of 4 hours according to their availability. The time included 0.5 hour’s demonstration and 3.5 hours’ recording. The demonstration time was designed for informants to see how their predecessors were recorded. It helped familiarise the informants with the recording procedure and helped them understand the mechanisms of the work session, enhancing their confidence and reducing their nervousness about being recorded. In addition, the informants were able to look at the word list to get an overall impression of what were being elicited.

The 3.5-hour recording session was broken into three formal elicitation sessions (3*1 hours) and one interaction session (0.5 hour). Each formal session of about one hour involved four small tasks:

- Reading monosyllabic tokens for citation tone elicitation,
- Reading disyllabic tokens for tone sandhi elicitation,
- Reading multi-disyllabic tokens for specific tone sandhi investigation,
- Reading supplementary tokens of local cultural relevance, for example, place names, numbers, food, and so on.

Breaks of about 5 minutes were taken between tasks. During the breaks, speakers could rest their voices and have water, dessert, or fruit to recover their energy, and I could save individual sound files. Normally, speakers might provide feedback about their recording experiences and correct items not properly produced. The three sessions actually involved the same tasks and elicited the same data sets but had differences in the ordering of tokens to control for variation in voice volume and intensity that frequently occurs when reading the same tokens in three times in sequence.

The 0.5-hour interaction session was used to elicit informants' personal information, including date of birth, birth place, current residence, language acquisition, education, and occupation (see Tables 2-2 and 2-3). If time allowed, informants were encouraged to read a short narrative, tell a local story, or produce other utterances that they were interested in and wanted to share.

2.2.2.3.2 Corpus presentation

Tokens to be elicited were all written in simplified Chinese characters and presented via Powerpoint, with one slide for one token, rather than using a typed list. Presenting corpus in Powerpoint offered several advantages.

- It ensured each token individually shown on the slide would be produced in a clear and unexaggerated voice, with balanced and well-controlled intensity and speech rate. On the contrary, using a typed list might cause speakers to articulate tokens with a higher F0 and amplitude at the beginning but with a lower or reduced F0 and amplitude at the end.
- It could minimise possible sandhi and intonation effect on the tokens being elicited. Using a typed list, on the other hand, might induce some speakers to read tokens in a quick succession, which often contains unnecessary sandhi and intonation information.
- It also controlled speakers' emotional state and motivated them to produce utterances in a natural and coherent state throughout the elicitation process.

All recordings were digitised at a sampling frequency of 44100 Hz in Praat, which is reported to be fast enough to capture the highest sinusoidal frequencies detectable by the human ear (Huckvale, 2012, p. 195). Each created sound file was then saved and named with a corresponding code for further data processing and analysis. The sound files for each individual speaker can be referred to in the Appendix C (attached USB).

This field trip proved to be very productive and enjoyable. Twenty-one speakers were recorded for experimental purposes and another six speakers for vocabulary documentation.

2.3. Auditory Data Processing

The processing of language data obtained from the field site largely involved three stages in this study: auditory processing, acoustic processing, and statistical testing. The auditory processing transformed the audio data into a graphic record with detailed phonetic information. The acoustic processing quantified the observations of speech production, and visualised the patterns in an objective way. The statistical testing process revealed the internal relation between variables of tonal parameters to assess whether the observed differences were attributed to the inherent variability or other factors, such as the influence of adjacent sounds or position effect. The three processes were closely related to each other and benefited from each other to produce a scientific, objective, measurable, and testable description and explanation of the nature of Zhangzhou tones.

An auditory description of speech sounds is important in phonetic and phonological studies. As Rose stated (2014, p. 101) "prior to phonological analysis and acoustic quantification, an auditory transcription was needed to represent the range of allophonic variation accurately and sufficiently and to signify the contrasting distinctions among speech sounds." In this thesis, Chapters 3, 4, and 6 focus on auditory descriptions of Zhangzhou segments, citation tones, and disyllabic tones, respectively. These chapters are intended to provide an overall segmental and tonal pattern of Zhangzhou with sufficient phonetic information to assess their distribution, phonemic status, and the interaction between segments and suprasegments. They also serve as a framework for subsequent acoustic quantification and statistical testing, and provide an impressionistic basis for theoretical considerations.

In this study, segments were all transcribed using IPA 2005 symbols. Diacritics were superimposed to encode sophisticated phonetic details, such as change in voice quality (e.g., breathy and creaky voice) or secondary articulation as a consequence of phonological processes (e.g., nasalisation and palatalisation).

The primary auditory correlate of tone—pitch—was transcribed using the five-level pitch notation system of Chao (1930), the most commonly used and preferred system in Asian tonal studies. This numerical system simply divides a speaker's pitch range into four equal parts and labels the five boundaries with the integers *1* through *5* to represent low, mid-low, medium, mid-high, and high pitch, respectively. For example, level tones are represented by two instances of the same number (e.g., [33] or [55]), and rising and falling tones are represented by increasing and decreasing numbers (e.g., [25] or [51]). When required, a third number is included to denote the turning point of a more complex shape. For example, the dipping tone in Mandarin is transcribed as [214].

This five-level approach has been claimed, on theoretical and typological grounds, to be sufficient to characterise tonal contrasts in the world's tonal languages, especially in Asia (Wang, 1967; Yip, 2002), because the maximum number of tonal pitch level contrasts is cross-linguistically claimed to be five (Wang, 1967). Nevertheless, the accuracy and sufficiency of this model has been questioned by a considerable number of scholars, especially in recent decades (Anderson, 1978; Wedekind, 1983; Maddieson, 1990; Duanmu, 2000; Steed, 2011; Zhu, 2012; Paterson, 2015; Rose, 2014, 2016).

For example, Rose (2014) took a likelihood-based quantitative evaluation approach to Chao's model using the data from two Chinese dialects—Wencheng Wu and Cantonese. The statistical result revealed Chao's five-point scale "is probably not an accurate reflection of the distribution of tonal pitch targets" (p. 110). As indicated in Figure 2-4, the pitch offset for the low falling tone is considerably lower than the lowest pitch level specified with tone 1 in Chao's system.

With respect to the problem of the insufficiency of Chao's model to characterise pitch realisation, scholars (Anderson, 1978; Zhu, 2012; Paterson III, 2015; Rose, 2014; 2016) have proposed different solutions, but they have not yet achieved general concensus. For example, Zhu (2012) proposed a new tonal model with the number 6 denoting the pitch value associated with falsetto voice in the high register, as found in Yueyang of the Xiang dialect in China. Paterson (2015) suggested the bar notation as a solution for phonetic pitch transcription without bias concerning phonological claims. Anderson (1978) and Rose (2016) suggested using up-step or down-step arrow to express tonal pitch height beyond Chao's five levels.

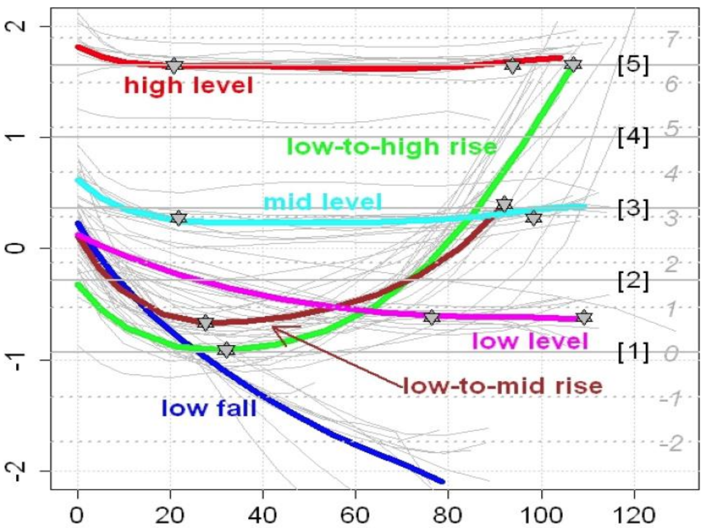


Figure 4: *Z-score normalisation of perceptually transformed tonal F0 for 10 Cantonese speakers. Stars indicate location of tonal pitch targets. X-axis = normalized duration (%), y-axis = standard deviations around mean. Dotted lines = semitones, solid lines = Chao pitch levels.*

Figure 2-4. Example of the inadequacy of Chao's pitch notation system (Rose, 2014, p. 104).

This study is focused on maximally reflecting and respecting the phonetic reality of tonal realisations from both the auditory and acoustic perspectives. Chao's pitch notation system is used to describe the auditory property of tonal pitch but is revised by adding additional tone number to denote those pitches that go beyond the pitch range of five levels. For example, the entire pitch range is discovered being shifted up for the tones in the phrase-initial environment where a new form with an extra pitch level and extra short length also emerges, thus, a pitch range of 2-6 is adopted for those phrase-initial tones (see Chapters 6 and 7).

2.4. Acoustic Data Processing

Phoneticians have a long tradition of quantifying the observations of production and perception, on the basis of which the hypothesis testing and regression techniques have been extensively used (Rasinger, 2013; Rose, 2016). An important portion of this thesis is devoted to quantifying two main acoustic correlates of tone—F0 and duration—to discover patterns in normalised data, and using pairwise *t*-test comparisons to determine the statistical relation among different normalised F0 and duration levels. This section describes how the procedure of acoustic processing was used in this study, involving acoustic measurement, acoustic normalisation, and normalisation assessment.

2.4.1. Acoustic measurement

The measurement of tonally relevant information—F0 and duration—in digital signals was conducted in Praat, version.5.3. Praat is a free phonetic software for analysing, synthesising, and manipulating speech and other sounds in a single integrated computer program, and allows one to perform acoustic analysis by inspecting visualised signals, such as the waveform, spectrogram, formant tracks, F0 curve, intensity curve, and spectrum (Boersma, 2001; Styler, 2013).

Three important aspects were taken into consideration while using Praat for measuring. (1) How is the tonally relevant duration in acoustic signals determined? (2) How should the F0 contours showing undesirably discontiguous trajectory from the perturbation effect of non-modal phonation be handled? (3) How is the tonally relevant information—F0 and duration—extracted from digital signals for further analysis?

2.4.1.1. Determining the tonally relevant duration

In alignment with most acoustic studies of Sinitic tones (e.g., Rose, 2008, 2014, 2016, 2016; Steed, 2011; Huang et al., 2016), the tonally relevant F0 duration was considered the rhyme portion of each monosyllable, which, in this study, included compulsory nucleus, optional pre-vocalic glide and/or postvocalic coda. The syllable onsets are generally not regarded as tone-bearing segments; while this is a theoretical possibility, it is not pursued further here.

The judgment of F0 contour onset and offset was made according to the acoustic manifestations of air pressure fluctuations in the waveform, for example, periodicity and amplitude, and of individual syllable components in the broad-band spectrogram, for example, VOT, formant pattern, and F0 curve. The onset of the F0 contour was generally set at the second strong glottal pulse where the amplitude of air pressure fluctuation began to increase, and periodicity of speech wave vibration tended to be regular in the waveform. Correspondingly, the formant patterns in the spectrogram appeared clearly stable and identifiable. Nevertheless, the third glottal pulse was often adopted, especially when the F0 contour was seriously perturbed by aspirated syllable onsets.

The offset of the F0 contour, on the other hand, was set at the point where the amplitude in the waveform decreased quickly, the waveform was strongly damped, and the periodicity was becoming irregular. Further, formant patterns for vocalic sounds tended to be blurred and were no longer visible and identifiable in the spectrogram.

The determination of the tonally relevant duration is shown in Figure 2-5, a monosyllabic example, and in Figure 2-6, a disyllabic example. In both figures, each syllable is segmented into syllable onset and tonally-relevant duration.

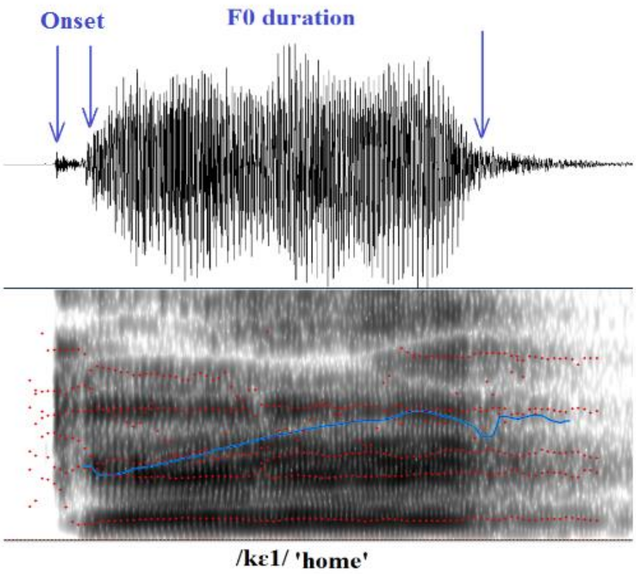


Figure 2-5. Praat labelled monosyllabic example /kɛ1/ ‘home’ (WYF, male).

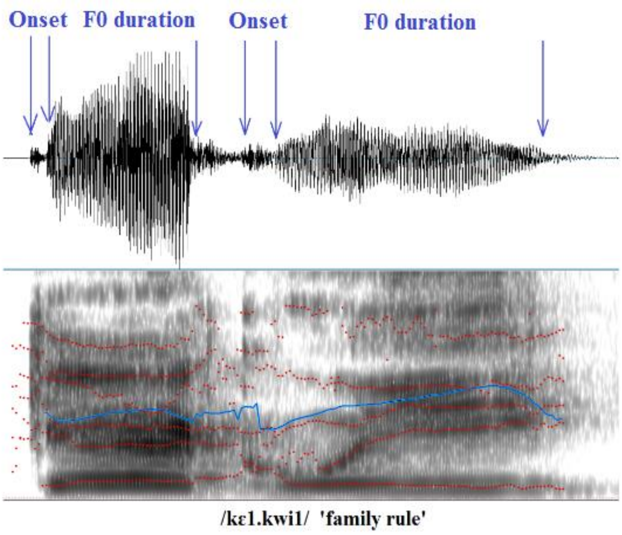


Figure 2-6. Praat labelled disyllabic example /kɛ1.kwi1/ ‘family rule’ (WYF, male).

2.4.1.2. Dealing with discontiguous F0 trajectories caused by non-modal phonation

The F0 contour is often perturbed by the articulation of tautosyllabic segments and the physical differences in individual vocal tracts. For example, it is very common to see an undesirable and unreasonable F0 trajectory presented in the spectrogram in Praat when the related sound is produced with a non-modal phonation, especially a creaky voice. As illustrated in Figure 2-7, the F0 contour in blue presents a discontiguous and jumping property since the related morpheme 'seven' in Zhangzhou is produced with a creaky voice.

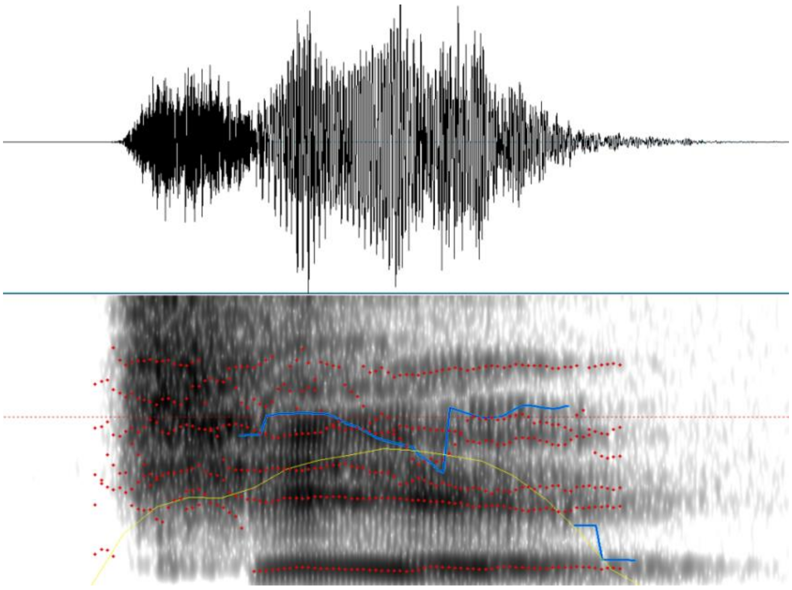


Figure 2-7. Praat example of unreasonable F0 trajectory of /tsʰit6/ 'seven' (WYF, male).

This discontiguous and variable property of the F0 trajectory often causes the F0 extraction algorithm to fail to operate and extract the related acoustic values. In dealing with an F0 contour of this kind, a manual modification was generally used to pull the problematic period back onto the right track in the Pitch window in Praat (Boersma, 2001; Styler, 2013), as indicated in the Figure 2-8.

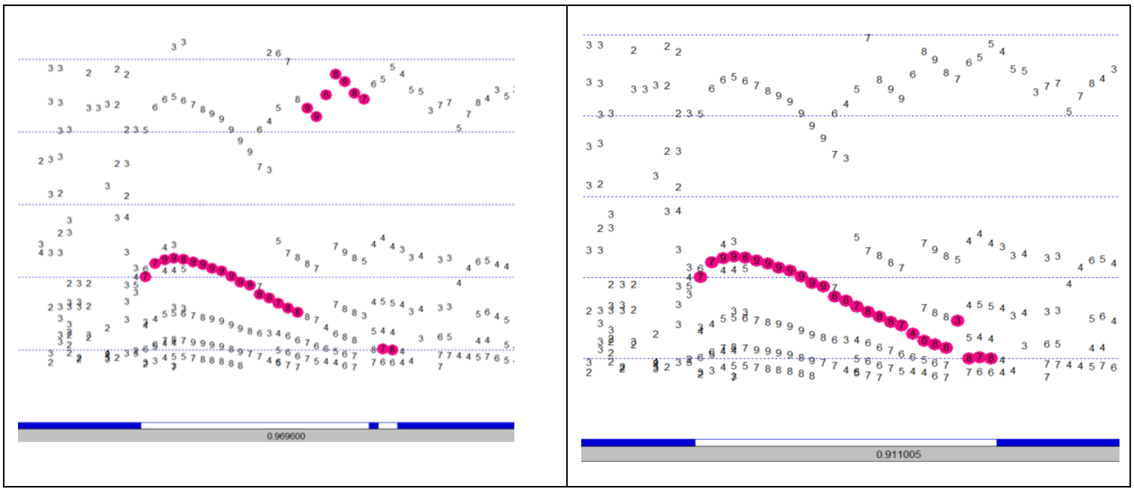


Figure 2-8. Praat example of dealing with unreasonable F0 trajectory of /tsʰit6/ 'seven' in the Pitch window (WYF, male).

The red contour in the left panel in Figure 2-8 is essentially a copy of the F0 contour in blue in Figure 2-7, both of which show discontiguous trajectories. The F0 contour in the right panel in Figure 2-8 is the modified result of the contour on the left, in which problematic points in the higher frequency area were largely realigned onto the path of red disks supposed to be the reasonable representation of the F0 curve of this token.

In this study, most of the discontiguous F0 trajectories were manually modified in this way. After the modification, a reasonably accurate F0 curve could be displayed, measured and extracted for further processing and analysis. Nevertheless, a very small number of problematic points may be retained even after adjusting the F0 tracking in the Pitch window. As shown in the right panel in Figure 2-8, one red disk still jumps off the main trajectory. In this case, the value from a nearby sampling point was selected in this study. Thus, constrained by the potential technical problems with Praat, it appears important to make use of a relatively large corpus for an acoustic study. The data from the 21 speakers in this study should be sufficient to derive a result that is representative of this language.

2.4.1.3. Extracting tonal F0 duration from the acoustic signal

Once the tonal durations of all syllables in target tokens were determined and labelled and the problematic F0 trajectories were modified to reflect a true representation, an F0 extraction script, designed by Prof. Phil Rose and revised by me, was run on the selected tonal duration to extract both F0 and duration values automatically at 10 equidistant sampling points. Having 10 sampling points largely ensured sufficient phonetic details of the selected portion were captured for acoustic analysis and statistical testing.

All the extracted F0 and duration values were first saved in Excel spreadsheets and then transferred to text files for further plotting and processing in R. Generally, 160 tokens were used for the quantitative analysis of citation tones, resulting in 33600 F0 values (= 160 tokens * 21 speakers * 10 sampling points) and 3360 duration values (= 160 tokens * 21 speakers) being extracted. In addition, about 598 disyllabic tokens with 1196 syllables were extracted from each speaker. Thus, 251,160 F0 values (= 1196 syllables * 21 speakers * 10 sampling points) and 25,116 duration values (= 1196 syllables * 21 speakers) were used for the quantitative and statistical analysis of disyllabic tones in this study.

2.4.2. Acoustic normalisation

The acoustic signals are highly variable because they carry not only the linguistic structures of the utterances but also a wealth of extralinguistic information, including a speaker’s sociocultural background, pragmatic intent, attitude and emotional status, and vocal tract anatomy and physiology of individual speakers (Anderson, 1978; Ladefoged, 1999; Harrington, 2010; Rose, 1986, 2000, 2016; Foulkes et al., 2013; Huang et al., 2016). For example, the anatomical and physiological differences in individual vocal tract structures can generate dynamic acoustic outputs for phonologically identical utterances, and it is common to see female speakers having higher F0 values than males because of their shorter and less massive vocal cords.

The acoustic variability resulting from the entwined linguistic and extralinguistic information makes it necessary to abstract the variable indexical content from invariable linguistic content in speech signals (Rose, 1986, 1996, Huang et al., 2016). The process of normalisation is designed to accomplish this goal and to derive a linguistic phonetic representation of the variety in question, making this variety comparable to other languages/varieties with respect to language-internal properties (Anderson, 1978; Rose, 1986, 1996, 2000, 2016; Ladefoged, 1999; Huang et al., 2016).

2.4.2.1. F0 normalisation

Several normalisation strategies have been proposed and compared for achieving an effective reduction in the between-speaker variation in tonal F0 (e.g., Earle, 1975; Shi, 1986; Shi et al., 2010; Rose, 1986, 1987, 1993, 2016; Zhu, 2004). For example, Earle (1975) proposed a *z*-score normalisation of the tones of standard Vietnamese. Shi (1986) proposed a *t*-value transform approach for a corpus with single speaker of Chinese languages, and Shi et al. (2010) suggested a revised *t*-value normalisation for a large corpus with multiple speakers. Zhu (2004) compared six different normalisation transforms while Rose (2016) conducted a comparative-quantitative study to judge the efficiency of seven normalisation strategies using the citation F0 data in four Chinese dialects.

Among different approaches, the *z*-score normalisation has been demonstrated superior to other strategies (Rose 1987, 1993, 2016), and it has been widely adopted in tonal studies (e.g., Steed, 2011; Shen & Rose, 2016; Huang et al, 2016). This study also used the *z*-score normalisation approach to reduce the individual-dependent variance of 21 speakers. The *z*-score normalised F0 value Z_i is calculated using the following formula (Huang et al., 2016):

$$Z_i = (X_i - m)/s.$$

In this formula, normalisation parameters *m* and *s*, separately, stand for raw mean F0 value and the standard deviation estimated from all sampling F0 values over all tokens of all tones from a given speaker under consideration. X_i is an observed F0 value at a given sampling point while Z_i is its corresponding normalised value derived as the distance from the mean F0 value, corresponding to speakers' neutral pitch. Therefore, the normalised F0 contour is expressed in the unit of standard deviation. Assuming the variables are distributed normally, nearly 95% of the normalised F0 values are supposed to be distributed between -2 and +2 standard deviations away from the mean value. In this study, each raw F0 value of individual speakers extracted from Praat was transformed into its corresponding normalised value, forming a new dataset to be further plotted and statistical tested with respect to specific research tasks, including

- What are individual citation tones realised in terms of *z*-score normalised F0 from 21 speakers? (Chapter 5)
- What are individual phrase-initial tones realised across their following tones in terms of *z*-score normalised F0 from 21 speakers? (Chapter 7)
- What are individual phrase-final tones realised across their preceding tones in terms of *z*-score normalised F0 from 21 speakers? (Chapter 8)

2.4.2.2 Duration normalisation

To retain the linguistic information of the absolute tonal duration, raw duration values are normalised using the following formula (Huang et al., 2016):

$$D_{norm} = (D/D_{mean}) * 100.$$

In this formula, the normalisation parameter D_{mean} represents the mean raw duration estimated from the average duration of all tokens over all tones from a given speaker being investigated. D is the duration observed for a given tone while its corresponding normalised value D_{norm} is expressed as a percentage of a speaker's average duration values over all their tonal production. For example, if a duration is calculated over 100%, it means the tone under consideration has a longer duration than the average. Similarly, each extracted raw duration value of individual speakers was also transformed into its corresponding normalised value, forming a new dataset to be further plotted and statistical tested with respect to specific research issues, including

- What are individual citation tones realised in terms of normalised duration from 21 speakers? (Chapter 5)
- What are individual phrase-initial tones realised across their following tones in terms of normalised duration from 21 speakers? (Chapter 7)
- What are individual phrase-final tones realised across their preceding tones in terms of normalised duration from 21 speakers (Chapter 8)

2.4.2.3. Normalisation evaluation

The efficiency of a normalisation process, as proposed by Earle (1975), can be quantified by the normalisation index (NI), which represents the ratio of the dispension coefficients of the normalised and unnormalised acoustic data (Rose, 1986, 1987). This index reflects how much proportion of variance, intermingled with unnormalised acoustic data and caused by physical differences in the mass and length of individual vocal folds, is reduced. The higher the NI value, the greater degree of speaker-dependent variation is abstracted, and the clearer linguistic-phonetic content of the signals is obtained (Huang et al., 2016).

The derived normalised data were further plotted into numbers to present the patterns in a visible, precise, and generalisable way. Further, the data were used for statistical testing concerning the assumptions based on both auditory and acoustic observation. The data plotting and statistical testing were conducted using a variety of R codes, originally provided by Prof. Phil Rose but improved by my colleague Siva Kalyan for their use with Zhangzhou data.

2.5. Statistical Testing

In spite of the best efforts to appropriately select and measure a solid body of high-quality data, certain remaining differences among the groups were attributed merely to the inherent variability of samples rather than observed differences in data values (Butler, 1985; Johnson, 2008). Thus, applying statistical techniques to test the hypotheses formulated from the observed patterns is crucial in determining whether the observed differences between sets of data could reasonably have been expected to occur by chance, for example, as a result of sampling variation or other factors (Butler, 1985; Dörnyei, 2007; Johnson, 2008; Woodrow, 2014; Levshina, 2015).

This study takes advantages of pairwise *t* test, hierarchical clustering algorithm, and the SplitsTree software to reveal those patterns and structures that remain hidden in the raw data, which are then presented in a logically structured, visually tangible, and scientifically reliable way.

2.5.1. Pairwise *t*-test comparison

While the *t* test is limited to comparing the means of two groups, the function of pairwise *t* tests allows for performing multiple *t* tests on a multitude of groups (Levshina, 2015). Thus, the ordinary *t* test is apparently not suitable for this study involving 8 citation tones and 64 disyllabic tones from 21 speakers. Correspondingly, it is appropriate to employ the pairwise *t* tests in this study to determine whether the variables (e.g., F0 or duration) among a population (e.g., sets of tone) differ from each other in a significant way under the assumption that the paired differences are independent and identically normally distributed.

For example, examining whether the eight Zhangzhou citation tones differ significantly in duration requires all possible pairwise comparisons of the values derived from acoustic quantification and normalisation. Figure 2-9 shows the comparison result among 28 (=8*7/2) paired differences.

	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 2-9. Pairwise *t*-test comparison of normalised duration among Zhangzhou citation tones.

During the application of pairwise *t* tests, the Bonferroni correction has to be performed to control for the Type I Error and achieve significance (Levshina, 2015). The corrected alpha is calculated by dividing the critical P value by the number of comparisons under consideration. For example, in this duration case, the corrected alpha is 0.00186 (= 0.05/28). If the calculated *t* value is less than the corrected alpha, then the paired difference is considered to be statistically significant, and vice versa. For example, the durational difference between tones 7 and 1 is statistically significant (2e-16<0.00186), while it is not significant between tone 8 and tone 2 (1>0.00186).

As discussed, pairwise *t* tests can help determine whether the observed variables between sets of data are statistically significantly different from each other. The testing result can be visualised using the hierarchical clustering algorithms, which can help assess how many levels the sets of data can be clustered into from the scientific point of view. For example, Figure 2-10 presents the clustering result of normalised length levels of Zhangzhou citation tones by pairwise *t* tests.

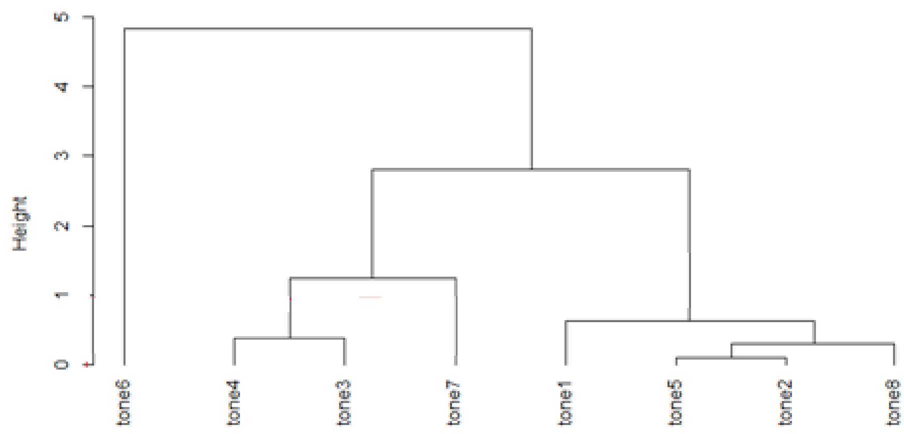


Figure 2-10. Clustering of normalised length levels of Zhangzhou citation tones by pairwise *t tests*.

The vertical lines in Figure 2-10 are branches representing the amount of significant difference in duration across different tones. The longer the branch in the vertical dimension, the larger the amount of significant difference in duration. The horizontal dimension indicates which tone connects to which in terms of length. The closer the tones, the higher the probability that they belong to the same length level. The scale at the far left of the figure is the threshold distance for significance. The higher the threshold, the lesser the length levels that can be generated.

This study primarily selected the threshold at 1 to determine the number of significantly different levels for normalised F0 and duration. The clustering results at this threshold were largely consistent with results determined on the basis of either the auditory impressions or the acoustic quantifications, which will be presented in Chapters 5, 7 and 8. Nevertheless, the threshold may be slightly modified for the data set having a relatively larger effect size. For example, the clustering of normalised F0 contours for citation tones involved 16 putative points to be pairwise tested and hierarchically clustered. A selection of threshold at 1.5 can group the putative points into higher-quality clusters (see Chapter 5). Once the number of the length level has been decided, the ranking can be further conducted for a better representation.

The pairwise *t* tests and the hierarchical clustering algorithms were mainly employed to address a series of specific research question in the thesis, including

- How many normalised F0 and duration levels are contrastive among Zhangzhou citation tones? (Chapter 5)
- Are the F0 and duration realisations of Zhangzhou phrase-initial tones affected by their following tones? If so, to what extent are they affected, and what conditions the variations? (Chapter 7)
- How many normalised F0 and duration levels are contrastive among Zhangzhou phrase-initial tones? (Chapter 7)
- Are the F0 and duration realisations of Zhangzhou phrase-final tones affected by their preceding tones? If so, to what extent are they affected, and what conditions the variations? (Chapter 8)
- How many normalised F0 and duration levels are contrastive among Zhangzhou phrase-final tones? (Chapter 8)

2.5.2. SplitsTree

This study asserts that tonal realisations in Zhangzhou are multidimensional, involving a variety of co-varying phonetic parameters that include pitch/F0, duration, vowel quality, voice quality, and obstruent coda. Each tone thus consists of a bundle of phonetic outputs, forming a multidimensional framework for the whole tonal system (see later chapters). Revealing the patterns and structures that remain hidden within the geometry is imperative for understanding the relatedness between tones from the phonological point of view.

This study employs the SplitsTree software (Kloepper & Huson, 2008) to generate phylograms in order to hierarchically visualise the mapping for a set of tones under consideration. The phylogenetic network is applied under the assumption that each tone in a sequence can change its phonetic outputs independently from the other tones. Before a phylogram is created, the phonetic outputs for a set of tones under investigation have to be transformed into a multiple sequence alignment to be computed in the SplitsTree software. The sequencing of the multiple alignment varies along with the changes of tonal realisations across different linguistic contexts.

Figure 2-11 demonstrates how the phylogenetic tree works for the mapping of phonological relatedness between Zhangzhou tones in this study. The root of the tree encodes what type of data set is under consideration (e.g., Zhangzhou citation tones). The horizontal lines are branches representing the amount of similarity in terms of multidimensional parameters across tones being investigated. The shorter the branch in the horizontal dimension, the stronger the similarity that tones share in phonetics, and vice versa. The vertical dimension indicates which tone connects to which other tone. The more divergent the tones, the higher the probability that they are phonologically unrelated, and vice versa. The bar on the top of the phylogram shows the length of the branch with an amount similarity change of 0.07.

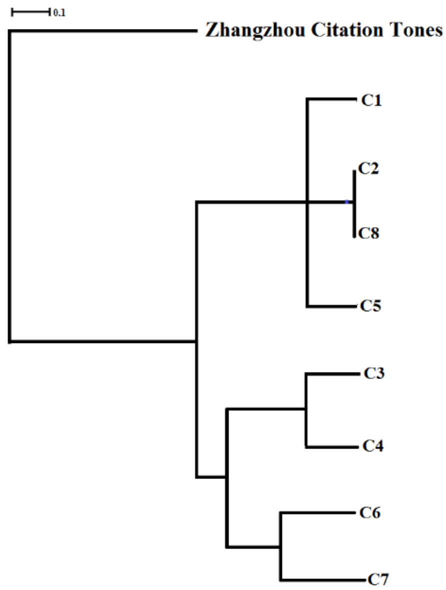


Figure 2-11. Example of phylogenetic mapping for Zhangzhou tones (C=Citation tone).

The SplitsTree software was mainly employed to address the relatedness between a set of tones from the phonological perspective, the results of which can significantly help to develop a conception of the totality of tonal contrasts and tonal neutralisations. The software was applied only when the multidimensional properties for each set of tones had been established and had been transformed into a multiple sequence alignment. The application of SplitsTree was mainly to address a series of specific research questions, including

- How are Zhangzhou tones related to each other in citation? (Chapter 5)
- How are Zhangzhou tones related to each other phrase initially? (Chapter 7)
- How are Zhangzhou tones related to each other phrase finally? (Chapter 8)
- How are Zhangzhou phrase-initial tones related to citation tones? (Chapter 9)
- How are Zhangzhou phrase-final tones related to citation tones? (Chapter 9)

The techniques of pairwise *t* testing, the hierarchical clustering algorithm, and the SplitsTree enabled the identifications of the relations among independent variables as logically structured, visually tangible, and objectively presented, which helped this study achieve a higher level of generalisation and interpretation. The procedure of pairwise *t* tests and hierarchical clustering used in this study was accomplished with the very kind assistance of my colleague, Dr. Siva Kalyan. The application of SplitsTree was conducted under the instruction of my supervisor, Dr. Paul Sidwell. Many concepts with respect to statistical testing were also enlightened by the precious feedback and advice from Prof. Phil Rose. I acknowledge their generous help. The codes for the pairwise *t*-test comparison and clustering, as well as the multiple sequence alignments for the SplitsTree operation are all provided in the attached USB.

2.6. Conclusion

This chapter has introduced the methodologies used in this project, involving the phonetic fieldwork, auditory data processing, acoustic data processing, and statistical testing. These methodologies are derived from but also constrained by the research questions proposed in Chapter 1. They form a cohesive and coherent framework for addressing the research questions from different perspectives.

This methodological-analytical framework ensures this study has a strong foundation of a generalisable sample, reliable and valid instrumentation, and credible and scientific testing. It allows for description and quantification of tonal realisation and interaction and testing the hypotheses formulated based on the observed auditory and acoustic patterns. The results of the applications of intersecting methodologies will be described and explained in the following chapters.

Chapter 3: Zhangzhou Syllables and Segments

This chapter provides a detailed description of Zhangzhou syllables and segments from the perspectives of phonetics and phonology. It largely addresses three issues. (1) How are syllables structured in Zhangzhou? (2) What are the syllable components? (3) What are the individual syllable components realised at the phonetic level? and how can they be treated phonologically? The descriptions and analyses offered in this chapter go beyond earlier studies by relying on more objective data and employing more up-to-date methods.

3.1. Syllables

This study asserts Zhangzhou exhibiting a C(G)V(C) syllable structure in which the syllable onset and nucleus are obligatory while the glide and coda are optional. The prevocalic glides are treated as an independent syllable component (G) while the postvocalic glides are considered as one type of syllable coda (C). The glottal stop is considered a phonemic syllable onset (C), but in the rhyme as a laryngeal feature that generally accompanies the creaky phonation in falling-pitched low vowels, rather than a contrastive syllable coda. Consonantal clusters do not occur. The reasons for these phonological treatments are discussed in detail in the rest of this chapter.

In addition, Zhangzhou essentially features four syllable types—CV, CGV, CVC, and CGVC—, as illustrated in Table 3-1 (in which segments are transcribed phonemically). The data from multiple speakers make it possible to generalise the phonemic inventory for individual syllable components—onset, glide, nucleus, and coda—, as shown in Table 3-2.

Table 3-1. Examples of different syllable types in Zhangzhou

CV	CGV	CVC	CGVC
ɗĩ22 'year'	sjɐ51 'write'	sim35 'heart'	tsjɐp41 'juice'
ɓɐ41 'meat'	tsʰjɵ41 'laugh'	tit41 'bamboo'	kwɐj35 'obdient'
ku33 'worn'	kwɐ35 'song'	kɐw22 'monkey'	kjɐm22 'salty'
tɵ51 'island'	tswɐ̃22 'spring water'	ħut221 'Buddha'	tsʰjɐt41 'chop'
ʔɔ35 'dark'	ʔjɔ̃22 'sheep'	ʔɔŋ33 'prosperous'	ʔjɐw35 'hungry'

Table 3-2. The phonemic inventory for individual syllable components in Zhangzhou

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

Nevertheless, it is important to realise that conducting a linguistic study of the sound system of a given language not only entails determining how many speech contrasts exist but also revealing and explaining a series of occurring questions, for example,

- How are phonemes realised phonetically?
- How are they distributed phonotactically?
- What conditions allophonic variation?
- How can the observed variation be interpreted formally?

The rest of this chapter is devoted to exploring and discussing Zhangzhou segments with respect to the above questions.

3.2. Onsets (C)

Zhangzhou shows a relatively small onset inventory of 15 phonemes: /p/, /pʰ/, /ɓ/, /t/, /tʰ/, /ɗ/, /k/, /kʰ/, /ɠ/, /ts/, /tsʰ/, /s/, /z/, /ħ/, and /ʔ/. Careful transcription and instrumental analyses reveal various allophones conditioned by various factors, as shown in Table 3-3. The symbol * signifies no relevant sound was found in the data.

Table 3-3. Onset allophony in Zhangzhou summarised

Manner	Place	Phoneme	Phonetic realisation				
			_/[i; j]	_/[u; w]	_/[Ṽ]	_/[N̩]	elsewhere
Occlussive	bilabial	p	p	pʷ	p	p	p
		pʰ	pʰ	pʰʷ	pʰ	*	pʰ
		ɓ	ɓ	βʷ	m	*	ɓ
		tʰ	t̪ʰ	t̻ʰʷ	tʰ	tʰ	tʰ
		ɗ	ɗ̪	lʷ	n	*	ɗ
	velar	k	kʲ	kʷ	k	k	k
		kʰ	kʰʲ	kʰʷ	kʰ	kʰ	kʰ
		ɠ	ɠʲ	ɣʷ	ŋ	*	ɠ
	glottal	ʔ	ʔʲ/∅	ʔʷ/∅	ʔ/∅	ʔ/∅	ʔ/∅
Fricative	alveolar	s	ɕ	ʃ	s	s	s
		z	ʝ	ʒ	*	*	*
	pharyngeal	ħ	ħʲ	hʷ	ɦ	ɦ	ħ
Affricative	alveolar	ts	tɕ	tʃ	ts	ts	ts
		tsʰ	tɕʰ	tʃʰ	tsʰ	tsʰ	tsʰ

As indicated in Table 3-3, the articulatory realisations of syllable onsets are diverse and dynamic, conditioned by a variety of factors, including the palatals ([i] and[j]) and labial-velars ([u] and

[w]), nasalised vowels [Ṽ], and syllabic nasals [N̩]. One onset can have several allophonic variants as a consequence of the operation of different phonological processes, such as nasalisation, lenition, labialisation, palatalisation, glottalisation, or laminalisation. For example, the bilabial implosive [ɓ] is weakened and labialised to a voiced fricative [βʷ] when it precedes the rounded vowel [u] but shifts to the nasal [m] before nasalised vowels, as shown in Table 3-4.

Table 3-4. Examples of allophonic variations of implosives in Zhangzhou

Onset	Before [u]	Before [Ṽ]	Elsewhere
/ɓ/	[βʷṳ51] 'dance'	[mĩ::33] 'noodle'	[ɓi̤::33] 'flavour'
/ɗ/	[lʷṳ51] 'female'	[nɔ̃::33] 'two'	[ɗɔ::33] 'road'
/ɠ/	[ɣʷṳ::22] 'cow'	[ŋɛ̃::33] 'stiff'	[ɠɛ::22] 'sprout'

In addition, the various processes tend not to be independent of each other but can function together to shape the dynamic realisations of Zhangzhou syllable onsets. This section provides a detailed description of how various allophones are derived in the relatively simple onset system of Zhangzhou and how they can be interpreted formally. What deserves further mentioning with respect to Table 3-3 is why /z/ is proposed if it seems never exist at the surface. The main reason for this treatment is that /z/ presents a neat pattern similar to its corresponding voiceless counterpart /s/. They are both found undergoing labialisation before rounded segments and palatalisation before palatal sounds. Similar to the [s] sound, /z/ is expected to be realised as a [z] sound elsewhere; however, there happens existing accidental gaps because no tokens were found in the local vocabularies. In this case, it is still proper to propose a /z/ phoneme to maintain a symmetrical pattern with the /s/ onset, although it seems never surface.

3.2.1. Palatal-conditioned variation

The palatals—[i] and [j]—condition two processes—dentalisation and palatalisation—in phonemes that are not labial plosives in Zhangzhou. They cause the alveolar occlusives—/t/, /tʰ/, and /ɗ/—to have dental allophones—[t̪], [t̪ʰ], and [ɗ̪]—at the surface, as expressed by Rule (1) and illustrated in Table 3-5. On the other hand, the palatal sounds condition the other nine onset phonemes that are non-labial and non-alveolar occlusives to be realised with a palatalised feature, as expressed in Rule (2) and illustrated in Table 3-6.

(1) Dentalisation: [+alveolar; + occlusive]→ [+dental]/_ [+palatal]

Table 3-5. Examples of the dentalisation process in Zhangzhou

Onset	Before [i]	Before [j]
/t/	[t̪i̤::35] 'pig'	[t̪jɐ::22] 'buy (rice)'
/tʰ/	[t̪ʰi̤41] 'shave'	[t̪ʰjɐ̰:221] 'overlay'
/ɗ/	[ɗ̪i̤51] 'you (2sg.)'	[ɗ̪jɐ̰:221] 'catch'

(2) Palatalisation: [-bilabial; -alveolar; +occlusive]→ [+palatal]/_[+palatal]

Table 3-6. Examples of the palatalisation process in Zhangzhou

Onset	Before [i]	Before [j]
/k/	[kʲi̤51] ‘raise’	[kʲjɐ̰41] ‘send’
/kʰ/	[kʰʲi̤41] ‘build’	[kʰʲjɐːː33] ‘stand’
/ɠ/	[ɠʲi̤51] ‘speech’	[ɠʲjɐːː22] ‘carry’
/s/	[ɕi̤ːː35] ‘silk’	[ɕjuːː35] ‘repair’
/z/	[ʝi̤ːː22] ‘character’	[ʝjuːː22] ‘gentle’
/ts/	[tɕi̤51] ‘elder sister’	[tɕjɵːː22] ‘stone’
/tsʰ/	[tɕʰi̤41] ‘test’	[tɕʰjɐːː35] ‘vehicle’
/ħ/	[ħʲi̤ːː22] ‘fish’	[ħʲjɐːː35] ‘there’
/ʔ/	[ʔʲi̤ːː22] ‘move’	[ʔʲjɐːː33] ‘scatter’

3.2.2. Labial-conditioned variation

The labial sounds—[u] and [w]—condition four different processes in Zhangzhou onsets, including lenition, laminalisation, glottalisation, and labialisation. All onset phonemes are labialised before the rounded sounds, suggesting that the lips are simultaneously rounded during the production of related sound, as expressed in Rule (3) and illustrated in Table 3-7. The auditory impression of the alveolar obstruents (/s/, /z/, /ts/ and /tsʰ/) is always rounded before the labial sound [u] or [w] in the data. It is thus appropriate to consider a labialisation process for their palatoaveolar allophones ([ʃ], [ʒ], [tʃ] and [tʃʰ]). Additionally, the processes of lenition, laminalisation, and glottalisation do not operate on their own but intersect with labialisation in Zhangzhou, as discussed below.

(3) Labialisation: C→ [+labial]/_[+labial]

Table 3-7. Examples of the labialisation in Zhangzhou

Onset	Before [u]	Before [w]
/p/	[pʷṳ41] ‘wealthy’	[pʷweːː35] ‘cup’
/k/	[kʷṳːː35] ‘turtle’	[kʷwɐːː35] ‘melon’
/ʔ/	[ʔʷṳːː33] ‘have’	[ʔʷweːː35] ‘pot’
/s/	[ʃṳːː35] ‘lose’	[ʃwɐːː35] ‘sand’
/z/	[ʒṳːː22] ‘such as’	[ʒwɐːː22] ‘hot’
/ts/	[tʃṳ51] ‘master’	[tʃwɐ̰51] ‘paper’
/tsʰ/	[tʃʰṳ41] ‘house’	[tʃʰwĩːː35] ‘cross’

The combined processes of lenition and labialisation largely occur in the peripheral implosives /ɓ/ and /ɠ/, which are realised as labialised and voiced fricatives [βʷ] and [ɣʷ], respectively, before the rounded segment [u] or [w]. The derivation is generalised in Rule (4) with examples in Table 3-8.

(4) Lenition and labialisation: /ɓ, ɠ/→[βʷ, ɣʷ]/_[+labial]

Table 3-8. Examples of combined processes of lenition and labialisation in Zhangzhou

Implosive	Before [u]	Before [w]
/ɓ/	[βʷṳ51] 'dance'	[βʷwi̤::35] 'smile'
/ɠ/	[ɣʷṳ::22] 'cow'	[ɣʷwɐ::33] 'external'

The combined processes of laminalisation and labialisation are seen in the voiceless alveolar occlusives /t/ and /tʰ/, which are realised as [t̪] and [t̪ʰ], respectively, in this labial-conditioned environment. The derivation is expressed by Rule (5) and is exemplified in Table 3-9.

(5) Laminalisation and labialisation: /t, tʰ/→[t̪ʷ, t̪ʰʷ]/_[+labial]

Table 3-9. Examples of combined processes of laminalisation and labialisation in Zhangzhou

Onset	Before [u]	Before [w]
/t/	[t̪ʷṳ::35] 'pile'	[t̪ʷwɐ::33] 'big'
/tʰ/	[t̪ʰʷṳ::41] 'dispute'	[t̪ʰʷwɐ::35] 'drag'

The complicated combination of lenition, laminalisation, and labialisation is seen with the alveolar implosive /ɗ/. It is realised as a labialised approximant [lʷ] in this environment as generalised in Rule (6) and illustrated in Table 3-10.

(6) Lenition, laminalisation and labialisation: /ɗ/→[lʷ]/_[+labial]

Table 3-10. Examples of combined processes of lenition, laminalisation and labialisation in Zhangzhou

Onset	Before [u]	Before [w]
/ɗ/	[lʷṳ51] 'female'	[lʷwɐ::22] 'bamboo basket'
	[lʷṳ::35] 'push'	[lʷwi::35] 'money'

The combined processes of glottalisation and labialisation are also seen in the voiceless pharyngeal fricative /ħ/, which is realised as a labialised glottal fricative [hʷ] before the rounded vowel [u] or the labial-velar glide [w]. This deviation is expressed in Rule (7) with examples in Table 3-11. The glottalisation in this study represents a partial constriction for this onset production at the glottis rather than in the pharyngeal cavity.

(7) Glottalisation and labialisation: /ħ/→[hʷ]/_[+labial]

Table 3-11. Examples of combined processes of glottalisation and labialisation in Zhangzhou

Onset	Before [u]	Before [w]
/ħ/	[hʷṳ::22] 'assist by the arms'	[hʷwɐ::35] 'flower'
	[hʷṳ::35] 'skin'	[hʷwi̤51] 'expense'

3.2.3. Nasality-conditioned variation

The nasalised vowels [Ṽ] and syllabic nasals [N̩] constitute the conditioning environment of [+nasal], motivating two different processes of nasalisation and glotalisation but only in specific sets of syllable onsets in Zhangzhou. Voiceless obstruents are observed being able to occur before the nasalised segments at the surface level, while the implosives are prohibited, instead, they are found undergoing nasalisation and become their homorganic nasal plosives before nasalised vowels, as generalised in Rule (8) and shown in Table 3-12.

(8) Nasalisation: [+voiced; +occlusives] → [+nasal]/_[Ṽ]

Table 3-12. Examples of the nasalisation process in Zhangzhou

Implosive	Before [V]	Before [Ṽ]
/ɓ/	[ɓi̤::33] 'flavour'	[mĩ::33] 'noodle'
/ɗ/	[ɗɔ::33] 'road'	[nɔ̃::33] 'two'
/ɠ/	[ɠɛ::33] 'shy'	[ŋɛ̃::33] 'stiff'
/p/	[pi̤::35] 'sorrow'	[pĩ::35] 'side'
/t/	[ti̤::35] 'pig'	[tĩ::35] 'sweet'
/k/	[ki̤::35] 'branch'	[kĩ::35] 'alkaline'

On the other hand, the voiceless pharyngeal fricative undergoes glottalisation and becomes a breathy sound [ɦ] before nasalised vowels and syllabic nasals. The derivation of the breathy variant is expressed by Rule (9), and examples are provided in Table 3-13.

(9) Glottalisation: /ħ/→[ɦ]/_[Ṽ] or [N̩]

Table 3-13. Examples of the glottalisation process in Zhangzhou

Pharyngeal	Before [Ṽ]	Before [N̩]
/ħ/	[ɦɛ̃41] 'rest'	[ɦm̩41] 'hit with a stick'
	[ɦɐ̃41] 'sizzling'	[ɦŋ̩::35] 'recipe'
	[ɦĩ41] 'fling'	[ɦŋ̩41] 'yep'
	[ɦjɔ̃::35] 'joss sticks'	[ɦŋ̩::33] 'moan'

Nevertheless, a series of theoretical concerns may arise with respect to the very abstract nature of the implosive-nasal alternation. For example, why are the implosives chosen to represent the phonemes, and why not just use the egressive occlusives /b, d, g/? Alternatively, why are the implosives considered to be the unmarked forms? Why are the pulmonic /b, d, g/ excluded from the onset phoneme inventory, and why are the nasals considered the derived forms? These questions essentially concern what IPA symbols to be used to represent phonemes having quite disparate allophones.

Three considerations are given for the phonemic treatment of implosives in this study. First, the egressive stops are seldmon used in Zhangzhou. Zhangzhou presents an asymmetric distribution with respect to the airstream mechanism involved in the onset production. All voiceless occlusives—[p, p^h, t, t^h, k, k^h]—are produced with the pulmonic egressive airstream while the voiced occlusives are produced with the glottalic ingressive airstream, forming three contrastive implosives [ɓ, ɗ, ɠ]. The egressive voiced stops—[b, d, g]—are less likely to be found in this language. In other words, they are marked.

Second, the implosives are widely used in Zhangzhou. They are found across different syllable types—for example, [ɓɐ̰41] 'meat', [ɓjɵː33] 'temple', [ɓɔmːː33] 'touch', and [ɗjɔŋːː22] 'dragon'—and across different oral vowels, but they undergo processes before the rounded segments [u] and [w], as discussed in subsection 3.2.2. Therefore, the implosives tend to be unmarked. Third, the occurrence of nasal consonants is strictly constrained in the syllable onset position. They are found only in CV syllable structures and before four nasalised vowels, which are found in only a small number of local vocabularies. In other words, the nasal consonant onsets are used in marked contexts. Considering these three factors, it is appropriate and plausible to propose the implosives as phonemes.

3.2.4. Summary

Therefore, in spite of Zhangzhou having a relatively small number of phonemic onsets, their phonetic realisations are substantially dynamic and diverse, motivated by a variety of processes, including dentalisation, palatalisation, lenition, labialisation, laminalisation, nasalisation, and glottalisation, which are conditioned by different factors. In addition, the processes are not necessarily independent of each other, often interacting to shape the dynamic realisations of Zhangzhou onsets. The auditory description of allophonic variation of syllable onsets sheds some interesting lights on future studies to conduct sophisticated acoustic and articulatory experiments to explore and explain why various processes occur and interact in the utterances of Zhangzhou, and to see whether male and female speakers have different manifestations.

3.3. Glides (G)

Two prevocalic glides—[j] and [w]—can be identified in the data. Nevertheless, three major issues arise with respect to their phonetic properties and theoretical status in syllables.

Phonotactically, the language tends to have an OCP_place effect (obligatory contour principle on the place of articulation) on the occurrence of glide in syllables. Rhymes, like/*wɵ/ and /*wɔ/, are prohibited to occur, while rhymes, like /jɵ/ and /wɐ/, are commonly seen in the data. Phonetically, the realisations of gliding features—[+palatal] and [+labial]—are essentially continuous, and specifying where the glide occurs and ends in the utterances is perceptually difficult. For example, syllable onsets are often undergoing palatalisation before [j], while undergoing labialisation before [w] glide, as discussed in the previous section.

Phonologically, the treatment of glides in syllables has appeared problematic for phonological theories. Many proposals have been suggested but without general agreement (Hooper, 1972; Blevins, 1995; Levi, 2011). For example, glides are grouped conventionally with the nucleus and coda to form a larger layer of rhyme in Sinitic linguistics. They can be grouped with consonant to form a complex syllable onset or grouped with vowel to form a layer of the nucleus. Alternatively, they can be treated as a secondary feature of onset without having a phonemic status in syllables. Nevertheless, different considerations can give rise to different representations of syllable modelling.

In this study, the glide is phonologically treated in the Sinitic fashion as an independent syllable component (G) and as one part of the rhyme. It is also considered a secondary feature of syllable onsets at the phonetic level, as shown in Table 3-14. Thus, the difference between [kj] and [kʲ] can be considered to be an issue of whether the onset [k] is palatalised due to the influence of the production of its following glide. This treatment of Zhangzhou glides can maximally respect the fact that the gliding features are continuous while ensuring the simplicity principle of phonological analysis. Additionally, the diverse and dynamic realisations of syllable onsets can be largely stated as a consequence of various processes conditioned by the continuous gliding features, as discussed in detail in section 3.3.

Table 3-14. Examples of the continuous realisation of glide feature

Phoneme	[j]	[w]
/k/	[kʲjɐ̰41] 'sent'	[kʷwɐ̰41] 'lid'
/kʰ/	[kʰʲjɐːː33] 'stand'	[kʰʷwɐ̰41] 'across'
/ɠ/	[ɠʲjɐːː22] 'raise'	[ɣʷwɐ̰41] 'I'

3.4. Nuclei (V)

This study identified 13 syllable nuclei consisting of seven oral vowels (/i, u, e, ɛ, ɐ, ɵ, ɔ/), four nasalised vowels (/ĩ, ɛ̃, ɐ̃, ɔ̃/), and two syllabic nasals (/m̩, ŋ̍/), as shown in Table 3-15. The nasality feature is used contrastively in the nucleus system. Vowels can be nasalised to distinguish different lexical items while nasal consonants can be used syllabically, as shown in Table 3-16 (in which segments are transcribed phonemically).

Table 3-15. The nucleus inventory of Zhangzhou

Oral vowel	Nasalised vowel	Syllabic nasal
i	ĩ	*
u	*	m̩
e	*	*
ɵ	*	ŋ̍
ɛ	ɛ̃	*
ɐ	ɐ̃	*
ɔ	ɔ̃	*

Table 3-16. Examples of the contrastive nasal feature in Zhangzhou

/V/	/Ṽ/	/m̩/	/ŋ̍/
ti35 'pig'	tĩ35 'sweet'	ʔm̩35 'drink'	tʰŋ̍35 'soup'
pɛ22 'crawl'	pɛ̃22 'flat'	ʔm̩22 'flower bud'	tŋ̍22 'long'
kɐ51 'twist'	kɐ̃51 'dare'	ʔm̩51 'aunt'	sŋ̍51 'play'
kjɐ41 'sent'	kjɐ̃41 'mirror'	ʔm̩41 'affirmative'	kŋ̍41 'steel'
ɗɔ33 'road'	ɗɔ̃33 'two'	ʔm̩33 'negative'	tsŋ̍33 'statement'

Several notable aspects exist with respect to their distribution, realisation and influence on syllable onsets. For example, why are the nasal consonants treated as allophones of voiced occlusives in the onset system but as contrastive phonemes in the nucleus system? This section provides a description and discussion of oral vowels, nasalised vowels, and syllabic nasals.

3.4.1. Oral vowels

Oral vowels have three-way contrasts with respect to either backness or height, as indicated in Table 3-17. Their realisations are also dynamic and diverse in terms of vowel quality and voice quality, which are primarily conditioned by the vowel height but also constrained by tone and the position that they occur in utterances.

Table 3-17. Feature distribution of Zhangzhou oral vowel phonemes

Backedness		Height	
Front	i, e, ɛ	High	i, u
Central	ɐ, ɵ	Mid	e, ɵ
Back	u, ɔ	Low	ɛ, ɐ, ɔ

3.4.1.1 Vowel quality realisation

Vowel quality in Zhangzhou can change with respect to tone and their occurrence position in utterances. The high vowels are observed being realised as diphthongs in the stopped tones 6 and

7 when their associated syllables are produced in isolation or in the final-position of utterances, as summarised in Table 3-18, but as monophthongs elsewhere, as generalised in Table 3-19. The non-high vowels, on the other hand, are realised consistently as monophthongs across tones and across utterance positions. The symbol * in the tables signifies no token was found in the data in the relevant environment.

Table 3-18. Vowel quality realisations across eight tones in the utterance-final position in Zhangzhou

Oral vowel		Tone 1	Tone 2	Tone 3	Tone 4	Tone 5	Tone 6	Tone 7	Tone 8
High	/i/	[i]	[i]	[i]	[i]	[i]	[iɛ̯]	[iɛ̯]	[i]
	/u/	[u]	[u]	[u]	[u]	[u]	[uɤ̯]	[uɤ̯]	[u]
Mid	/e/	[e]	[e]	[e]	[e]	[e]	*	*	[e]
	/ɵ/	[ɵ]	[ɵ]	[ɵ]	[ɵ]	[ɵ]	*	*	[ɵ]
Low	/ɛ/	[ɛ]	[ɛ]	[ɛ]	[ɛ]	[ɛ]	*	*	[ɛ]
	/ɐ/	[ɐ]	[ɐ]	[ɐ]	[ɐ]	[ɐ]	[ɐ]	[ɐ]	[ɐ]
	/ɔ/	[ɔ]	[ɔ]	[ɔ̰]	[ɔ̰]	[ɔ]	[ɔ]	[ɔ]	*

Table 3-19. Vowel quality realisations across eight tones in the non-final position in Zhangzhou

Oral vowel		Tone 1	Tone 2	Tone 3	Tone 4	Tone 5	Tone 6	Tone 7	Tone 8
High	/i/	[i]	[i]	[i]	[i]	[i]	[i]	[i]	[i]
	/u/	[u]	[u]	[u]	[u]	[u]	[u]	[u]	[u]
Mid	/e/	[e]	[e]	[e]	[e]	[e]	*	*	[e]
	/ɵ/	[ɵ]	[ɵ]	[ɵ]	[ɵ]	[ɵ]	*	*	[ɵ]
Low	/ɛ/	[ɛ]	[ɛ]	[ɛ]	[ɛ]	[ɛ]	*	*	[ɛ]
	/ɐ/	[ɐ]	[ɐ]	[ɐ]	[ɐ]	[ɐ]	[ɐ]	[ɐ]	[ɐ]
	/ɔ/	[ɔ]	[ɔ]	[ɔ̰]	[ɔ̰]	[ɔ]	[ɔ]	[ɔ]	*

As Tables 3-18 and 3-19 indicate, the high vowels in Zhangzhou alternate between monophthongs—[i] and [u]—in the unstopped tones and diphthongs—[iɛ̯] and [uɤ̯]—in stopped tones in the utterance-final position, respectively. The high vowel diphthongisation is claimed to be one of the phonetic effects of the non-realisation of obstruent codas at the utterance-final context. In addition, in this thesis, vowel quality is considered one important parameter in shaping the multidimensional realisations of Zhangzhou tones, discussed in the following chapters across three environments—monosyllabic, phrase-initial, and phrase-final—from the perspectives of auditory and acoustic phonetics. The information described here aims to provide a brief generalisation of how vowel quality is realised across tones; however, the description will be presented in detail in related sections concerning the interaction between tone and vowel quality across Chapters 4, 5, 6, 7 and 8, supplemented by examples of phonetic transcriptions, acoustic manifestations, and corresponding articulatory explanations.

Although vowel quality in Zhangzhou is changed in certain circumstances, the oral vowels can significantly impact the realisation of syllable onsets. For example, the [+palatal] feature of the high and front vowel [i] can motivate dentalisation and palatalisation, while the [+labial] feature of the back high vowel [u] can also cause various phonological processes, including labialisation, lenition, laminalisation, and glottalisation, in syllable onsets, as discussed in Section 3.2.

3.4.1.2 Voice quality realisation

The realisation of voice quality is diverse and dynamic in Zhangzhou. The data reveal four different phonations—breathy, creaky, falsetto, and modal—exist for Zhangzhou oral vowels, as described in later chapters. Different types of phonation are not independent of tones to multiply the number of tonal contrasts, as found in many tonal languages in South Asia, but occurring as one important parameter to make multidimensional realisations of Zhangzhou tones.

The realisation of Zhangzhou voice quality is primarily conditioned by vowel quality, pitch and tone as a category. The breathy voice is dominantly found in high vowels in unstopped tones. The creaky voice largely occurs in low vowels with falling pitch contours, but occasionally, the creaky phonation can also be perceived in the falling-pitched mid-vowel [e], especially among male speakers. The falsetto voice only occurs in the stopped tone 6 in the non-final position of utterances, regardless of vowel height. The modal voice, on the other hand, is largely found in mid-vowels across tones and in low vowels but only with non-falling pitch contours.

Table 3-20 summaries the realisation of Zhangzhou voice quality across eight citation tones in the monosyllabic setting. Table 3-21 shows their realisations in non-utterance-final position. The symbol * signifies no token was found in the data in the related environment.

Table 3-20. Voice quality realisation across eight tones in the monosyllabic setting in Zhangzhou

Oral vowel		Tone 1 [35]	Tone 2 [22]	Tone 3 [51]	Tone 4 [41]	Tone 5 [33]	Tone 6 [41]	Tone 7 [221]	Tone 8 [22]
High	/i/	breathy	breathy	breathy	breathy	breathy	creaky	creaky	breathy
	/u/	breathy	breathy	breathy	breathy	breathy	creaky	creaky	breathy
Mid	/e/	modal	modal	modal/ creaky	modal/ creaky	modal	*	*	modal
	/ɵ/	modal	modal	modal	modal	modal	*	*	modal
Low	/ɛ/	modal	modal	creaky	creaky	modal	*	*	modal
	/ɐ/	modal	modal	creaky	creaky	modal	creaky	creaky	modal
	/ɔ/	modal	modal	creaky	creaky	modal	creaky	creaky	*

Table 3-21. Voice quality realisation across eight tones in the non-utterance-final position in Zhangzhou

Oral vowel		Tone 1 [33]	Tone 2 [33]	Tone 3 [35]	Tone 4 [63]	Tone 5 [32]	Tone 6 [65]	Tone 7 [43]	Tone 8 [32]
High	/i/	breathy	breathy	breathy	breathy	breathy	falsetto	breathy	breathy
	/u/	breathy	breathy	breathy	breathy	breathy	falsetto	breathy	breathy
Mid	/e/	modal	modal	modal	modal/ creaky	modal/ creaky	*	*	modal/ creaky
	/ɵ/	modal	modal	modal	modal	modal	*	*	modal
Low	/ɛ/	modal	modal	modal	creaky	creaky	*	*	creaky
	/ɐ/	modal	modal	modal	creaky	creaky	falsetto	creaky	creaky
	/ɔ/	modal	modal	modal	creaky	creaky	falsetto	creaky	*

As Tables 3-20 and 3-21 show, the voice quality realisation in Zhangzhou is complicated and diverse. It can be realised completely differently across different vowel heights, such as in the falling pitched tones. It can also be realised consistently similar across different vowel heights, such as with stopped tone 6 in the non-final position. Alternatively, one vowel can have different phonations across different tonal environments, such as the low vowels having three types of voice quality in the non-final position. The tables are to provide a brief introduction of the complicated but systematic correlations between voice quality, tone, and vowel quality; however, details of phonetic transcriptions, acoustic manifestations, and articulatory explanations will be presented across three different linguistic contexts in Chapters 4, 5, 6, 7, and 8.

3.4.2. Nasalised vowels

Nasalised vowels are used contrastively to distinguish lexical items in Zhangzhou. For example, the morpheme /ti1/ means 'pig' while /tĩ1/ indicates 'sweet'. Nevertheless, not every oral vowel can be nasalised. Only four nasalised vowels—/ĩ, ɛ̃, ɐ̃, ɔ̃/—appear in the nucleus system of Zhangzhou, as shown in Table 3-15 and exemplified in Table 3-16. The existence of nasalised vowels, on the other hand, can trigger several processes on preceding syllable onsets. For example, implosives can assimilate to the nasal feature and become nasal consonants. The pharyngeal fricative /ħ/ is found undergoing glottalisation and being realised as a breathy sound [ɦ] before nasalised vowels, as discussed in Section 3.2.3.

3.4.3. Syllabic nasals

Many tokens are produced with syllabic nasals in Zhangzhou. For example, the syllabic /m̩/ can carry five different pitch contours for five lexical meanings, including /ʔm̩24/ 'drink', /ʔm̩22/ 'flower bud', /ʔm̩51/ 'aunt', /ʔm̩41/ 'affirmative', and /ʔm̩33/ 'negative'. Nevertheless, as far as their distribution is concerned, one crucial issue must be addressed. As reflected in Table 3-15, the two syllabic nasals—[m̩] and [ŋ̩]—occur in contexts where the nasalised vowels [*ũ] and [*ɵ̃] are not found in the data. It thus seems logically plausible to consider the underlying forms of [m̩] and [ŋ̩] as /*ũ/ and /*ɵ̃/, respectively, rather than as /m̩/ and /ŋ̩/. This study posits the syllabic nasals are independent phonemes rather than allophonic variants of nasalised vowels /*ũ/ and /*ə̃/ for two major reasons.

First, the analysis of phonology should be grounded in phonetics (Lass, 1984; Blumstein, 1991; Scobbie, 2007; Cohn, 2007; Ohala, 2010). For example, Lass (1984, p.25) asserted "good descriptions are as simple as is consistent with the facts, as coherent, orderly, highly patterned and structured as possible". In the Zhangzhou data, a considerable number of tokens are produced with syllabic nasals, especially with the velar [ŋ̩]. In contrast, no tokens were found having the nasalised vowels [*ũ] and [*ə̃]. Thus it is appropriate to propose syllabic nasals as the underlying representations which have phonetic reality in utterances, rather than considering the sounds that lack reality in phonetics or in facts as phonemes.

Second, proposing syllabic nasals as phonemes makes the system symmetrical and predictable. Similar to nasalised vowels /ĩ, ɛ̃, ɐ̃, ɔ̃/, the existence of syllabic nasals can also motivate processes on syllable onsets and affect their realisations at the surface. For example, the pharyngeal fricative /ħ/ is found to undergo glottalisation and is realised as a breathy sound [ɦ] before [m̩] and [ŋ̩], as discussed in Section 3.2.3. Thus, considering syllabic nasals as phonemes is not only consistent with the facts, but also allows the pattern to remain coherent, predictable and symmetrical.

3.4.4. Summary

Therefore, the realisations of Zhangzhou oral nuclei are also diverse and dynamic, showing complicated interactions with tone and other phonetic parameters in terms of both vowel quality and voice quality. The nasality feature is contrastively used, forming categories of nasalised vowels and syllabic nasals. Considering the phonemic status of syllabic nasals should address the issues of whether the proposed phonemes have reality in phonetics and whether the treatment can make the system symmetrical and predictable. Additionally, the length realisation of vowels is also observed dynamic, varying with respect to different tones and different linguistic contexts, which will be discussed in later chapters 5, 7 and 8 from the acoustic and statistical perspectives.

3.5. Codas (C)

Three types of syllable coda are identified in this study—glide (/j, w/), nasal (/m, n, ŋ/), and obstruent (/p, t, k/)—which optionally occur in syllables. This section primarily addresses the following four questions.

- Why are postvocalic glides treated as one type of syllable coda (C) rather than as an independent syllable component (G), as are prevocalic glides?
- Why are obstruent codas proposed to be phonemic if they are not always realised?
- Why is the glottal stop considered as a laryngeal feature of rhymes rather than a coda?
- Why are the nasal codas proposed to be contrastive if they seem complementary to the obstruent codas?

3.5.1. Glide codas

Two types of glides—/j/ and /w/—are used contrastively in the postvocalic position in Zhangzhou. They are considered one type of syllable coda (C), patterning with obstruent and nasal consonants. Three major considerations are taken into account in identifying their phonological status within syllables in this study.

First, postvocalic glides are not commonly used in the local vocabulary. They can only occur after vowels /ɐ/ and /ɐ̃/ to form the glide-ending rhymes: /ɐw/, /ɐ̃w/, /ɐj/, /ɐ̃j/, /jɐw/, /jɐ̃w/, /wɐj/, and /wɐ̃j/, as illustrated in Table 3-22. The other nine nuclei, on other hand, are constrained from occurring before the two glides, for examples, rhymes like /*ɛj/, /*ɛw/, /*ej/, /*ew/, /*ɔj/, /*ɔw/, /*ɛ̃j/, and /*ɔ̃j/ were not found in the data.

Second, a good phonological analysis is as simple as is consistent with the facts (Lass, 1984, p. 25). Proposing postvocalic glides as having one independent status in syllables would create an unnecessary degree of complexity with respect to syllable modelling for tokens that occur very little in the whole vocabulary. In contrast, proposing they are one type of syllable coda makes the phonological analysis more simple and economic. It simplifies the generalisation of Zhangzhou syllable structure as C(G)V(C), which logically can have a maximum of four components rather than the five components in the structure of C(G)V(G)(C). Correspondingly, it can reduce the number of syllable types; for example, the types of CVG and CGVG are not necessary in the pattern inventory.

Third, the glide codas group with nasal codas to form a larger category of sonorant codas distinguished from another category of obstruent codas. The classification of syllable codas into sonorant and obstruent has a significant impact on the tonal distinctions in Zhangzhou, as shown in Table 3-23. As discussed, combining the three factors, it is desirable to propose the postvocalic glides /j/ and /w/ as syllable coda in Zhangzhou.

Table 3-22. Examples of morphemes with a postvocalic glide coda

CVC_j	CVC_w	CG_jVC_w	CG_wVC_j
pɐj22 'row'	tɐw35 'home'	tjɐw22 'dynasty'	kwɐj35 'obedient'
sɐj41 'compete'	kɐw51 'dog'	tsjɐw51 'bird'	kwɐj41 'strange'
tsɐ̃j35 'know'	ɗɐ̃w51 'brain'	ɠjɐ̃w35 'itch'	swɐ̃j33 'mango'
ʔɐ̃j33 'carry'	ɠɐ̃w22 'stew'	ɠjɐ̃w51 'scoop'	ʔwɐ̃j51 'sprain'

Table 3-23. Distribution of Zhangzhou tones with respect to syllable codas

Tone	Zero coda	Glide coda	Nasal coda	Obstruent coda
Stopped	-	-	-	+
Unstopped	+	+	+	-

3.5.2. Obstruent codas

Three obstruent coda phonemes—/p, t, k/—are proposed for Zhangzhou. Nevertheless, two important issues needs further justification: (1) Why are the obstruent codas proposed as contrastive phonemes while they are not always realised in the utterances? (2) Why is the glottal stop not considered as a contrastive coda while it is realised in the data?

3.5.2.1 Obstruent codas /p/, /t/, and /k/

Perceptually, identifying the obstruent codas—[p], [t], and [k]—at the ends of utterances is difficult, although occasionally the bilabial coda [p] is perceivable. Two major reasons are given for why they are proposed as contrastive phonemes at the underlying level while they are not always realised in Zhangzhou. First, although obstruent codas tend not to be realised at the utterance-final position, they can give rise to a series of phonetic effects on the whole syllable, including syllable lengthening, high vowel diphthongisation, vowel laryngealisation, and pitch contour depression, which will be discussed in detail from the auditory and acoustic points of view in Chapters 4, 5, 6, and 8.

Second, in contrast to their non-realisation in utterance-final position, they are largely identifiable in the non-final position and cause other sets of phonetic effects on the whole syllable, including syllable shortening and falsetto voice for all vowels in the stopped tone 6. The realisations of obstruent codas and their relations to other phonetic parameters in causing tonal contrasts are discussed in detail from the auditory, acoustic, and articulatory perspectives and supported by statistical testing in Chapters 6 and 7.

Therefore, although the obstruent codas are not realised utterance finally, they are perceivable in the non-final position. Whether they are realised or not, they can cause various phonetic effects on whole syllables. It is thus appropriate to propose the obstruent codas as contrastive phonemes. In addition, while the obstruent codas are realised in the non-final position, their distribution is asymmetrical and, to a certain degree, constrained by vowel quality, as shown in Table 3-24.

Table 3-24. Realisation of obstruent codas at the non-utterance-final position across vowels in Zhangzhou

Vowel	/p/	/t/	/k/
[ɐ]	[p̚]	[t̚]	[k̚]
[i]	[p̚]	[t̚]	[k̚]
[u]	*	[t̚]	*
[ɔ]	[p̚]	*	[ʡ]

As indicated in Table 3-24, obstruent codas are asymmetrically distributed across the rounded vowels [u] and [ɔ]. The labial and velar obstruent codas—[p] and [k]—do not occur after the rounded vowel [u] while the alveolar obstruent coda [t] does not appear after the vowel [ɔ]. Therefore, the obstruent codas seem subject to a segment-sequencing constraint, which can be generalised as OCP_place (obligatory contour principle on the place of articulation). If the adjacent segments share a common place feature, such as the [+bilabial] feature for /*up/ and the [+velar] feature for /*uk/, they are prohibited from occurring simultaneously.

Additionally, the realisation of the velar obstruent coda /k/ as an epiglottal stop [ʡ] after the back vowel [ɔ] is impressionistically salient but also articulatorily understandable. The low back vowel [ɔ] in Zhangzhou creates an impression of a relatively tight constriction in the pharynx, which can also be physically sensed by the tongue movement in my own oral cavity as a native speaker. The lowering jaw and retracted tongue root render a complete constriction that forms in the pharyngeal cavity for the following obstruent coda production, which is reasonably considered as an epiglottal stop [ʡ] in this study.

3.5.2.2 Glottal stop coda

The glottal stop coda [ʔ] is readily perceived in Zhangzhou. Nevertheless, this glottal stop coda is considered a laryngeal feature of rhymes rather than a contrastive syllable coda in this study for the following two reasons. First, the glottal stop is synchronically widespread across both stopped and unstopped tones, rather than being limited to the specific tones 6 and 7 as the obstruent stop coda. There are two major environments in which the glottal stop can be perceived: (a) all low vowels, occasionally including the mid-vowel [e], in the falling tonal environments, for example, tones 3, 4, 6, and 7 in citation, and (b) all vowels that can occur in stopped tones 6 and 7 in the monosyllabic setting.

Additionally, as far as the common feature shared by the segments in the two main environments is concerned, the glottal stop coda essentially can occur anywhere if the vowels are produced with a creaky voice, regardless of in the stopped or unstopped tones. As shown in Table 3-25, it is common for the low vowels to occur with a creaky voice across tones.

Table 3-25. Examples of the creaky low vowels in falling tones

Tone	Pitch	/ɛ/	/ɐ/	/ɔ/
1	low rising	[kɛ::35] ‘home’	[kɐ::35] ‘glue’	[kɔ::35] ‘mushroom’
2	mid-low level	[kʰɛ::22] ‘get stuck’	[tsʰɐ::22] ‘firework’	[kɔ::22] ‘paste’
3	high falling	[kʰɛ̰51] ‘arise’	[kɐ̰51] ‘grind’	[kɔ̰51] ‘drum’
4	mid-high falling	[kɛ̰41] ‘shelf’	[kɐ̰41] ‘teach’	[kɔ̰41] ‘look after’
5	mid level	[kɛ::33] ‘low’	[kɐ::33] ‘bite’	[tɔ::33] ‘degree’
6	stopped mid-high falling	*	[kɐ̰̆41] ‘horn’	[kɔ̰̆41] ‘country’
7	stopped mid-low level	*	[kɐ̰:221] ‘throw away’	[kɔ̰:221] ‘cook’
8	mid-low level	[pɛ::22] ‘white’	[tɐ::22] ‘step on’	*

Second, most of the syllables that are diachronically assumed to have a phonemic glottal stop coda (Lin, 1992; Ma, 1994; FCCEC, 1998; ZCCEC, 1999; Gao, 1999; Zhou, 2006; Yang, 2008; Guo, 2014) are effectively produced without a glottal stop coda in spontaneous data. Previous studies used the glottal stop coda phonemically before different types of nuclei but only for the stopped tones of Yinru and Yangru in terms of the Middle Chinese (MC) tonal categories, in which tones are categorised into the four classes of Ping, Shang, Qu, and Ru and into the two registers of Yin and Yang for each tonal class. Nevertheless, the data for this study reveal those syllables are largely produced without a glottal stop at the end of utterances in MC Yinru and Yangru tones. If syllables have a phonetic glottal stop, they should have low vowels and falling pitch contours, as shown in Table 3-26.

Table 3-26. Examples of the synchronic disappearance of the glottal stop coda in Zhangzhou

Ma (1994)	**This study**	MC YinRu (tone 4)	MC Yangru (tone 8)
/iʔ/	**/i/**	[t̪ʰi̤41] ‘iron’	[tɕi̤ːː22] ‘tongue’
/uʔ/	**/u/**	[tʃṳ41] ‘squeeze’	[t̪ʷṳːː22] ‘piece’
/oʔ/	**/ɵ/**	[tɵ41] ‘table’	[tɵːː22] ‘burn’
/eʔ/	**/e/**	[pḛ41] ‘eight’	[ʔeːː22] ‘narrow’
/aʔ/	**/ɐ/**	[ɓɐ̰41] ‘meat’	[tɐːː22] ‘step on’
/ɔʔ/	*****	*	*
/ɛʔ/	**/ɛ/**	[tsʰɛ̰41] ‘book’	[pɛːː22] ‘white’
/ĩʔ/	**/ĩ/**	[sĩ41] ‘flash’	[mĩːː22] ‘stuff’
/ɔ̃ʔ/	**/ɔ̃/**	[mɔ̃41] ‘sneak’	[kɔ̃ːː22] ‘snore’
/ãʔ/	**/ɐ̃/**	[sɐ̃41] ‘shinning’	*
/ɛ̃ʔ/	**/ɛ̃/**	[ɦɛ̃41] ‘rest’	[mɛ̃ːː22] ‘pulse’
/m̩ʔ/	**/m̩/**	[ɦm̩41] ‘hit with a stick’	[ɦm̩ːː22] ‘silent’
/ŋ̩ʔ/	**/ŋ̩/**	*	[sŋ̩ːː22] ‘sound of wheeling’
/iaʔ/	**/jɐ/**	[t̪ʰjɐ̰41] ‘dismantle’	[tɕjɐːː22] ‘eat’
/iãʔ/	**/jɐ̃/**	[ɦjɐ̃41] ‘fetch (cloth)’	*
/ioʔ/	**/jɵ/**	[tɕjɵ41] ‘borrow’	[tɕjɵːː22] ‘stone’
/iɔʔ/	**/jɔ/**	*	[ħʲjɔːː22] ‘affirmative’
/uaʔ/	**/wɐ/**	[βʷwɐ̰41] ‘smudge’	[ʔʷwɐːː22] ‘alive’
/uãʔ/	*****	*	*
/ueʔ/	**/we/**	[hʷwe41] ‘blood’	[ɣʷweːː22] ‘moon’
/auʔ/	**/ɐw/**	[pʰɐw41] ‘baby fat’	[ɗɐwːː22] ‘drop’
/ãuʔ/	**/ɐ̃w/**	[mɐ̃w41] ‘lisping’	[ɦɐ̃wːː22] ‘mealy’
/iauʔ/	**/jɐw/**	[ħʲjɐw41] ‘fall off’	[kʲjawːː22] ‘stiff status’
/iãuʔ/	**/jɐ̃w/**	[ŋʲjɐw41] ‘mean’	[ŋʲjɐwːː22] ‘wriggle’
/uãiʔ/	**/wɐ̃j/**	[ʔʷwɐj42] ‘teeter’	[hʷwɐːː22] ‘fat status’

In Table 3-26, the leftmost column contains the 25 glottal-ending rhymes summarised by Ma (1994). The second column shows their corresponding phonemic representations proposed in this study, with examples in both MC Yinru (present tone 4) and Yangru (present tone 8) tones, which are transcribed phonetically according to their auditory impression. As indicated in this table, no creaky-voiced syllables are found in MC Yangru tone in this study, but their rhymes were documented as ending in the glottal stop in Ma (1994). In contrast, some morphemes in MC Yinru tone, which has a mid-high falling pitch contour, are found being produced with a creaky voice, but they are limited to low vowels and the mid-vowel [e]. The high vowels are realised with a breathy voice while the mid-vowel [ɵ] is largely produced in modal phonation. In other words, the laryngeal feature is preditable to associate with low vowels and falling pitch contour.

Therefore, it is preditable, plausible and economic to propose the glottal stop as a laryngeal feature accompanying the low vowels in the falling pitch contexts and all vowels in the stopped tones. In contrast, if the glottal stop coda is represented as a phonemic coda, a considerable number of rhymes ending in the glottal stop are indicated, creating a substantial degree of redundancy in rhyme table and dictionary production in Zhangzhou. For example, rhymes ending in the glottal stop occupied more than 28% of the rhyme inventory in previous studies (Ma, 1994; FCCEC, 1998; Gao, 1999; Guo, 2014).

Additionally, syllables transcribed in previous studies in MC Yinru with a glottal stop coda actually have been merged into those of tone 4 (MC Yinqu) with respect to vowel quality, voice quality, length, and pitch because they do not present the properties of those syllables in stopped tone 6 (MC Yinru), such as vowel laryngealisation and high vowel diphthongisation. The syllables that were transcribed in Yangru with a glottal stop coda in previous studies effectively form a new tonal category of tone 8 in this study because they have different realisations from the syllables in stopped tone 7, which end in obstruent codas (see later chapters). Therefore, there is no glottal coda phoneme in the synchronic system. The historical Yinru with glottal stop is merged with Yinqu (tone 4), while the historical Yangru with glottal stop is realised without a glottal stop and forms a separate tone category (tone 8, Yangru). The laryngealisation is thus a phonetic effect in present day Zhangzhou and should not be equated with the historical glottal stop.

3.5.3. Nasal codas

Three nasal codas—[m], [n], and [ŋ]—are perceived in Zhangzhou data, as examplied in Table 3-27. Nevertheless, they are arguably complementary to their homorganic oral counterparts—[p], [t], and [k]—conditioned by tones: the unstopped tones for the nasal codas and the stopped tones for the obstruent codas, as shown in Table 3-23. It thus seems logically plausible to consider the nasal codas as allophonic variants of obstruent codas or to treat the obstruent codas as allophonic variants of nasal codas.

This study considers the nasal codas are used phonemically, independent of the obstruent codas, for three reasons. First, a large number of tokens end in nasal codas in the local vocabularies. Second, nasal codas, along with glide codas, form a sonorant-ending category, different from the obstruent-ending category, to classify tones and syllables. Third, while the occurrence of obstruent codas is strictly limited to stopped tones 6 and 7, whether they are realised or not in utterances, they can always create a series of phonetic effects on whole syllables, as discussed in Section 3.5.2. Combining these three factors, it is phonologically plausible and simple to consider the nasal codas as contrastive phonemes.

Table 3-27. Examples of nasal coda contrasts in Zhangzhou

Coda	[i]	[ɐ]
/m/	kim35 ‘gold’	tɐm22 ‘discuss’
/n/	kin35 ‘half a kilogram’	tɐn22 ‘Chen, family name’
/ŋ/	kiŋ35 ‘through’	tɐŋ22 ‘copper’

3.5.4. Summary

Phonological analysis of segments should not be judged simply on the basis of their distributions in utterances, for example, whether they are contrastive or complementary to other segments. Taking into account the phonetic facts, for instance, the rate of usage in local vocabularies and their correlations with surrounding elements, is crucial for developping a simple but also plausible phonological representation.

3.6. Previous Studies

Zhangzhou has been documented in various studies since the late 1950s (Dong, 1959; Lin, 1992; Ma, 1994; FCCEC, 1998; ZCCEC, 1999; Gao, 1999; Zhou, 2006; Yang, 2008; Huang, 2009, 2010; Guo, 2014), but those studies were largely impressionistic and heavily informed by the highly conventionalised expectations of Mainland Chinese linguistics. Consequently, the previous studies largely lacked phonetic detail and provided only broad transcriptions of both segments and suprasegmentals. That approach undermined the objectivity of those reported results and left readers unable to assess the reasonableness of the phonological analyses (both overt and implied).

In addition, both agreements and disagreements occur between the results of this study and the nearly unanimous results of published studies since 1959, as illustrated in Tables 3-28 and 3-29.

Table 3-28. Comparisons of Zhangzhou syllable onset inventory in different studies

Place	Dong 1959	Lin 1992	Ma 1994	FCCEC 1998	ZCCEC 1999	Gao 1999	Zhou 2006	Yang 2008	Guo 2014	**This study**
Labial	p	p	p	p	p	p	p	p	p	**p**
	p^h	p^h	p^h	p^h	p^h	p^h	p^h	p^h	p^h	**p^h**
	b	b	b	b	b	b	b	b	b	**ɓ**
Alveolar	t	t	t	t	t	t	t	t	t	**t**
	t^h	t^h	t^h	t^h	t^h	t^h	t^h	t^h	t^h	**t^h**
	l	l	l	l	l	l	l	l	l	**ɗ**
	ts	ts	ts	ts	ts	ts	ts	ts	ts	**ts**
	ts^h	ts^h	ts^h	ts^h	ts^h	ts^h	ts^h	ts^h	ts^h	**ts^h**
	s	s	s	s	s	s	s	s	s	**s**
	dz	dz	dz	dz	dz	dz	dz	dz	dʒ	**z**
Velar	k	k	k	k	k	k	k	k	k	**k**
	k^h	k^h	k^h	k^h	k^h	k^h	k^h	k^h	k^h	**k^h**
	g	g	g	g	g	g	g	g	g	**ɠ**
Glottal	h	h	h	h	h	h	h	h	h	**ɦ**
Zero	∅	∅	∅	∅	∅	∅	∅	∅	∅	**ʔ**

Table 3-29. Comparisons of Zhangzhou nucleus inventory in different studies

Dong 1959	Lin 1992	Ma 1994	FCCEC 1998	ZCCEC 1999	Gao 1999	Zhou 2006	Yang 2008	Huang 2010	Guo 2014	**This study**
i	i	i	i	i	i	i	i	i	i	**i**
e	e	e	e	e	e	e	e	e	e	**e**
ɛ	ɛ	ɛ	ɛ	ɛ	ɛ	ɛ	ɛ	ɛ	ɛ	**ɛ**
a	a	a	a	a	a	a	a	a	a	**ɐ**
ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	**ɔ**
o	o	o	o	o	o	o	o	o	o	**ɵ**
u	u	u	u	u	u	u	u	u	u	**u**
ĩ	ĩ	ĩ	ĩ	ĩ	ĩ	ĩ	ĩ	ĩ	ĩ	**ĩ**
ẽ	ẽ	*	*	*	*	*	*	*	*	*****
*	ɛ̃	ɛ̃	ɛ̃	ɛ̃	ɛ̃	ɛ̃	ɛ̃	ɛ̃	ɛ̃	**ɛ̃**
ã	ã	ã	ã	ã	ã	ã	ã	ã	ã	**ɐ̃**
ɔ̃	ɔ̃	ɔ̃	ɔ̃	ɔ̃	ɔ̃	ɔ̃	ɔ̃	ɔ̃	ɔ̃	**ɔ̃**
*	ũ	*	*	*	*	*	*	*	*	*****
m̩	m̩	m̩	m̩	m̩	m̩	m̩	m̩	m̩	m̩	**m̩**
*	*	*	*	*	n̩	*	*	*	*	*****
ŋ̩	ŋ̩	ŋ̩	ŋ̩	ŋ̩	ŋ̩	ŋ̩	ŋ̩	ŋ̩	ŋ̩	**ŋ̩**

The differences relate largely to the fact that those studies placed less importance on phonetic details and were satisfied with producing a convenient set of labels for what were assumed to be phonemes in the received structuralist perspective (such as codified by Pike, 1947). For example, in this current study, many of the syllable onsets, even in relatively unmarked environments, had characteristics not adequately represented by the symbols used in earlier studies, nor were these details adequately described in those works. In particular,

- Most scholars proposed /dz/ while Guo (2014) represented this phoneme as /dʒ/. This segment is effectively identified as a voiced fricative /z/ with some conditioned allophony (e.g., [z], [ʒ], and [ʝ]) in this study. As shown in Figure 3-1, no voicing period is initiated before vowel production and no voiced bar is generated in the spectrogram, illustrating that no voiced stop is produced before the fricative onset /z/. Thus, this sound is apparently not an affricative, but rather a fricative.
- No unambiguously egressive voiced occlusives are produced in the data ([*b], [*d], [*g]), contradicting the findings of prior studies. Rather, voiced implosives [ɓ], [ɗ], and [ɠ] are typically observed in unmarked environments. As shown in Figure 3-2, the voicing is initiated before vowel production and the amplitude remains relatively stable or slightly increases as a function of time, signifying the production of an implosive. Additionally, previous studies also consistently indicated a lateral /l/ for the onset that, in the present data, is predominantly pronounced as an implosive stop [ɗ], and in a pattern with labial and velar implosives.
- A glottal fricative [h^w] is found only in the environment of a preceding high, central, rounded vowel [u] or the prevocalic glide [w]; in most other cases, it is produced as a pharyngeal fricative [ħ]. Additionally, it can be realised as a breathy [ɦ] before nasalised vowels or syllabic nasals. As shown in Figure 3-3, prominent acoustic energy peaks at lower resonant frequencies, indicating a long front cavity of vocal tract and a posterior constriction. The spectrogram of energy peaks is distinct and dominant rather than being a cross between a fricative and a vowel as for glottal fricative [h], because of the coupling of the front and back cavity. Thus, it is plausible to propose the pharyngeal glottal as a phoneme.
- Previous studies consistently incorporated a zero syllable onset /∅/ into the inventory, yet the present data indicate that [ʔ] onset is frequently articulated, with no indication that it contrasts with a zero onset (either segmentally or in any suprasegmental context). Given the otherwise strong tendency for Zhangzhou syllables to have onsets, it is preferable to recognise all syllables as having an obligatory onset that, in the case of /ʔ/, can be omitted without interference.

Why the studies examined consistently used some symbols indicating values divergent from the phonetic values found associated with the real unmarked articulations of these onsets is not clear. One possible reason could be substantial deference among Chinese scholars toward earlier studies, such that they were not effectively independent works. Thus, it would not be appropriate to regard the sheer number of unanimous results as deserving special weight.

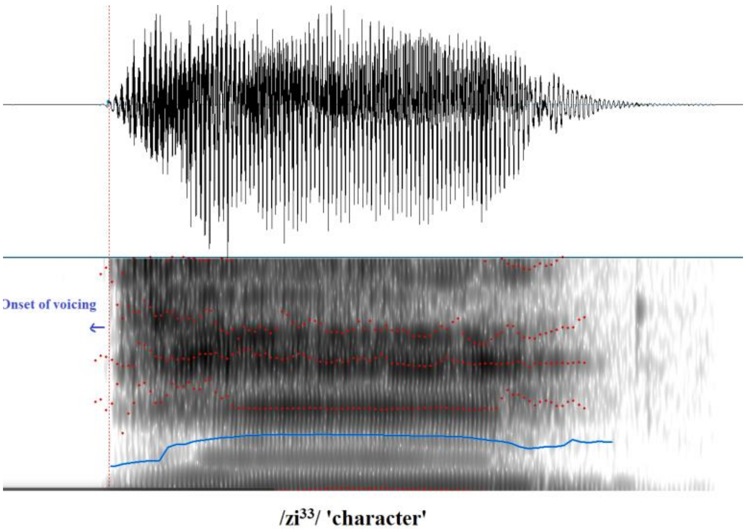


Figure 3-1. Acoustic spectrogram of voiced fricative onset /z/ in Zhangzhou (WYF, male).

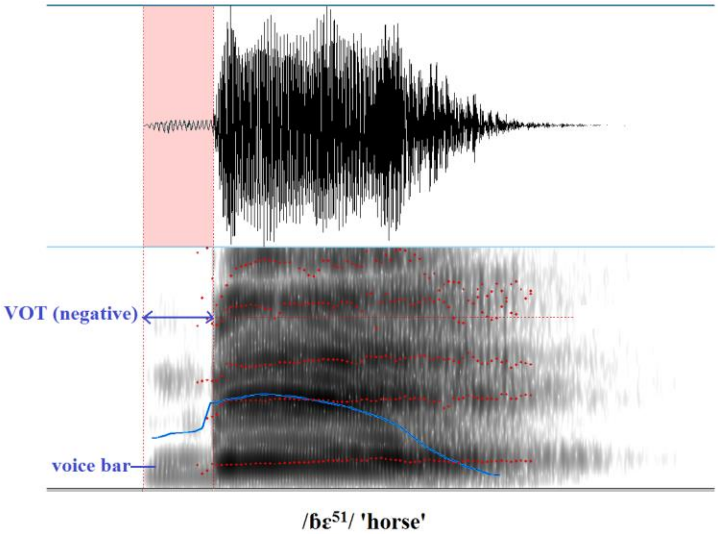


Figure 3-2. Acoustic spectrogram of bilabial implosive /ɓ/ in Zhangzhou (WYF, male).

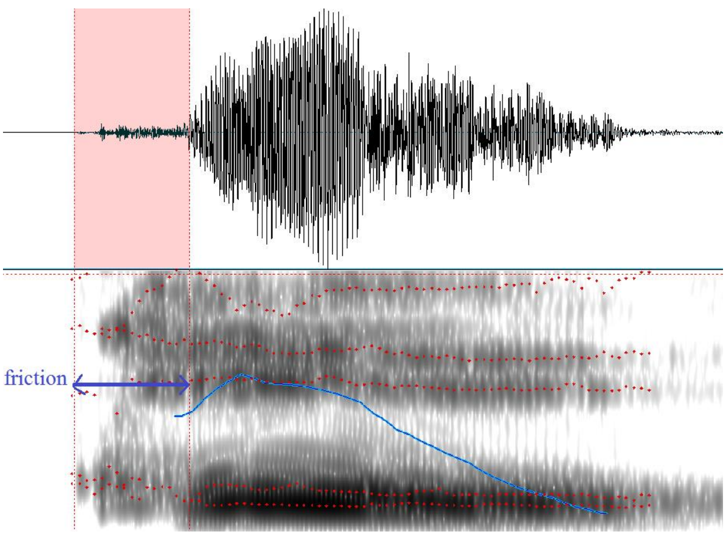


Figure 3-3. Acoustic spectrogramof pharyngeal fricative /ħ/ in Zhangzhou (WYF, male).

3.7. Conclusion

This chapter has provided a detailed description of Zhangzhou syllables and segments from the perspectives of both phonetics and phonology. In spite of Zhangzhou having a relatively small number of segmental contrasts to constitute different syllable components, their phonological status and phonetic realisations vary somewhat. The descriptions in this study advance understanding of Zhangzhou speech sounds in terms of universal and language-specific principles that shape the diverse and dynamic properties of human speech. This chapter has also provided a foundation for the discussion in the remaining chapters concerning how segmentals and suprasegmentals function together to create Zhangzhou tonal realisations and distinctions.

Part B: Zhangzhou Citation Tones

Chapter 4: Auditory Properties of Zhangzhou Citation Tones

This thesis asserts Zhangzhou has eight citation tones, although to fully appreciate this finding needs to understand their multidimensional characteristics across different linguistic contexts. This is because the single dimension of pitch in the single context of citation is not adequate to characterise Zhangzhou tonal contrasts. This chapter presents an auditory description of Zhangzhou monosyllabic citation tones. It mainly addresses two research questions: (1) What are the eight tones realised with respect to pitch in monosyllables? (2) How do segmental and suprasegmental parameters interact to make tonal distinctions in Zhangzhou?

4.1. Pitch

A considerable amount of variation can be observed in terms of the pitch realisation of individual citation tones among the 21 speakers recorded for this study, although tendencies can be generalised as well. The impressionistic descriptive result is summarised in Table 4-1, using the Chao pitch notation system.

The following section describes the pitch realisation of individual Zhangzhou citation tones abstracted broadly from the data. Each tone is given its corresponding name in Middle Chinese (MC) tonal categories; for example, tone 1 corresponds to MC Yinping tone, making this study comparable, both synchronically and diachronically, with previous studies of Zhangzhou and with studies of other Sinitic dialects and their varieties that generally categorise synchronic tones according to the MC tonal categories. To highlight the pitch distinction, segments of various examples are shown in their phonemic transcriptions, but their complicated realisations and interactions with pitch are discussed in fuller detail in Section 4.2.

4.1.1. Tone 1–Yinping [35]

Tone 1 has a rising contour, but speakers differ somehow in terms of pitch onset and offset height. Four variations can be perceived from the empirical data, including [35] (CSM, HJH, HQM, LRX, LRZ, ZBQ, HZB, LMY, and YBR); [34] (YML, HMC, HYS, and WJG); [24] (HHX, HTR, and LYY), and [25] (HSX, HYH, LCH, LZP, and WYF). Because more speakers have the onset realisation at 3 and offset at 5, this tone is labelled [35] in this study. Examples of morphemes of this tone are shown in Table 4-2. Thus, this tone is apparently neither a mid-high level [44] pitch, as described previously by most scholars (Lin, 1992; Ma, 1994; FCCEC, 1998; ZCCEC, 1999; Zhou, 2006; Chen, 2007; Yang, 2008; Guo, 2014) nor a high-rise contour of [45] as Gao (1999) described.

Table 4-1. Pitch realisation of Zhangzhou citation tones among 21 speakers

Gender	Speaker	Tone 1	Tone 2	Tone 3	Tone 4	Tone 5	Tone 6	Tone 7	Tone 8
Female	CSM	35	22	52	41	33	41	21	22
	HHX	24	211	51	41	33	41	211	211
	HJH	35	22	51	41	33	42	221	22
	HQM	35	211	51	41	33	41	211	211
	HSX	25	22	51	41	33	41	221	22
	HYH	25	212	51	41	33	41	212	212
	LCH	25	22	51	41	33	41	21	22
	LRX	35	22	52	41	33	41	221	22
	LRZ	35	22	51	41	33	41	22	22
	LZP	25	22	51	41	33	43	21	22
	YML	34	22	51	41	33	42	221	22
	ZBQ	35	22	51	41	33	41	221	22
Male	HMC	34	22	51	41	33	41	21	22
	HTR	24	211	51	41	33	42	211	211
	HYS	34	22	51	41	33	42	22	22
	HZB	35	22	51	41	44	42	22	22
	LMY	35	22	51	41	33	43	22	22
	LYY	24	211	51	41	33	42	21	211
	WJG	34	211	51	41	33	42	211	211
	WYF	25	22	51	31	33	41	221	221
	YBR	35	22	51	31	33	42	22	22

Table 4-2. Examples of morphemes of tone 1 (Yinping) in Zhangzhou

CV	CGV	CVC	CGVC
si35 ‘poetry’	swẽ35’mountain’	sim35 ‘heart’	sjɐn35 ‘immortal’
t^{h}ĩ35 ‘sky’	t^{h}jẽ35 ‘hall’	t^{h}ɐŋ35 ‘window’	ts^{h}jɔŋ35 ‘rush’
kɔ35 ‘mushroom’	kjɔ̃35 ‘ginger’	kɔŋ35 ‘grandfather’	kjɐw35 ‘delicate’
ɗɐ35 ‘pull’	ɓwi35 ‘smile’	ħɐj35 ‘big’	ʛjɐn35 ‘ring (v.)’
ʔm̩35 ‘drink’	ʔwi35 ‘power’	ʔɐw35 ‘cup’	ʔjɐw35 ‘hungry’

4.1.2. Tone 2–Yangping [22]

Tone 2 is dominantly realised as a mid-low level contour [22], but variants with a low fall with level plateau [211] (HRX, HQM, HTR, LYY and WJG) and a low dipping [212] (HYH) can also be perceived. It is not realised as a rising contour of [13] or [12] as described previously (Lin, 1992; FCCEC, 1998; ZCCEC, 1999; Gao, 1999; Ma, 1999; Zhou, 2006; Chen, 2007; Yang, 2008; Guo, 2014). This tone is broadly transcribed as [22] in this study. Examples are shown in Table 4-3.

Table 4-3. Examples of morphemes of tone 2 (Yangping) in Zhangzhou

CV	CGV	CVC	CGVC
si22 ‘time’	swi22 ‘follow’	sin22 ‘deity’	sjɐw22 ‘silly’
pɛ̃22 ‘flat’	tjɐ̃22 ‘yard’	tɐj22 ‘platform’	tswɐn22 ‘entire’
kɔ22 ‘paste’	kju22 ‘ball’	kɔŋ22 ‘in rush’	kwɐn22 ‘high’
ɓɐ22 ‘palsy’	ɓwɐ̃22 ‘sesame’	ɗɐŋ22 ‘people’	ʄjɐm22 ‘sticky’
ʔm̩22 ‘bud’	ʔjɔ̃22 ‘sheep’	ʔɐw22 ‘throat’	ʔjɐm22 ‘salt’

4.1.3. Tone 3–Shang [51]

Tone 3 is a high falling tone. Its pitch falls through most speakers’ ranges from the top to the bottom [51], but some speakers have a slightly higher offset [52] (CSM and LRX). The observation is similar to previous descriptions of [53] (Dong, 1959; Lin, 1992; Ma, 1994; FCCEC, 1998; ZCCEC, 1999; Gao, 1999; Zhou, 2006; Chen, 2007; Yang, 2008; Guo, 2014) but with a lower pitch offset. This tone is transcribed as [51] in this study. Examples of this tone are shown in Table 4-4.

Table 4-4. Examples of morphemes of tone 3 (Shang) in Zhangzhou

CV	CGV	CVC	CGVC
si51 ‘die’	swi51 ‘beautiful’	sim51 ‘judge’	swɐn51 ‘choose’
tsɛ̃51 ‘wells’	tjɐ̃51 ‘fried pan’	ts^hɐw51 ‘glass’	tsjɐw51 ‘bird’
kɔ51 ‘drum’	kwe51 ‘rice cake’	kɔŋ51 ‘speak’	kwɐn51 ‘govern’
pɐ51 ‘full’	ʠwɐ51 ‘I’	ɗɐŋ51 ‘we’	ɓjɐn51 ‘exempt’
ʔm̩51 ‘aunt’	ʔjɔ̃51 ‘ladle out’	ɦɐj51 ‘sea’	ɦjɐm51 ‘danger’

4.1.4. Tone 4–Yinqu [41]

Tone 4 has a falling contour. It largely falls from the mid-third of the speaker’s pitch range to low and occasionally falls from the mid-pitch range for some speakers (WYF and YBR). It is apparently not a low-falling contour [21] as most scholars described previously (Lin, 1992; Ma, 1994; FCCEC, 1998; ZCCEC, 1999; Gao, 1999; Zhou, 2006; Chen, 2007; Yang, 2008; Guo, 2014) or [32] (Dong, 1959). This tone is transcribed as [41] in this study. Examples of morphemes of this tone appear in Table 4-5.

Table 4-5. Examples of morphemes of tone 4 (Yinqu) in Zhangzhou

CV	CGV	CVC	CGVC
si41 ‘four’	swɐ̃41 ‘thread’	sin41 ‘believe’	sjɐw41 ‘accounts’
sɛ̃41 ‘surname’	tʰjɐ̃41 ‘painful’	tun41 ‘shield’	tjɐm41 ‘store’
kɔ41 ‘look after’	kju41 ‘rescue’	kɔŋ41 ‘hit’	kwɐj41 ‘blame’
ɗu41 ‘scrub’	kwɐ41 ‘hang’	ʛɐn41 ‘chilly’	ɗjɔŋ41 ‘pedal’
ʔm̩41 ‘yes’	ʔjɵ41 ‘guess’	ʔɐj41 ‘love’	ʔjɐm41 ‘hate’

4.1.5. Tone 5–Yangqu [33]

Tone 5 is realised as a mid-level tone among the speakers but as a mid-high level [44] for the male speaker HZB. Perceptually, it has a higher pitch height than a mid-low level [22], as most previous studies have indicated (Lin, 1992; Ma, 1994; FCCEC, 1998; ZCCEC, 1999; Zhou, 2006; Chen, 2007; Yang, 2008; Guo, 2014). This tone is transcribed as [33], which is identical to the descriptions of Dong (1959) and Gao (1999). Examples are shown in Table 4-6.

Table 4-6. Examples of morphemes of tone 5 (Yangqu) in Zhangzhou

CV	CGV	CVC	CGVC
si33 ‘affirmative’	sjɐ33 ‘shoot’	sin33 ‘remain’	sjɐw33 ‘whip’
pɛ̃33 ‘illness’	tjɐ̃33 ‘stop’	tɐj33 ‘replace’	tjɐn33 ‘electricity’
ħɔ33 ‘rain’	sjɔ̃33 ‘miss’	tɔŋ33 ‘hole’	kwɐn33 ‘county’
ɗɔ̃33 ‘road’	ʛwɐ33 ‘outside’	ɗun33 ‘tender’	ɗjɐn33 ‘practice’
ʔm̩33 ‘no’	ʔwɐ33 ‘speech’	ʔɐw33 ‘rear’	ʔjɔŋ33 ‘use’

4.1.6. Tone 6–Yinru [41]

Tone 6 is perceptually realised as a mid-high falling contour [41] for most speakers but has variations in the pitch offset, including [42] (CSM, YML, HTR, HYS, LYY, WJG, and YBR) and [43] (LZP and LMY). It has a slightly higher pitch onset than that of tone 4 but has a shorter duration. Syllables in this tone have an obstruent coda at the underlying level, but the codas tend to be realised only in non-utterance-final position. This tone is transcribed as [41] and is labelled as a stopped mid-high falling tone to signify its similarity to (e.g., pitch) and difference from (e.g., duration and syllable coda types) the unstopped tone 4. The transcription is different from previous descriptions of [32] (Dong, 1959; Lin, 1992; Ma, 1994; FCCEC, 1998; ZCCEC, 1999; Zhou, 2006; Chen, 2007; Yang, 2008), [21] (Gao, 1999), and [31] (Guo, 2014). Examples of this tone appear in Table 4-7.

Table 4-7. Examples of morphemes of tone 6 (Yinru) in Zhangzhou

CVC	CGVC
sip41 'humid'	sjɐp41 'astringent'
tik41 'bamboo'	tsʰjɐt41 'chop'
kɔk41 'country'	kjɵk41 'chrysanthemum'
ɗut41 'fall off'	ħwɐt41 'law'
ʔɐk41 'water (v.)'	ʔwɐt41 'turn (v.)'

4.1.7. Tone 7–Yangru [221]

Tone 7 has a similar pitch contour to tone 2 but with a slight final fall because of the association between creaky phonation and pitch depression. Its pitch realisation varies from [221] (HJH, HSX, LRX, YML, ZBQ, and WYF), [211] (HHX, HQM, and WJG), [22] (LRZ, HYS, HZB, LMY, and YBR), [21] (CSM, LCH, LZP, HMC, and LYY), and [212] (HYH). It is labelled as [221] in this study. Syllables of this [221] pitch have underlying obstruent codas that tend not to be realised in citation. Examples of this tone are shown in Table 4-8. These observations are different from previous studies that largely transcribed this tone with a convex contour [121] (Ma, 1994; FCCEC, 1998; ZCCEC, 1999; Gao, 1999; Zhou, 2006; Chen, 2007; Yang, 2008; Guo, 2014) or a low rise [13] (Dong, 1959) or [12] (Lin, 1992).

Table 4-8. Examples of morphemes of tone 7 (Yangru) in Zhangzhou

CVC	CGVC
tsɐp221 'ten'	kjɵk221 'bureau'
tɔk221 'poison	kjɐk221 'drama'
ɗik221 'green'	ɓjɐt221 'destroy'
ɓɐk221 'eye'	ɠjɔk221 'jade'
ħut221 'Buddha'	ʔjɐt221 'fan (v.)'

4.1.8. Tone 8-Yangru [22]

Tone 8 is largely realised as a mid-low level contour [22], as tone 2, but similarly, several variants can be observed, including [211] (HHX, HQM, HTR, LYY, and WJG), [221] (WYF), and [212] (HYH). This tone is labelled as Yangru [22] in this study. Examples are provided below in Table 4-9.

The reason for proposing tone 8 as an independent tone rather than classifying it as an allotone of Tone 2 is primarily because of its different realisations in sandhi environments (see Chapters 6 and 7). In addition, this tone is classified as Yangru in MC tonal categories because syllables of this tone have been consistently classified in the MC Yangru tone with a glottal coda /ʔ/ (Dong, 1959; Lin, 1992; Ma, 1994; FCCEC, 1998; ZCCEC, 1999; Gao, 1999; Zhou, 2006). To keep diachronic track of syllables in this tone, tone 8 is labelled as Yangru in this study, although no glottal stop is found in related syllables, as discussed in Chapter 3.

Table 4-9. Examples of morphemes of tone 8 (Yangru) in Zhangzhou

CV	CGV	CVC	CGVC
pɛ22 'white'	tjɵ22 'right'	ɦɐw22 'mealy'	k^hjɐw22 'stiff status'
tsi22 'tongue'	tsjɐ22 'eat'	ɓɐw22 'buy up'	ŋjɐ̃w22 'wriggle'
ɗɐ22 'wax'	ɗwɐ22 'spicy'	ɗɐw22 'drop'	swɐ̃j22 'sound of moving fan'
ʔe22 'narrow'	ʔwɐ22 'alive'	*	ɦwɐ̃j22 'fat status'

4.1.9. Summary

As noted previously, the pitch realisation of Zhangzhou citation tones involves variation in both pitch contour—rising, level, and falling—and pitch height—mid-low, mid, mid-high, and high. The overall pitch system of the eight citation tones is summarised in Table 4-10, with examples of minimal and/or semi-minimal pairs. In addition, several notable aspects deserve further mentioning with respect to the relations among different citation tones in terms of pitch.

Table 4-10. Pitch system of Zhangzhou citation tones

Tone	MC	Pitch	Morpheme	Morpheme
1	Yinping	mid rising [35]	/si/ 'poetry'	/kɔ/ 'mushroom'
2	Yangping	mid-low level [22]	/si/ 'time'	/kɔ/ 'glue'
3	Shang	high falling [51]	/si/ 'die'	/kɔ/ 'drum'
4	Yinqu	mid-high falling [41]	/si/ 'four'	/kɔ/ 'look after'
5	Yangqu	mid level [33]	/si/ 'affirmative'	/ɦɔ/ 'rain'
6	Yinru	stopped mid-high falling [41]	/sik/ 'colour'	/kɔk/ 'country'
7	Yangru	stopped mid-low level [221]	/sit/ 'solid'	/tɔk/ 'poison'
8	Yangru	mid-low level [22]	/tsi/ 'tongue'	/kɔ̃/ 'snore'

First, tones of similar pitch contours are not separated maximally from each other in terms of pitch height. For example, the level tones—2 and 5—are not in a high–low distinction (e.g., [55] as opposed to [22] or [44] as opposed to [11]) but rather in a mid–mid-low contrast ([33] as opposed to [22]). Similarly, the falling tones—3 and 4—are not in a high–low distinction (e.g., [51] as opposed to [21]) but rather in a high–mid-high contrast ([51] as opposed to [41]). Therefore, the question is whether these can be regarded as high-low contrasts phonologically.

Second, different tones can share very similar pitch realisations in the monosyllabic setting. For example, tones 4 and 6 share the very similar mid-high falling contour in citation while tones 2, 7, and 8 have a similar mid-low level contour. It is perceptually difficult to distinguish them in terms of pitch; thus, the question remains as to why consider them as separate tones rather than grouping them into one single category.

The reason they are treated as contrastive is primarily because the realisations of Zhangzhou tones are discovered being multidimensional, involving a variety of phonetic parameters besides pitch: the tones that share similar pitch realisation can have considerable differences in other parametric realisations. For example, as shown in Table 4-11, the unstopped tone 4 has a longer duration than the stopped tone 6. The high vowels in tone 4 are realised as monophthongs with a breathy voice rather than being diphthongised and laryngealised as in tone 6.

Table 4-11. Comparison of multidimensional realisations of citation tones 4 and 6 in Zhangzhou

Parameter	Tone 4	Tone 6
Pitch	[41]	[41]
Length	medium	short
High vowel	breathy	creaky
Mid vowel	modal	*
Low vowel	creaky	creaky
Diphthongisation	*	high vowel diphthongisation
Coda type	sonorant	not-realised obstruent
Sandhi pitch	[63]	[65] (extra high)

Therefore, one single parameter of pitch is arguably not sufficient to characterise the nature of tonal contrasts in Zhangzhou. Instead, pitches are in active and complicated interactions with other phonetic parameters to shape the multidimensional realisations of Zhangzhou tones. The next section is devoted to a discussion of how segmental and suprasegmental parameters interact with each other to construct tonal distinctions.

4.2. Interactions Among Auditory Parameters

This section primarily explores how the diverse phonetic parameters—length, vowel quality, voice quality, and syllable coda—interact to characterise Zhangzhou tones multidimensionally. It is hoped this discussion provides a deeper understanding of the nature of tonal realisations and distinctions in Zhangzhou, but also stimulates a fresh look at other Southern China dialects and languages of other regions.

4.2.1. Duration

Perceptually, tonal length varies with respect to different citation tones, but a considerable amount of variation can be observed in terms of the number of length levels and the ranking among different levels across the 21 speakers. The descriptive results are summarised in Table 4-12, in which *1* indicates the longest and *5* the shortest in individuals' length ranges.

Table 4-12. Length realisation of Zhangzhou citation tones among 21 speakers

Gender	Speaker	Tone 1	Tone 2	Tone 3	Tone 4	Tone 5	Tone 6	Tone 7	Tone 8
Female	CSM	3	2	4	4	2	5	3	1
	HHX	2	2	3	3	2	4	2	1
	HJH	2	1	3	3	1	4	2	1
	HQM	1	1	2	3	1	4	2	1
	HSX	1	1	2	3	1	4	2	1
	HYH	3	1	4	4	2	5	3	1
	LCH	1	1	3	3	1	4	2	1
	LRX	2	1	3	3	1	4	2	1
	LRZ	2	1	3	3	1	4	2	1
	LZP	2	1	4	4	2	5	3	1
	YML	1	2	4	4	1	5	3	1
	ZBQ	2	1	4	4	1	5	3	2
Male	HMC	1	1	2	2	1	3	2	1
	HTR	1	2	3	4	1	5	2	1
	HYS	1	1	3	3	1	4	2	1
	HZB	2	1	3	3	1	4	2	1
	LMY	2	1	3	3	2	4	2	1
	LYY	2	1	4	4	1	5	3	2
	WJG	1	2	2	3	1	4	3	1
	WYF	2	1	3	3	1	4	2	1
	YBR	1	1	3	3	1	4	2	1

As Table 4-12 indicates, in terms of the number of length levels, four length levels are the most likely to be identified in speakers' utterances, but variations of three levels (HMC) and five levels (CSM, HYH, LZP, YML, ZBQ, HTR, LYY) can also be observed in the data.

As for the ranking among different levels, tones 2, 5, and 8, sharing level pitch contours, are largely among the longest while tone 6, stopped mid-high falling contour, is dominantly the shortest for most speakers. Tones 3 and 4, sharing falling pitch, are perceptually difficult to be distinguished in terms of length, but both are considerably shorter than level tones 2, 5, and 8. In addition, the length of tone 1 (rising pitch contour) mediates between the longest and the second longest. It is perceived as the longest among 9 out of 21 speakers and the second longest among 10 speakers. The length of tone 7 (stopped mid-low level) is impressionistically longer than that of tones 3 and 4 (falling contour) but shorter than tone 1 (rising contour) for most speakers.

Therefore, apparent variation exists in the length realisation across the 21 speakers. Nevertheless, central tendencies can be generalised based on the information encoded in Table 4-12. Four tonal length levels are most commonly found among speakers, and the ranking can be generalised as follows: tones 1, 2, 5, and 8 > tone 7 > tones 3 and 4 > tone 6. In alignment with the pitch realisation, the overall length system of Zhangzhou citation tones is summarised in Table 4-13, and the four length levels are labelled as extra long [Vːː], long [Vː], medium [V], and short [V̌].

Table 4-13. Length realisations of Zhangzhou citation tones

Contour	Tone	Pitch	Length
Level/rising	1	mid rising [35]	extra long [Vːː]
	2	mid-low level [22]	
	5	mid level [33]	
	8	mid-low level [22]	
Stopped level	7	stopped mid-low level [221]	long [Vː]
Falling	3	high falling [51]	medium [V]
	4	mid-high falling [41]	
Stopped falling	6	stopped mid-high falling [41]	short [V̌]

As Table 4-13 shows, several interesting aspects are worthy of further mentioning with respect to the interaction between tonal pitch and duration among Zhangzhou citation tones.

- The length realisation of Zhangzhou citation tones varies across pitch contours. Falling contours are perceptually significantly shorter than the non-falling contours of rising and level.
- The length realisation of Zhangzhou non-stopped citation tones tends to have little association with the pitch height from auditory point of view. For example, identifying whether a mid-low level (tones 2 and 8) is shorter or longer than a mid-level contour (tone 5) is perceptually difficult. Therefore, conducting experiments of acoustic measuring and statistical testing is crucial to help specify the relation between duration and pitch height.
- The stopped tones generally have shorter durations than corresponding unstopped tones sharing similar pitch contours. For example, the stopped tone 7 shares a similar mid-low level contour with the unstopped tone 2, but its duration is shorter.
- The stopped tones are not always the shortest, as has been conventionally assumed in Sinitic studies (Ma, 1994; FCCEC, 1998; ZCCEC, 1999; Gao, 1999; Zhou, 2006; Chen, 2007; Yang, 2008; Guo, 2014). Instead, it can be much longer than many other

tones in a language, for example, the stopped tone 7 is longer than tones 3, 4 and 6 in citation in Zhangzhou.

- In comparison to their length realisation in the non-utterance-final position, the stopped tones have considerably longer durations in the monosyllabic setting. This phenomenon is considered to be syllable lengthening effect induced by the non-realisation of obstruent codas in citation. (See Section 4.2.4 for details.)

Therefore, the parameter of length is important in characterising Zhangzhou citation tones. The variation of length realisation can help enhance the perception and identification of a specific tone and its corresponding pitch realisation in Zhangzhou. For example, if a syllable is produced with the longest duration, it is predicted to have a level pitch contour. If a syllable is produced with the shortest duration, it is thought to be in tone 6.

4.2.2. Vowel quality

Vowel quality is observed changing with tonal categories in Zhangzhou. The high vowels undergo an active alternation between monophthongs in the unstopped tones and diphthongs in the stopped tones in Zhangzhou. On the other hand, the non-high vowels (either mid or low) are realised consistently as monophthongs without changing their qualities across different tones, but the mid vowels are not found in the stopped tones. Table 4-14 shows the vowel quality realisations with respect to Zhangzhou citation tones.

Table 4-14. Examples of vowel quality realisation with respect to Zhangzhou monosyllabic citation tones

Tone		High vowel /i/	High vowel /u/	Mid vowel /e/	Low vowel /ɔ/
Unstopped	1	[ʔʲi̤ːː35] ‘he/she’	[ʃṳːː35] ‘lose’	[keːː35] ‘chicken’	[tɔːː35] ‘capital’
	2	[ʔʲi̤ːː22] ‘move’	[ʃṳːː22] ‘word’	[teːː22] ‘question’	[tɔːː22] ‘map’
	3	[ʔʲi̤51] ‘chair’	[ʃṳ51] ‘history’	[tḛ51] ‘short’	[tɔ̰51] ‘block’
	4	[ʔʲi̤41] ‘intention’	[ʃṳ41] ‘four’	[tḛ41] ‘emperor’	[tɔ̰41] ‘contest’
	5	[ʔʲi̤ːː33] ‘play’	[ʃṳːː33] ‘affair’	[teːː33] ‘land’	[tɔːː33] ‘degree’
	8	[tɕi̤ːː22] ‘tongue’	[t̻ʷṳːː22] ‘piece’	[tʰeːː22] ‘take away’	*
Stopped	6	[ʔʲiɛ̰̆41] ‘one’	[ʃuɤ̰̆41] ‘whip’	*	[tɔ̰̆41] ‘govern’
	7	[ʔʲiɛ̰ː221] ‘bathe’	[ʃuɤ̰ː221] ‘skill’	*	[tɔ̰ː221] ‘poison’

In Table 4-14, the high front vowel /i/ is realised as a diphthong [iɛ̰], and the high back vowel /u/ is realised as a diphthong [uɤ̰] in the stopped tones 6 and 7 while both are realised as monophthongs in other non-stopped tonal environments.

Thus, tones motivate a diphthongisation process in the high vowels in Zhangzhou; conversely, this diphthongisation process helps enhance the perception and identification of specific tones and vowels. For example, the diphthongisation indicates high vowels in the stopped tones.

Nevertheless, the question remains with respect to why the high vowels undergo the diphthongisation process in the stopped tones. I suggest such a high-vowel diphthongisation process is essentially one of the phonetic effects induced by the non-realisation of obstruent codas in the monosyllabic context, as discussed in Chapter 3 and in following section 4.2.4.

4.2.3. Voice quality

Phonation also plays a role in characterising Zhangzhou citation tones. The non-modal phonation—creaky and breathy—occurs consistently as allophonic variants of modal phonation in Zhangzhou, but their distribution is conditioned by both vowel quality and tone. As Tables 4-14 and 4-15 show, the vowel quality largely constrains the realisation of voice quality. The breathy voice tends to occur only on the high vowels while the creaky voice dominantly occurs on the low vowels. The modal phonation generally dominates on the mid vowels.

Table 4-15. Voice quality realisation with respect to Zhangzhou vowels and tones

Tone	Pitch	High vowel /i/, /u/	Mid vowel /e/, /ɵ/	Low vowel /ɛ/, /ɐ/, /ɔ/
1	mid rising [35]	breathy	modal	modal
2	mid-low level [22]	breathy	modal	modal
3	high falling [51]	breathy	modal/creaky	creaky
4	mid-high falling [41]	breathy	modal/creaky	creaky
5	mid level [33]	breathy	modal	modal
6	stopped mid-high falling [41]	creaky	*	creaky
7	stopped mid-low level [221]	creaky	*	creaky
8	mid-low level [22]	breathy	modal	modal

Nevertheless, the realisation of voice quality can further vary in certain tonal circumstances. For example, in the stopped tones 6 and 7, the voice quality of high vowels shifts to creaky from a breathy voice. In the non-falling tonal environments (tones 1, 2, 5, and 8), the voice quality of low vowels changes to a modal voice from creaky. As for the mid vowels, the mid vowel [e] is occasionally found with a creaky voice in falling pitch, especially among male speakers.

Although the realisation of voice quality is constrained by tone and vowel quality in Zhangzhou, voice quality enhances the perception and identification of specific tones and vowels. For example, vowels being produced with a breathy voice indicates they are high vowels and in unstopped tones. Similarly, low vowels being produced with a modal phonation indicates they occur either in the rising or level tones. Segments being produced with a modal phonation in the falling tonal environments indicates they are mid vowels. This suggests a kind of three-register system based on phonation, although it is not so neat as this.

Furthermore, the non-realisation of obstruent codas in the monosyllabic setting also condition vowels that can occur in the stopped tones to be laryngealised, as described in the next section. Therefore, pitch, vowel quality, voice quality, and syllable codas are dependent on each other to shape the realisation and distinction of Zhangzhou tones.

4.2.4. Syllable coda

Three types of syllable coda were proposed in Chapter 3: glides (/j, w/), nasals (/m, n, ŋ/), and obstruents (/p, t, k/); however, their distribution is asymmetric with respect to tones. The glide, nasal, and zero codas are aligned with the unstopped tones while the obstruent codas are aligned with the stopped tones, as shown in Table 4-16.

Table 4-16. The distribution of syllable codas with respect to Zhangzhou citation tones

Tone		Pitch	Zero	Glide	Nasal	Obstruent
Unstopped	1	mid rising [35]	+	+	+	-
	2	mid-low level [22]	+	+	+	-
	3	high falling [51]	+	+	+	-
	4	mid-high falling [41]	+	+	+	-
	5	mid level [33]	+	+	+	-
	8	mid-low level [22]	+	+	-	-
Stopped	6	stopped mid-high falling [41]	-	-	-	+
	7	stopped mid-low level [221]	-	-	-	+

While the obstruent coda serves as an important parameter for distinguishing the stopped tones from the unstopped tones in Zhangzhou, identifying obstruent codas in the monosyllabic setting is perceptually difficult, although the bilabial coda [p] is occasionally perceivable in some speakers' utterances. Nevertheless, they can generally be identified but without audible obstruent release when the related syllables are produced in non-utterance-final positions, also referred to as *sandhi positions*, in Zhangzhou (see Chapters 6 and 7), as shown in Table 4-17, in which the superscript *f* signifies falsetto voice.

The non-realisation of obstruent codas in the monosyllabic setting gives rise to several phonetic effects on the whole syllables, including syllable lengthening, vowel laryngealisation, high vowel diphthongisation, and pitch contour depression.

- Comparing length realisation in the non-final position, the related syllables with underlying obstruent codas have significantly longer duration in the citation form. (See Chapters 5, 6, and 7 for details.)
- The vowels found in the stopped tones, basically high and low vowels, are all laryngealised and become creaky, as indicated in Table 4-18.

- The high vowels /i/ and /u/ undergo diphthongisation and become [iɛ̰] and [uɤ̰], respectively, in the stopped tonal environments as discussed in the above section.
- The pitch contours of stopped tones are depressed in association with creaky phonation. Tone 7 clearly has a similar pitch contour as tones 2 and 8, but its final portion has been depressed and released as [221], thus being distinguishable from the [22]-pitched tones 2 and 8.

Table 4-17. Obstruent coda realisation with respect to the citation and sandhi forms of Zhangzhou stopped tones

Tone	Example	Citation	Sandhi
6	/sip/ 'humid'	[ɕiɛ̰̆41] 'humid'	[ɕɪ̆ᶠp̚65.tɔ::33] 'humidity'
	/kut/ 'bone'	[kʷuɤ̰̆41] 'bone'	[kʷuᶠt̚65.tʰɐw::311] 'bones'
	/kɔk/ 'country'	[kɔ̰̆41] 'country'	[kɔ̆ᶠʔ65.kɛ::24] 'country'
7	/tsɐp/ 'ten'	[tsɐ̰:221] 'ten'	[tsɐ̆p̚32.hʷe41] 'ten-year-old'
	/tit/ 'straight'	[t̪iɛ̰:221] 'straight'	[t̪ɪ̆t̚32.ɗɔ::33] 'straight road'
	/tɔk/ 'poison'	[tɔ̰:221] 'poison'	[tɔ̆ʔ32. kʰʲi̤41] 'poison gas'

Table 4-18. Phonetic effects caused by the non-realisation of obstruent coda on preceding vowels of Zhangzhou speech

Coda	Tone	High vowel /i/	High vowel /u/	Low vowel /ɐ/	Low vowel /ɔ/
Sonorant	1	[ʔʲi̤::35] 'he/she'	[kʷṳ::35] 'turtle'	[kɐ::35] 'glue'	[tɔ::35] 'capital'
	2	[ʔʲi̤::22] 'move'	[kʰʷṳ::22] 'squat'	[ɓɐ::22] 'numb'	[tɔ::22] 'map'
	3	[ʔʲḭ51] 'chair'	[kʷṵ51] 'long (time)'	[kɐ̰51] 'crush'	[tɔ̰51] 'block'
	4	[ʔʲi̤41] 'intention'	[kʷṳ41] 'sentence'	[kɐ̰41] 'teach'	[tɔ̰41] 'contest'
	5	[ʔʲi̤::33] 'play'	[kʷṳ::33] 'worn'	[kɐ::33] 'bite'	[tɔ::33] 'degree'
	8	[tɕi̤::22] 'tongue'	[t̪ʷṳ::22] 'piece'	[tɐ::22] 'step on'	*
Obstruent	6	[ʔʲiɛ̰̆41] 'one'	[kʷuɤ̰̆41] 'bone'	[kɐ̰̆41] 'horn'	[tɔ̰̆41] 'govern'
	7	[ʔʲiɛ̰:221] 'bathe'	[kʷuɤ̰:221] 'slippy'	[ɓɐ̰:221] 'eye'	[tɔ̰:221] 'poison'

In addition, the non-realisation of obstruent codas in Zhangzhou essentially reflects a well-established tendency in the history of the Sino-Tibetan language family and other languages for obstruent codas to lenite to a glottal stop and then to rhyme laryngealisation. Because the articulatory gesture for laryngealisation production involves a raised larynx, it reduces the supraglottal cavity (Ladefoged, 1971; 2003; Laver, 1980; 1994), conditioning the tongue body to move downwards and causing the high vowels to inglide to a lower target, giving rise to diphthongisation.

Thus, I suggest processes of syllable lengthening, vowel laryngealisation, diphthongisation, and pitch depression occurring in the stopped tones are motivated essentially by the non-realisation of obstruent codas. These processes, conversely, make the stopped tones distinguishable from their corresponding unstopped tones that share similar pitch contours. Therefore, the obstruent codas also help enhance the perception and identification of the stopped tones and segments in Zhangzhou. For example, obstruent codas indicate vowels are laryngealised, the high vowels are diphthongised, and the related tones are either tone 6 or tone 7 in Zhangzhou.

4.2.5. Summary

As described, the tonal realisations of Zhangzhou are multidimensional. Tones differ not only in terms of pitch but also in other segmental and suprasegmental parameters. Each phonetic parameter does not independently exist in characterising the tonal contrasts; rather, they interact with each other in an active and systematic way. Therefore, the phonological tones in Zhangzhou are essentially a complex of co-varying phonetic features, as shown in Table 4-19.

Thus, a high vowel in tone 1 is expected to have a mid-rising pitch contour, extra-long duration, and a breathy voice, and it precedes a sonorant coda. Tone 3 indicates a high-falling pitch contour, a medium duration, a sonorant coda, and three different phonation types: breathy voice for the high vowels, modal voice for the mid vowels, and creaky voice for the low vowels. Similarly, tone 6 indicates a mid-high falling contour similar to tone 4, but it has the shortest duration. Further, all vowels in tone 6 are creaky, and all syllables have an underlying obstruent coda generally not realised in a monosyllabic utterance. In addition, the high vowels are predicted to undergo a process of diphthongisation. Tone 8 seems indistinguishable from tone 2 in the monosyllabic setting with respect to all the parameters shown in Table 4-19, however, it is realised differently from tone 2 in multisyllabic contexts, as discussed in Chapters 6, 7 and 9.

Such multidimensional realisations of Zhangzhou tones render the conventional definition of tones as the lexical phonemicisation of pitch distinctions inadequate for understanding their nature. Thus, conducting a linguistic study of tones does not simply include providing a list of pitch contrasts but must include exploring how tones are essentially realised and how segmental and suprasegmental parameters interact in actual use of tones lexically.

Table 4-19. Auditory properties of Zhangzhou citation tones

Parameter	Tone 1	Tone 2	Tone 3	Tone 4	Tone 5	Tone 6	Tone 7	Tone 8
Pitch	[35]	[22]	[51]	[41]	[33]	[41]	[221]	[22]
Length	extra long	extra long	medium	medium	extra long	short	long	extra long
High vowel	breathy	breathy	breathy	breathy	breathy	creaky	creaky	breathy
Mid vowel	modal	modal	modal	modal	modal	*	*	modal
Low vowel	modal	modal	creaky	creaky	modal	creaky	creaky	modal
Diphthongisation	-	-	-	-	-	+	+	-
Coda type	sonorant	sonorant	sonorant	sonorant	sonorant	not realised obstruent	not realised obstruent	sonorant

4.3. Previous Studies

Numerous descriptions of Zhangzhou citation tones can be seen in the literature (Dong, 1959; Lin, 1992; Ma, 1994; FCCEC, 1998; ZCCEC, 1999; Gao, 1999; Zhou, 2006; Chen, 2007; Yang, 2008; Guo, 2014; Huang et al., 2016). These previous studies described tones almost exclusively with respect to one dimension of pitch and identified a seven-way tonal contrast. Nevertheless, their descriptions differed not only from each other but also largely from the result of this study, with the exception of Huang et al. (2016), as shown in Table 4-20, where *I*, *II*, *III*, and *IV* correspond to the Middle Chinese tonal categories of *Ping*, *Shang*, *Qu*, and *Ru*, respectively, while *a* and *b* represent the *Yin* and *Yang* registers, respectively.

Table 4-20. Previous descriptions of Zhangzhou citation tones

Author	Year	Tone 1 (Ia)	Tone 2 (Ib)	Tone 3 (II)	Tone 4 (IIIa)	Tone 5 (IIIb)	Tone 6 (IVa)	Tone 7 (IVb)	Tone 8 (IVb)
Dong	1959	24	212	53	32	33	32	13	*
Lin	1992	44	13	53	21	22	32	12	*
Ma	1994	44	12	53	21	22	32	121	*
FCCEC	1998	44	13	53	21	22	32	121	*
ZCCEC	1999	44	13	53	21	22	32	121	*
Gao	1999	45	23	53	21	33	21	121	*
Zhou	2006	44	13	53	21	22	32	121	*
Chen	2007	44	13	53	21	22	32	121	*
Yang	2008	44	13	53	21	22	32	121	*
Guo	2014	44	13	53	21	22	31	121	*
Huang et al.	2016	35	22	51	41	33	41	221	*
This study	**2018**	**35**	**22**	**51**	**41**	**33**	**41**	**221**	**22**

The reason for such descriptive variation is unclear. One may ascribe it to a range of factors, such as idiolectal or sub-regional differences and between-speaker or between-transcriber variations. For example, scholars conduct the impressionistic description commonly based on a single speaker's utterances; thus, it is perhaps reasonable to see such disagreements and differences as a result of between-speaker variations. Further, differing transcriptional skills or modifications can cause the variations in the descriptions; however, there seems no way to determine the specific reasons.

The existing issue of inconsistent descriptions can give rise to several serious consequences.

- It leads to questioning the accuracy and reliability of previous pitch transcriptions and descriptions of monosyllables.
- It seems unreliable to derive a system of tonal representation of Zhangzhou as a whole variety on the basis of the inconsistently descriptive results.

- It also leads to questioning the accuracy and reliability of previous studies of tone sandhi, which depend exclusively on such impressionistic results of citation tones.
- It creates confusion and ambiguous results for tonal modelling. For example, given the system of Yip (1980), tone 1 can have three different representations, including [+upper, HH] (Lin, 1992; Ma, 1994; FCCEC, 1998; ZCCEC, 1999; Zhou, 2006; Chen, 2007; Yang, 2008; Guo, 2014); [+upper, LH] (Gao, 1999), and [-upper, LH] (Huang et al., 2016).

Another important problem with previous studies is that they described tones almost exclusively with respect to the single dimension of pitch. No descriptive work has investigated the way in which pitch interacts with other phonetic features to shape the rich and complex paradigm of tonal distinctions in Zhangzhou. Therefore, we are still far from a satisfactory and sufficient understanding of the nature of tones in Zhangzhou if we rely simply on previous descriptions.

4.4. Conclusion

This chapter included a systematic description of the auditory properties of Zhangzhou monosyllabic citation tones. It showed the tonal realisations in Zhangzhou are multidimensional and each phonological tone of Zhangzhou is essentially a complex of co-occurring phonetic features. They differ from each other not only in pitch height and contour but also in a variety of segmental (vowel quality and syllable coda) and suprasegmental parameters (pitch, length, and voice quality), which interact dependently with each other in a complicated but systematic way.

This description has factored out the problems of inconsistency and inadequacy of the previous studies and has accomplished one of the primary goals of this thesis that uses modern linguistic theory to characterise the auditory properties of Zhangzhou tones from multiple speakers' production. This description supersedes the results of previous research and offers an improved understanding of the nature of tonal realisations in Zhangzhou.

The description provides a descriptive framework for subsequent quantification of the acoustic properties of Zhangzhou citation tones in Chapter 5. It also provides a foundation to explore and explain the interaction of these citation tones in the disyllabic setting in Chapters 6, 7, and 8 and the mapping between citation and disyllabic tones in Chapter 9.

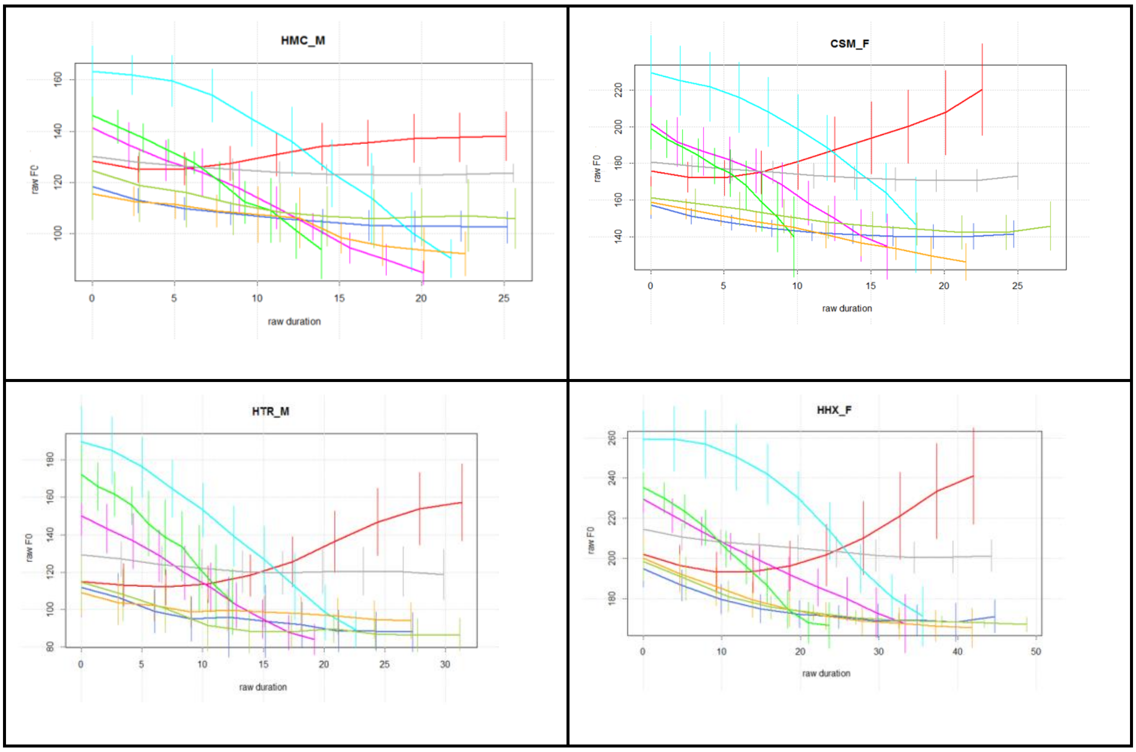
Chapter 5: Acoustic Properties of Zhangzhou Citation Tones

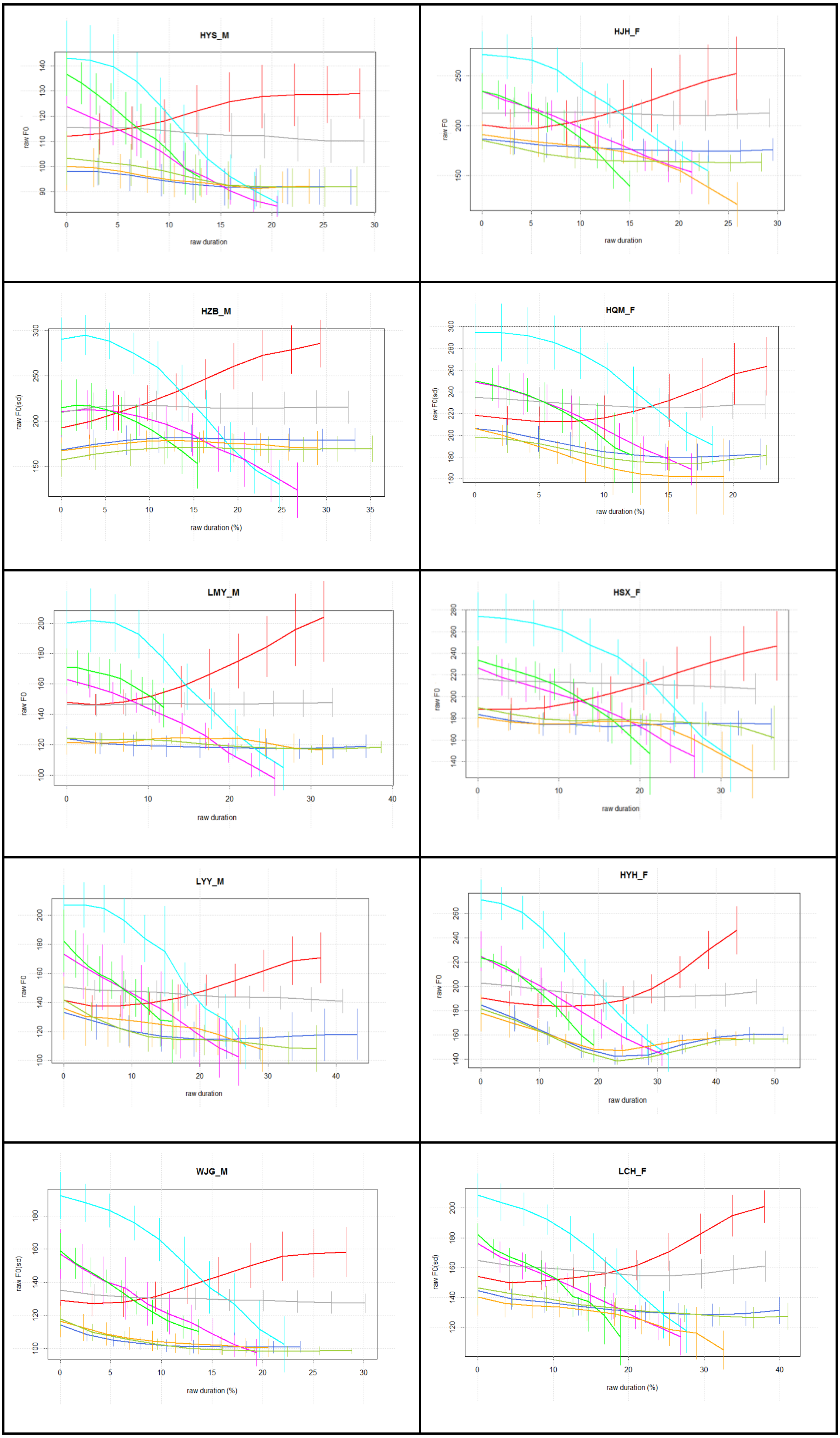
Chapter 4 reveals that tonal realisations of Zhangzhou citation tones are multidimensional from the auditory point of views. This chapter aims to explore what Zhangzhou citation tones are essentially realised in acoustics, with the additional goal of examining the accuracy of auditory observations and classifications. The theory of articulatory phonetics is drawn upon to explain how different vocal tract shapes and actions contribute to the acoustic variation for different tonal production. Pairwise *t* tests and hierarchical clustering algorithms were employed to determine how many normalised F0 and duration levels are contrastive among Zhangzhou citation tones. The SplitsTree software was used to address how Zhangzhou citation tones are related to each other from the phonological perspective.

5.1. Acoustic Variation in Tonal F0 and Duration

Chapter 4 revealed a considerable amount of variation with respect to the pitch and length realisations of Zhangzhou citation tones across individual speakers. The auditory observation is acoustically justified by the quantified results of the raw mean F0 shapes of individual citation tones as a function of raw mean duration across the 21 speakers (plotted in Figure 5-1). The values shown in the figures are absoluate mean values without normalisation.

Throughout Figure 5-1, the results from male speakers are shown in the left panels and the plottings from female speakers, except LZP and ZBQ, are shown in the right panels. Vertical lines across individual F0 curves indicate one standard deviation away from the mean values at the sampling points concerned.




HYS_M
raw F0
raw duration
HJH_F
raw F0
raw duration
HZB_M
raw F0(sd)
raw duration (%)
HQM_F
raw F0(sd)
raw duration (%)
LMY_M
raw F0
raw duration
HSX_F
raw F0
raw duration
LYY_M
raw F0
raw duration
HYH_F
raw F0
raw duration
WJG_M
raw F0(sd)
raw duration (%)
LCH_F
raw F0(sd)
raw duration (%)

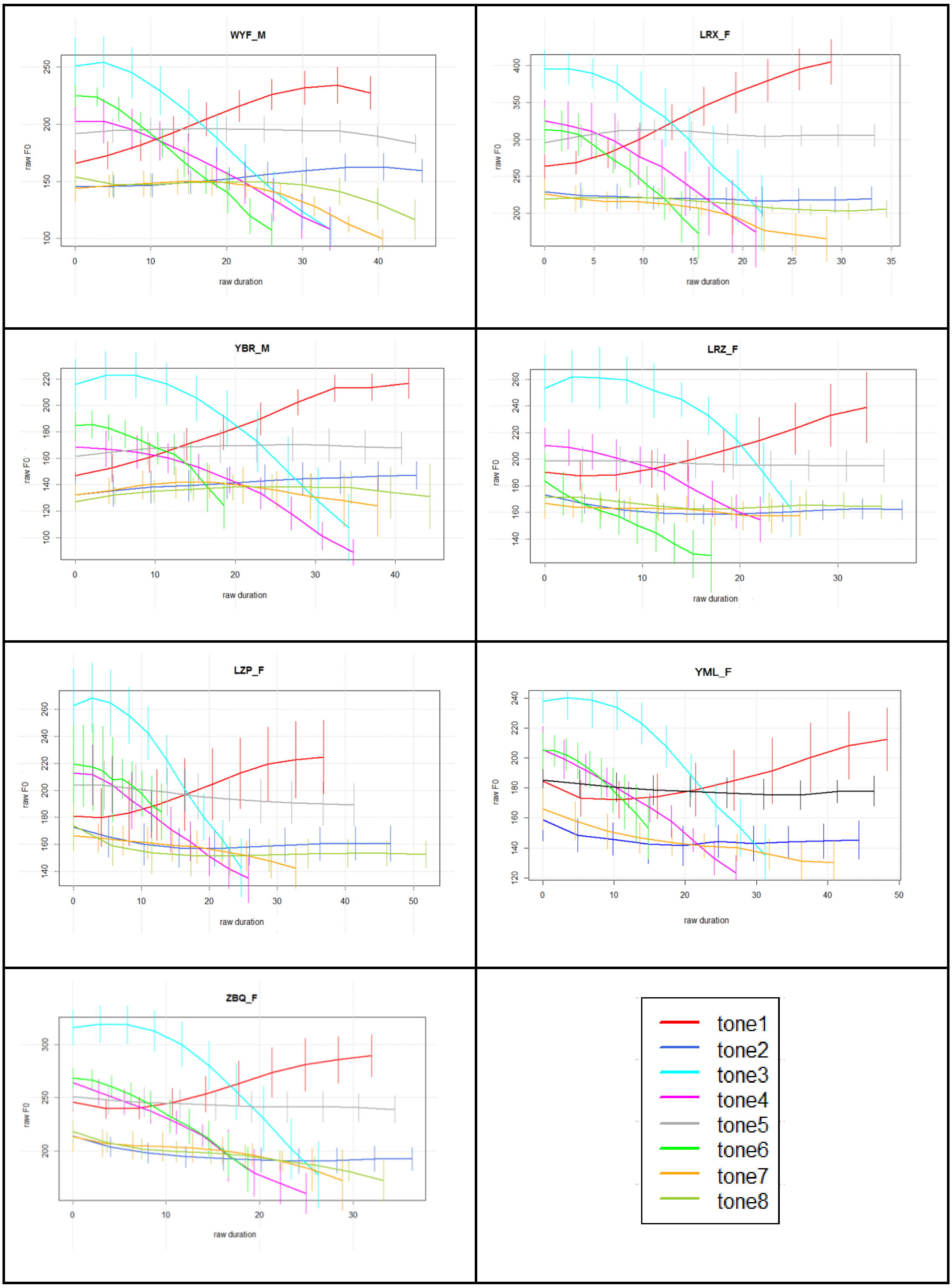


Figure 5-1. Raw mean F0 shapes of Zhangzhou citation tones for individual speakers.

As indicated in Figure 5-1, several notable aspects show numerically significant differences across speakers, for example, in the distribution of F0 ranges, individual F0 curves, and individual F0 curve duration.

5.1.1. Acoustic differences in F0 range

Female speakers largely have a greater F0 range than male speakers because of the shorter and less massive structure of their vocal folds (Rose, 1987, 2000). As summarised in Tables 5-1 and

5-2, the average F0 range of Zhangzhou male speakers is 180 Hz, which is about three quarters the average female F0 range of 235 Hz. Nevertheless, exceptional cases do exist. Male speakers HZB and WYF had a noticeably greater F0 range than many female speakers, such as CSM, HHX, HYH, LCH, and YML.

Further, significant inter-gender differences also occur in terms of the average F0 range. For example, the male speaker HZB has the widest F0 range of 267 Hz, which is nearly 2.5 times greater than that of HMC, with the narrowest F0 range of 107 Hz. Similarly, the female speaker LRX has the widest F0 range of 388 Hz, more than 2.5 times greater than the narrowest F0 range of 149 Hz of LCH.

Given the significant inter- and intra-gender differences in the F0 range, it is not surprising to observe that tones belonging to different categories have similar F0 realisation among speakers, even when one speaker's mid-level tone is acoustically lower than another speaker's mid-low-level tone. For example, the mid-level F0 contour (tone 5) of female speaker LCH is lower than the mid-low-level F0 contours (tone 2) of most other females, except CSM, LZP, and YML.

Table 5-1. F0 ranges of Zhangzhou male speakers (Hz.)

F0	HMC	HTR	HYS	HZB	LMY	LYY	WJG	WYF	YBR	Ave
Max	184	222	170	344	251	250	217	301	249	258
Min	77	75	78	77	77	78	83	80	77	78
Range	107	147	92	267	174	172	134	221	172	180

Table 5-2. F0 ranges of Zhangzhou female speakers (Hz.)

F0	CSM	HHX	HJH	HQM	HSX	HYH	LCH	LRX	LRZ	LZP	YML	ZBQ	Ave
Max	260	304	319	343	357	298	232	480	307	350	262	363	326
Min	88	126	88	83	79	106	83	92	94	83	83	89	91
Range	172	178	231	260	278	192	149	388	213	267	179	274	235

5.1.2 Acoustic differences in F0 realisation

Considerable variations also occur with respect to the F0 realisation of Zhangzhou citation tones, which seem less gender sensitive but rather speaker dependent. In terms of F0 contour shapes, variations are most apparent in tones 2, 7, and 8. For example, tone 7 was largely realised as a mid-low-falling contour, but speakers differ slightly concerning the point where the slope start to occur. Some speakers (HJH, HSX, LRX, YML, ZBQ, WYF) show a mid-low-level plateau during the greater part of the contour but with a final fall while some other speakers (HTR, WJG, HHX, HQM) present a low-level plateau but with an initial fall. Exceptional realisations of low level

(HYS, HZB, LMY, YBR, LRZ), low fall (HMC, LYY, CSM, LCH, LZP), and low-dipping (HYH) also occur (Figure 5-1).

As for the contour relative height, a substantial degree of between-speaker variation can also be observed, especially in the rising contour (tone 1) and the stopped mid-high falling contour (tone 6). For example, the offset heights of F0 contours show notable variation in tone 6 among speakers. It is dominantly realised as a mid-high falling contour, with the lowest F0 offset, but some speakers realise it with a slightly higher offset (HTR, HYS, HZB, HQM, WJG, LYY, YBR, ZBQ) or with a much higher offset (LMY, LZP).

5.1.3. Acoustic differences in length realisation

The acoustically quantified result of tonal absolute length values also shows notable variations across tones and across speakers, as shown in Table 5-3 for male speakers and in Table 5-4 for female speakers. In general, female speakers tend to have a longer duration than their male counterparts for most citation tones; however, exceptional cases can also be seen with respect to individual speakers. For example, the duration values from male speakers WYF and YBR appear greater than the values from some female speakers, such as CSM, HJH, and HQM.

As for the individual tones, tone 6 appears the shortest among the eight citation tones for all male and female speakers. The duration values of tones 3 and 4 are largely similar to most speakers, which is slightly shorter than tone 7 but significantly shorter than the rest of other tones. Tones 1, 2, 5, and 8 apparently have much longer duration, but the ranking among them tends to be speaker-independent. For example, speakers MHC, and HQM have very similar duration values for these four tones. Speakers LYY, ZBQ, and LRZ have the largest duration value for tone 2. Speaker WJG has the longest duration in tone 5; while many female speakers, such as CSM, HHX, HYH, LCH, LRX, and LZP, present the largest duration values in tone 8.

As indicated, the length realisation of Zhangzhou citation tones also presents various variation as a consequence of significant inter- and intra-gender effects. In addition, the length tends to have different scales with respect to different tonal categories. However, substantial statistical tests are needed in order to derive a length system representing Zhangzhou tonal speech, and to examine whether the internal relation between tonal duration and F0 in Zhangzhou also reflects the across-linguistic tendency of an inverse association or presents a language-specific pattern. This issue will be mainly addressed in Section 5.3.1 for citation tones, and in Chapters 7 and 8 for disyllabic tones in Zhangzhou.

Table 5-3. Raw mean duration values of Zhangzhou citation tones among male speakers (csec.)

Dur.	HMC	HTR	HYS	HZB	LMY	LYY	WJG	WYF	YBR	Ave
Tone 1	25.1	31.3	28.6	29.3	31.6	39.3	28.2	38.5	41.7	32.6
Tone 2	25.2	27.3	27.7	33.3	36.8	43.1	23.7	45.7	42.6	33.9
Tone 3	21.8	22.6	20.6	24.7	26.6	26.7	22.1	33.7	34.2	25.9
Tone 4	20.1	19.2	20.5	26.7	25.5	25.7	19.4	33.6	34.7	25.0
Tone 5	25.6	29.8	28.9	32.5	32.6	40.9	30.1	44.9	40.7	34.0
Tone 6	13.9	12.5	13.0	15.4	11.9	15.8	13.7	25.9	18.6	15.6
Tone 7	22.7	27.1	23.6	29.0	31.4	30.8	20.6	40.6	37.8	29.3
Tone 8	25.7	31.2	28.3	35.2	37.1	37.1	28.8	44.8	42.5	34.5

Table 5-4. Raw mean duration values of Zhangzhou citation tones among female speakers (csec.)

Dur.	CSM	HHX	HJH	HQM	HSX	HYH	LCH	LRX	LRZ	LZP	YML	ZBQ	Ave
T1	22.6	42.0	25.8	22.6	36.9	43.6	37.8	28.9	33.1	37.4	48.3	32.0	34.2
T2	24.7	44.7	29.5	22.2	36.3	51.3	39.9	33.0	36.6	46.6	44.4	36.3	37.1
T3	18.0	35.5	22.9	18.4	31.2	31.7	27.6	21.9	25.2	24.7	31.2	26.3	26.2
T4	16.1	33.3	21.3	16.8	26.7	30.8	26.9	21.3	22.0	27.6	27.1	25.0	24.6
T5	25.0	44.3	29.2	22.5	34.3	46.8	38.1	33.5	34.7	41.3	46.5	34.4	35.9
T6	9.7	23.7	15.0	12.4	21.2	19.4	18.9	15.5	17.1	13.0	14.0	18.7	16.6
T7	21.5	41.8	25.9	19.3	33.9	43.3	32.6	28.5	26.1	32.7	36.4	28.8	30.9
T8	26.6	48.8	28.6	22.6	36.2	52.2	41.1	34.5	34.4	51.8	49.1	33.0	38.2

5.1.4. Summary

As described, although the acoustic signals were readily quantifiable, they inevitably contained the imprints of individual vocal tract differences and other extralinguistic information, giving rise to inter- and intra-gender variations in the quantified results. Therefore, one single speaker was insufficient to represent a language as an individual variety. Given this acoustic variability, it is necessary to apply normalisation to abstract the speaker-dependent acoustic material and to extract, as much as possible, the linguistic-phonetic contents that entwined in speech signals in order to derive a system representing the variety being considered as a whole.

5.2. *Z*-Score Normalised F0 Model

This section aims to achieve a linguistically-phonetic F0 representation of Zhangzhou citation tones by applying the methods of acoustic quantification, normalisation, and statistical testing. It also examines the plausibility of the auditory observations made in Chapter 4. Two specific issues are mainly addressed: (1) what individual citation tones are realised in terms of *z*-score normalised F0 from 21 speakers, and (2) how many normalised F0 levels are contrastive.

The methodology introduction can be referred back to sections 2.4 and 2.5 in Chapter 2. The normalised F0 values, as well as the statistical testing results, are provided in Appendix B.

5.2.1. *Z*-score normalised F0

Figure 5-2 shows the *z*-score normalised F0 contours of individual citation tones as a function of their corresponding normalised duration values from 21 speakers. All but tone 1 showed a downward trend over the course of utterances. Such a F0 declination has been cross-linguistically reported in the literature (e.g., Lieberman, 1967; Pierrehumbert, 1989; Maeda, 1976; Cohen & Collier, 1982; Ladd, 1984; Yuan & Liberman, 2010; Rose, 2014), and the reasons are argued from different perspectives, for example, drop in subglottal pressure and laryngeal tension (Lieberman, 1967; Pierrehumbert, 1987) or the tracheal pull hypothesis (Maeda, 1976).

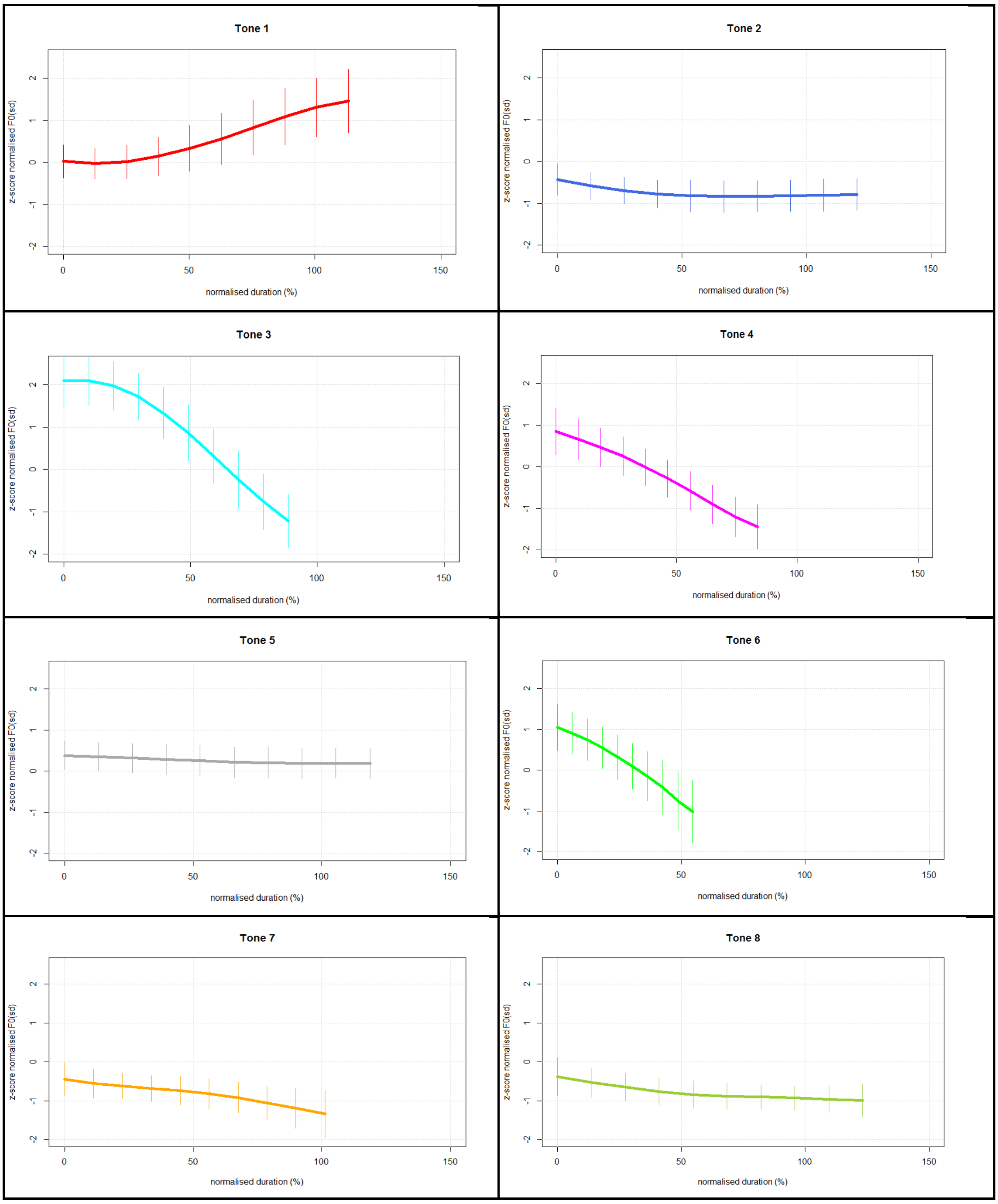


Figure 5-2. *Z*-score normalised F0 contour of individual Zhangzhou citation tone as a function of normalised duration from 21 speakers.

Therefore, taking into account the F0 declination effect is imperative in plotting the acoustic correlate of tonal pitch. Figure 5-3 shows the result of the F0 declination adjustment, rendering level contours as effectively level and the slope of the contoured tones not as steep as those shown in Figure 5-2. The curves in Figure 5-3 thus constitute the linguistically phonetic F0 system of Zhangzhou citation tones, representing the central tendency of Zhangzhou as an independent variety. 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). The normalisation index (NI) of 8.58 indicates more than eight folds of speaker-dependent variances entwined in the raw F0 values have been reduced by the *z*-score normalisation.

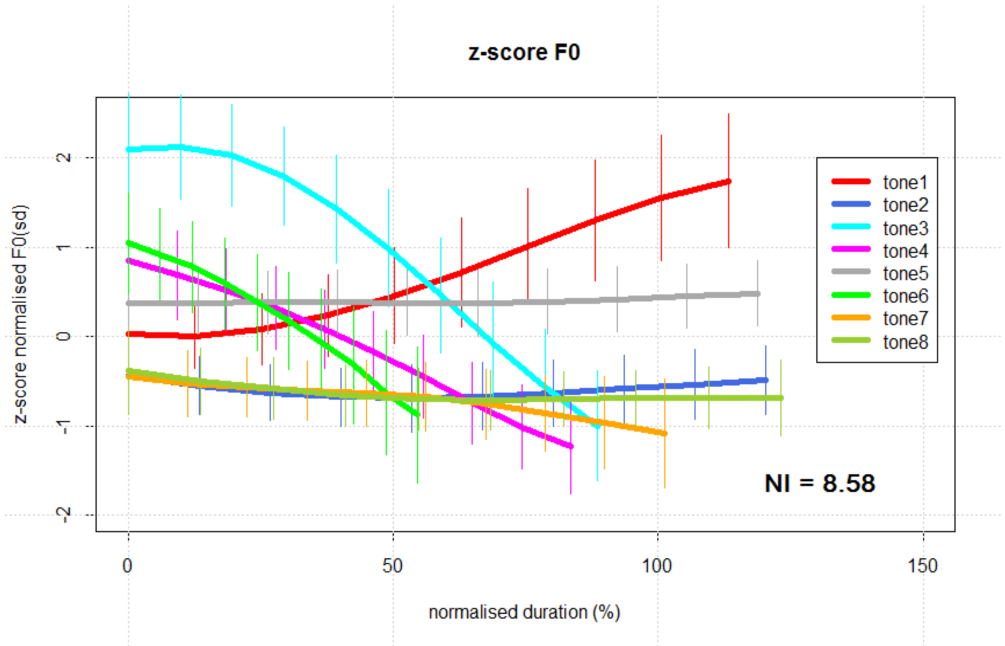


Figure 5-3. A linguistic-phonetic F0 representation of Zhangzhou citation tone system from 21 speakers.

5.2.2. Statistical testing

As described, Zhangzhou has a relatively simple F0 inventory in terms of contour shape; however, how many normalised F0 levels are categorically distinguished, and how can they be determined scientifically? Each normalised F0 contour has two putative targets of onset and offset; thus, they logically may have 16 (= 2 * 8) putative levels, as indicated in Figure 5-4. The left panel shows the variations of normalised F0 levels at the 10% sampling point (onset) while the variations at the 100% sampling point (offset) are shown on the right. Given the logically possible 16 normalised F0 levels, how do we decide which level is categorically higher than others and which levels are categorically the same?

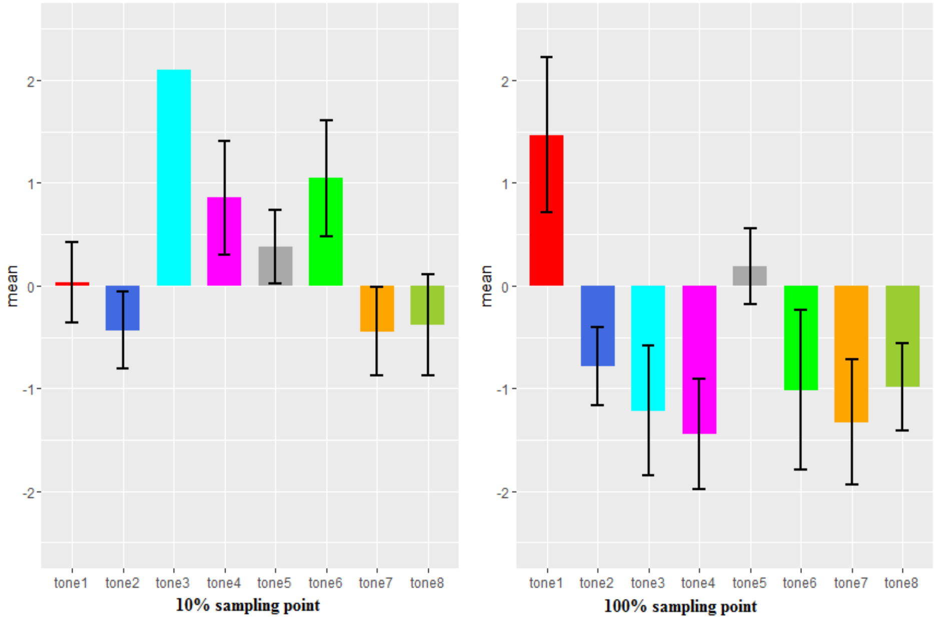


Figure 5-4. Variation of normalised F0 levels at both 10% and 100% sampling points by 21 speakers.

To address this issue, the technique of pairwise *t* tests was applied to compare 120 (=16*15/2) paired differences, as shown in Table B3 in Appendix B, under the assumption that all putative F0 levels are independent and identically normally distributed. The Bonferroni corrected alpha of 0.00041667 (=0.05/120) was performed to achieve significance. To visualise the testing result, the hierarchical clustering algorithm was further applied to generate a network reflecting the relation between all logically possible putative F0 levels for the eight citation tones, as shown in Figure 5-5. The vertical lines represent the amount of statistically significant difference in terms of normalised F0 values across 16 putative targets. The horizontal lines instead indicate which citation tone connects to which other tone. For example, tone 3 at the 10% sampling point connects to tone 1 at the 100% sampling point in the normalised F0 value.

Nevertheless, the population size of 16 putative targets creates a difficulty for the threshold selection to cluster the targets into different classes. For example, selecting a threshold at 1, as mostly used in this thesis (see later chapters), would isolate tone 1 at the 10% sampling point forming a separate F0 level and would create six levels for the F0 system of Zhangzhou citation tones. Taking into consideration the factors of acoustic quantified result, auditory impression, and typological properties of the tonal system, it is appropriate to modify the threshold to 1.5 and cluster these putative points into five reasonable levels as shown in the figure.

The levels have been further ranked in sequence from the highest (*5*) to the lowest (*1*) in accordance with the acoustic quantified and normalised result. For example, the normalised F0 value of tone 3 at the 10% sampling point clustered with the normalised F0 value of tone 1 at the 100% sampling point, and both are ranked at the highest level. On the contrary, the normalised F0 offset of falling tones 3, 4, 6, and 7 clustered at the lowest level.

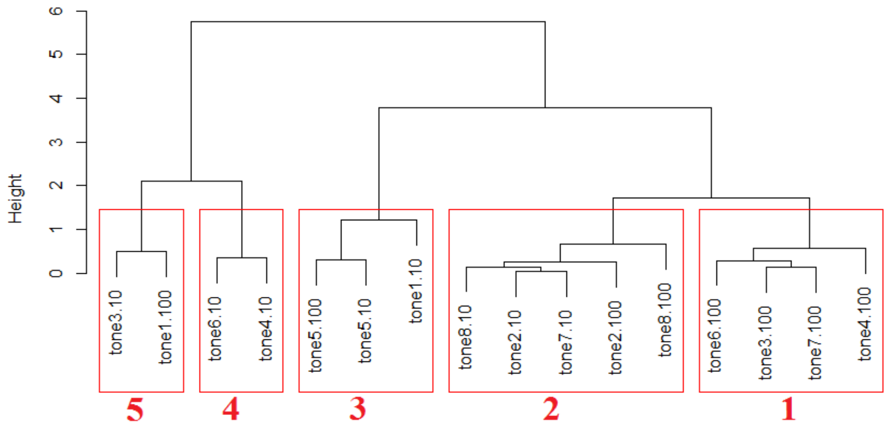


Figure 5-5. Clustering of normalised F0 levels of Zhangzhou citation tones at 10% and 100% sampling points based on pairwise *t* tests.

5.2.3. Summary

On the basis of the acoustic quantifying, statistical testing, and clustering result, the invariant linguistic-tonetic F0 information of Zhangzhou citation tones can be indicated using numerical notation, with *5* indicating the highest F0 level and *1* the lowest, as shown in Table 5-5. The system represents the central tendency of Zhangzhou as an independent variety. What is worth further mentioning is why the F0 contour of tone 7 is transcribed as [221] rather than [22]. The major reason for the treatment is that over 70% portion of the F0 contour of tone 7 is overlapped with that of tones 2 and 8 and present a mid-low level tendency, but tone 7 has a statistically significant lower offset which is claimed to be depressed by the articulation of creaky phonation (see Section 5.3.2 for acoustic details). It is thus proper to denote the F0 contour of tone 7 as [221] rather than [21].

Table 5-5. The linguistically phonetic F0 system of Zhangzhou citation tones

Tone	F0 contour description	Numerical representation
1	mid rising	[35]
2	mid-low level	[22]
3	high falling	[51]
4	mid-high falling	[41]
5	mid level	[33]
6	stopped mid-high falling	[41]
7	stopped mid-low level	[221]
8	mid-low level	[22]

In addition, as Table 5-3 shows, the pitches of tones 2 and 8 were statistically indistinguishable, with both presenting a mid-low level contour [22]. Similarly, the normalised F0 contours and heights of tones 4 and 6 were also statistically indistinguishable. Thus, distinguishing tone 2 from tone 8 and tone 4 from tone 6 is not statistically possible in terms of the single parameter of normalised F0. This generalisation justifies the auditory observation in Chapter 4 that the single dimension of pitch is not sufficient to distinguish Zhangzhou tones. Therefore, other phonetic parameters should be taken into account for the exploration of the nature of Zhangzhou tonal contrasts, as discussed in the rest of this chapter and later chapters.

5.3 Interactions Among Acoustic Parameters

Chapter 4 indicated the auditory realisations of Zhangzhou tones are multidimensional, involving a variety of segmental and suprasegmental parameters beside pitch. This section adopts the acoustic and statistical methods to examine the length realisation of Zhangzhou citation tones and the statistical relation among different normalised levels. It also uses the theories of acoustic and articulatory phonetics to explore and explain how these parameters interact to shape the tonal contrasts in acoustic signals and how the acoustic manifestations of various parameters vary along with the adjustments of articulatory configuration for different tonal productions.

5.3.1. Duration

The tonal length was described in Chapter 4 as changing with respect to tonal categories in Zhangzhou, although speakers had slight differences in terms of the number of length levels and length ranking across different tones. This subsection aims to obtain a linguistically acoustic length system of Zhangzhou citation tones by addressing two specific issues: (1) what individual citation tones are realised in terms of normalised duration from 21 speakers, and (2) how many normalised length levels are contrastive statistically.

The bars in Figure 5-6 present the acoustic quantification results of tonal duration from 21 speakers, which are expressed as a percentage of the average duration value of all citation tones. The normalised duration values for each tone appear in Appendix B (Tables B4).

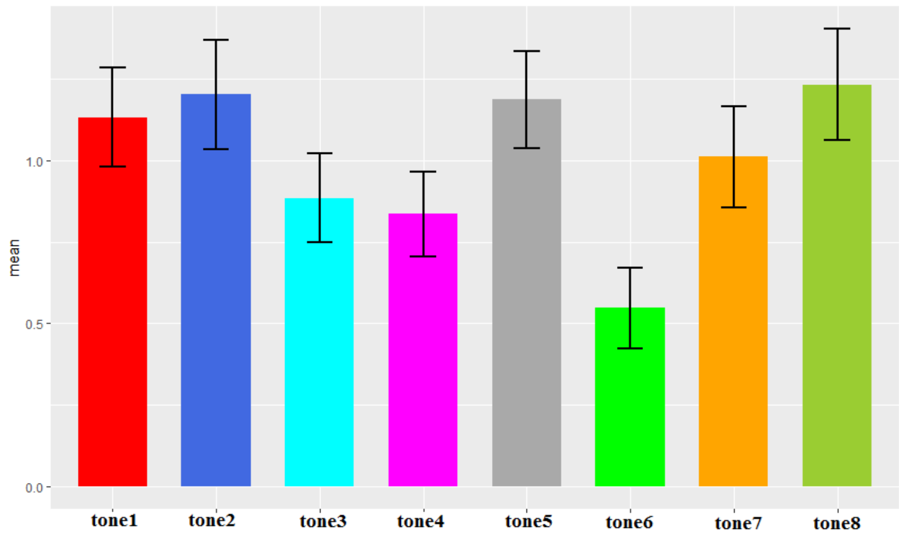


Figure 5-6. Normalised F0 duration of Zhangzhou citation tones from 21 speakers.

As Figure 5-6 shows, tonal length does vary across different tones. Tone 6 had the shortest duration, while the level contours (tones 2, 5, and 8) tended to be the longest and the falling contours (tones 3, 4, 6, and 7) tended to have shorter durations than both rising (tone 1) and level (tones 2, 5, and 8) contours in Zhangzhou. However, whether the length differences of the eight citation tones are statistically significant needs to be tested by conducting exhaustive comparisons on 28 (=8*7/2) pairs with the Bonferroni corrected alpha of 0.00186 (=0.05/28) to control for the Type I Error. The assumption for the pairwise *t* tests is that the paired duration differences are independent and identically normally distributed.

The statistical testing result was tabulated in Appendix B (Tables B5) for reference, and was also visualised hierarchically as shown in Figure 5-7 using the clustering algorithms. The threshold selected at 1 further clusters the eight citation tones into four different classes with *1* representing the longest and *4* the shortest, in accordance with the acoustic quantified and normalised result. For example, the length of tone 6 represents the shortest level, while tones 1, 2, 5, and 8 cluster together and are among the longest. Therefore, a linguistic-tonetic length representation of Zhangzhou citation tones is derived as shown in Table 5-6.

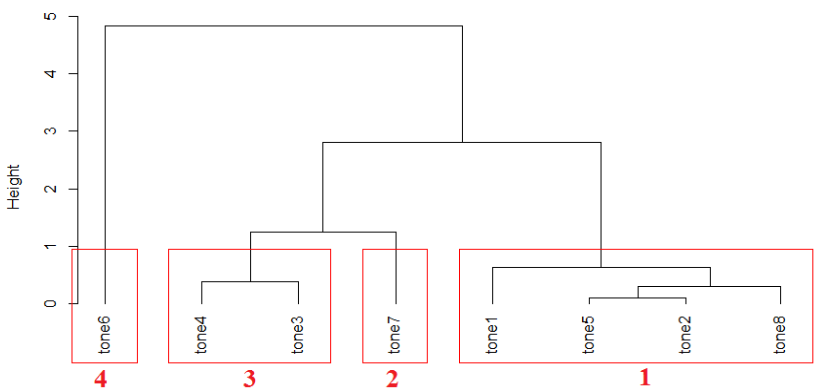


Figure 5-7. Clustering of normalised tonal length levels by pairwise *t* tests based on pairwise *t*-tests.

Table 5-6. A linguistically phonetic length system of Zhangzhou citation tones

Tone	F0 contour	Length	Notation	Length ranking
1	mid rising [35]	extra long	[V::]	1
2	mid-low level [22]			
5	mid level [33]			
8	mid-low level [22]			
7	stopped mid-low level [221]	long	[V:]	2
3	high falling [51]	medium	[V]	3
4	mid-high falling [41]			
6	stopped mid-high falling [41]	short	[V̌]	4

In addition, some generalisations can be made concerning the statistical relation among different normalised length levels with respect to tonal categories and pitch.

- No statistically significant differences existed among level contours (tones 2, 5, and 8) and among unstopped falling contours (tones 3 and 4), failing to support the measurably positive interaction between the raw mean values of F0 height and F0 duration in Zhangzhou and the cross-linguistic assumption of an inverse correlation between pitch height and duration.
- The rising contour (tone 1) was statistically grouped with the level contours but was longer than any other falling tones, supporting the universal assumption that an upward F0 has a marked tendency to take longer time than a downward change (Gandour, 1977; Ohala & Ewan, 1972).
- The stopped tones were consistently statistically significantly shorter than their corresponding unstopped tones that share similar F0 realisation, supporting both the auditory observation in Chapter 4 and the acoustic quantified result in Section 5.1.3.
- The stopped tone 7 of a mid-low-level contour with a final fall was statistically shorter than both rising and level contours but longer than other types of falling contours, supporting the auditory observation in Chapter 4 that the stopped tone is not always the shortest but can be longer than many other tones.

Therefore, the parameter of length is apparently an important part of tonal realisations in Zhangzhou. It can function with the parameter of pitch/F0 for categorising tonal contrasts.

5.3.2. Vowel quality

In Chapter 4, the vowel quality of high vowels—/i/ and /u/—was observed undergoing alteration with respect to tone in Zhangzhou. They are realised as monophthongs [i] and [u] in the non-stopped tones but are diphthongised to [iɛ̯] and [uɤ̯], respectively, in the stopped tones 6 and 7. From the articulatory point of view, the diphthongisation process indicates two articulatory trajectories during the articulation of related sounds (Laver, 1994; Harrington, 2010; Lorenz, 2013). In the case of Zhangzhou, the tongue tends to move from a more peripheral area of the vocoid space towards a lower and more central zone of the speaker's oral cavity.

According to the acoustic theory of speech production, diphthongisation indicates that the acoustic correlates of vowels—mainly the first two formants (F1 and F2)—change as a function of time (Johnson, 2011; Ladefoged, 2003; Baart, 2010). The acoustic realisation of Zhangzhou vowel quality with respect to the citation tones is illustrated in the following figures. Since no mid vowels were found in the stopped tonal environments, only the high and low vowels are shown below.

In Figure 5-8, the two monosyllabic morphemes have the same high vowel /i/ at the underlying level and similar falling F0 contour at the surface level, but they present different formant patterns. In the token /ʔi4/ 'intention', the F1 and F2 curves are relatively steady without significant changes across the whole duration. The relatively steady formant patterns of F1 and F2 indicate the tongue remains in the same position during the articulation. Therefore, this high vowel /i/ is realised as a monophthong in this unstopped tone.

Nevertheless, in the token /ʔit6/ ‘one’, the F1 and F2 change gradually once the peak of sonority is reached. For example, at the 10% point, the frequency values for F1 and F2 are 365 Hz and 2145 Hz, respectively, but at the 90% point, the F1 value is raised to 465 Hz while the F2 value is lowered to 1827 Hz. The increasing F1 value indicates a lowering trend of tongue movement while the decreasing F2 value indicates a backward movement of the tongue. Therefore, during the articulation of the token, the tongue moves to a lower and more central position so that this high vowel is realised as a diphthong [iɛ̯] in this stopped tone.

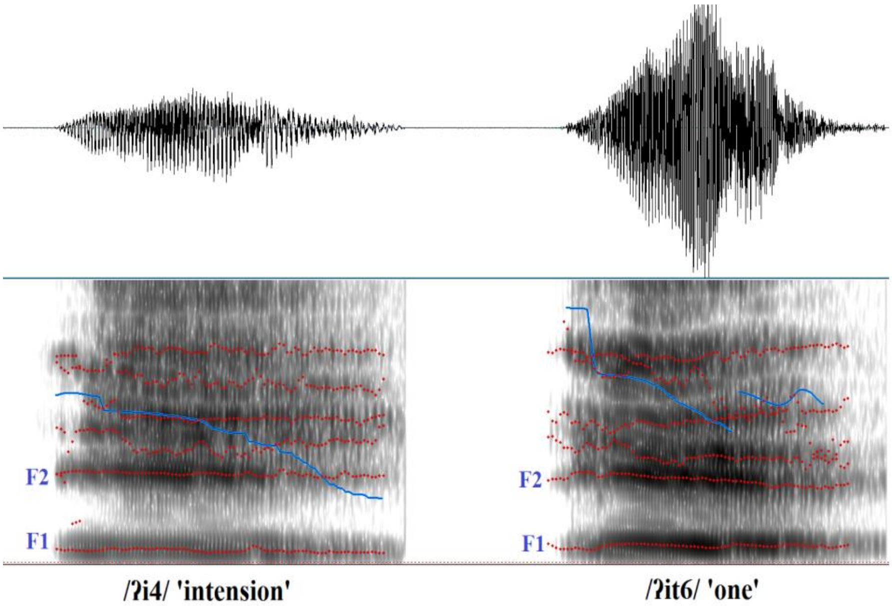


Figure 5-8. Acoustic realisations of high vowel /i/ in the unstopped tone 4 and stopped tone 6 (WYF, male).

In Figure 5-9, the two morphemes have the same high vowel /u/ at the underlying level with similar F0 realisations, but their formant patterns are quite different from each other. In the token /pu4/ ‘wealthy’, the F1 curve is essentially stable across the utterance while the F2 curve is largely steady but with reasonably slight fluctuation during the final portion. For example, the F1 and F2 values are 454 Hz and 956 Hz, respectively, at the 10% sampling point but are 394 and 909 Hz, respectively, at the 90% sampling point. The relatively stable formant patterns indicate the tongue largely remains at the same position during articulation; thus, this high back vowel /u/ is realised as a monophthong [u] in this unstopped tonal environment.

Further, in the token /kut6/ ‘bone’, both F1 and F2 curves show an upward change during the second half of the utterance. At the 10% sampling point, the F1 and F2 values are 392 Hz and 947 Hz, respectively, but at the 90% sampling point, the values are increased to 533 Hz and 1387 Hz, respectively. The increasing F1 indicates a downward movement of the tongue while the increasing F2 curve indicates a fronting tongue movement. Therefore, the production of this token involves two articulatory trajectories, and this high back vowel /u/ is realised as a diphthong [uɤ̯] rather than a monophthong [u] in this stopped tone.

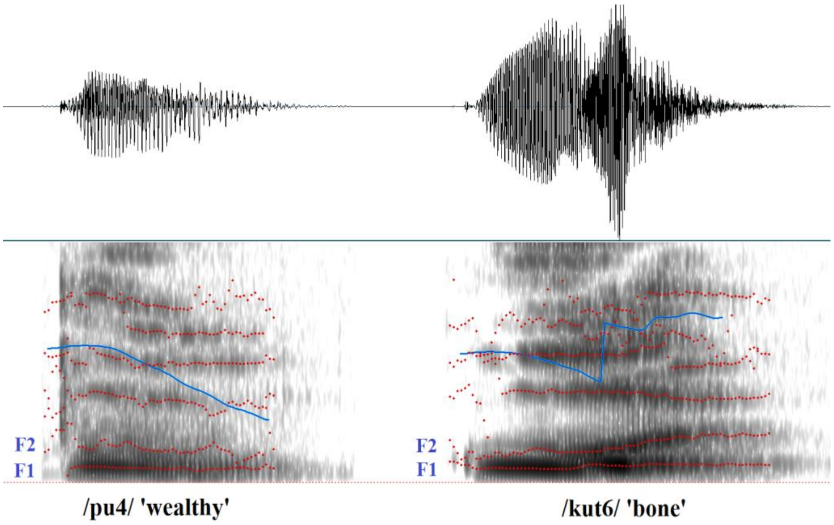


Figure 5-9. Acoustic realisations of high vowel /u/ in the unstopped tone 4 and stopped tone 6 (WYF, male).

In Figure 5-10, the two morphemes have the same underlying low vowel /ɐ/ and similar falling F0 contour at the surface level. Their F1 and F2 curves both present as relatively stable across time, except the first 10% portion shows a dynamic contour as an influence of the pre-vocalic articulation of the velar stop. The consistently similar and steady manifestations of formant patterns indicate a monophthong realisation of the low vowel /ɐ/ across both unstopped and stopped tones.

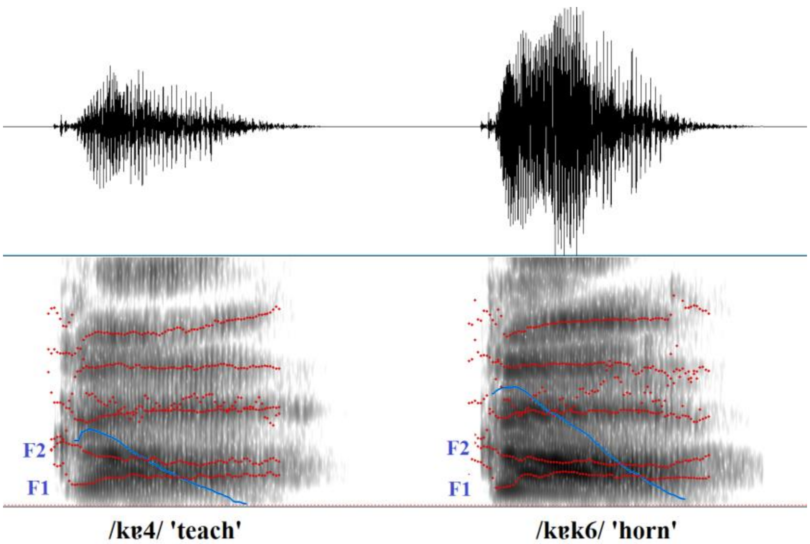


Figure 5-10. Acoustic realisations of low vowel /ɐ/ in the unstopped tone 4 and stopped tone 6 (WYF, male).

Therefore, high vowel diphthongisation with respect to stopped tones in Zhangzhou is basically supported acoustically and articulatorily. Changing the position of the tongue during the production of utterances gives rise to the acoustic correlates of formant patterns changing as a function of time. However, conducting a systematic acoustic quantification of vowel quality is imperative for a deeper understanding of the correlation between vowel quantity and tones, which has been proposed as a promising direction for further studies in Chapter 10.

5.3.3. Voice quality

In Chapter 4, the non-modal voice qualities of creaky and breathy were perceived as allophonic variants of modal phonation in Zhangzhou, with a distribution conditioned by both vowel quality and tone. The breathy voice is more likely to occur on the high vowels while the creaky phonation tends to occur on the low vowels in falling contours. From the articulatory point of view, the dynamic realisation of phonation with respect to vowel quality and tone in Zhangzhou effectively has robust articulatory bases (Ladefoged, 1971, 2003; Laver, 1980, 1994; Lotto et al., 1997; Thurgood, 2000; Gordon & Ladefoged, 2001), as shown in Table 5-7.

Table 5-7. Articulatory configurations for non-modal phonation production in Zhangzhou

Parameter		Creaky	Breathy
Correlation	pitch contour	falling	no restriction in unstopped tones
	vowel quality	low	high
Laryngeal	laryx	tense/raised	lax/lowered
	glottis	tightly adducted	abducted
	vocal fold structure	thick and compressed	weak medial compressed
	vocal fold vibration	irregular and slow	inefficient and loose
Supraglottal	supraglottal cavity	reduced	distended
	tongue root	retracted	advanced
	tongue body	lowered	raised
Transglottal airflow		low rate	high rate (frication noise)

From the acoustic point of view, the variation of articulatory configurations for the non-modal production gives rise to corresponding changes in the acoustic signals (Johnson, 2011; Baart, 2010; Ladefoged, 2003). The Figures 5-11, 5-12, and 5-13 show how the acoustic realisations of voice quality change with respect to the adjustments of laryngeal and supraglottal configuration and the mode of transglottal airflow for different vowels (high, mid, and low) and tonal production (rising, level, and falling) in Zhangzhou. The left panel of each figure shows the shape of glottal pulses while the right panel presents their corresponding spectral tilts, especially the amplitude differences between the first harmonic and the second harmonic (H1-H2), one of the most reliable acoustic parameters for phonation distinction (Stevens, 1977; Huffman, 1987; Holmberg et al., 1995; Andruski & Ratliff, 2000; Ladefoged, 2003; Baart, 2010; Johnson, 2011; Esposito, 2012). The waveforms and spectral slices are extracted from the last 10% of the related sounds.

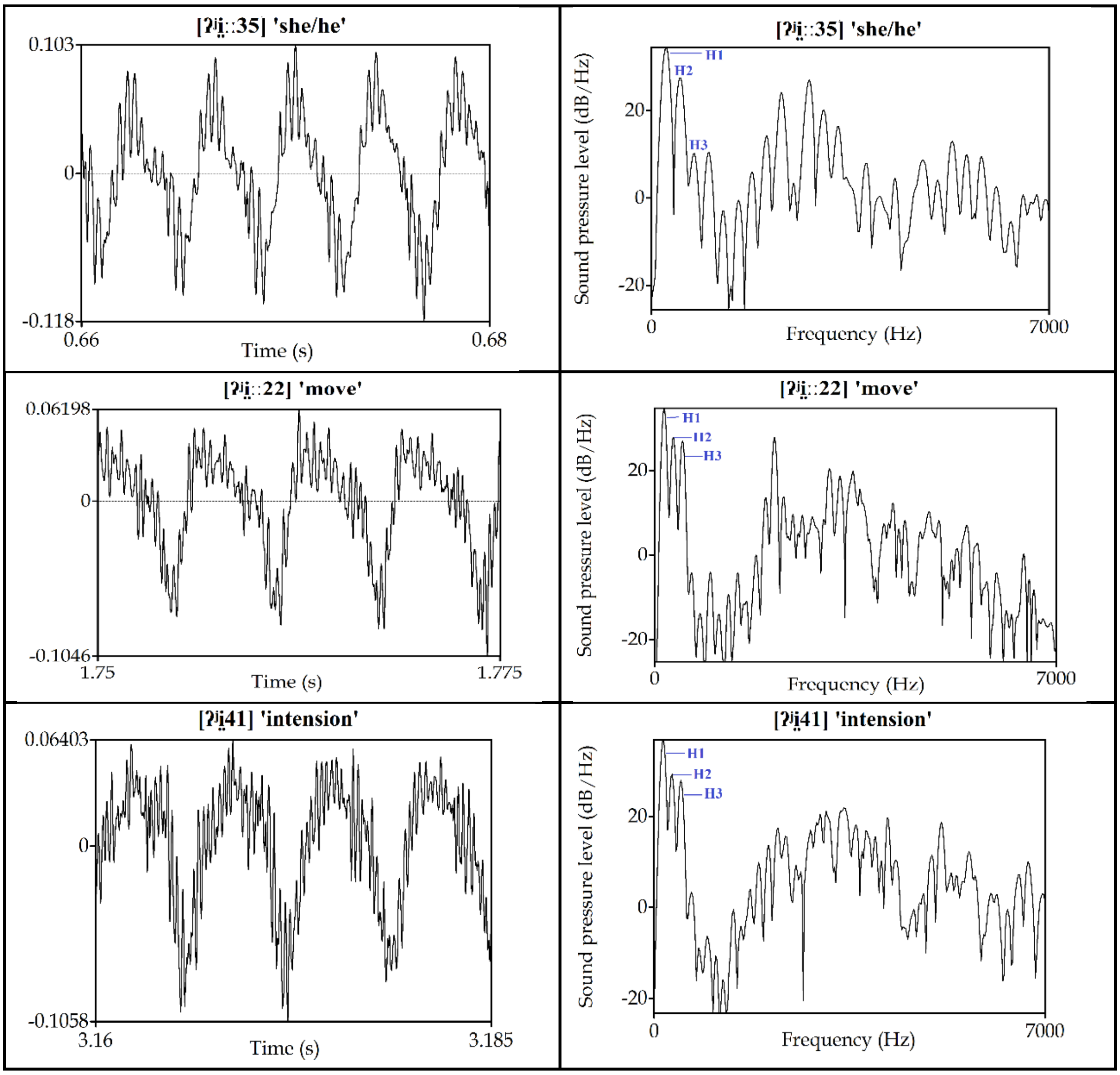


Figure 5-11. Phonation realisations of high vowel /i/ in rising, level, and falling F0 contexts (WYF, male).

Figure 5-11 shows the voice quality realisation of the high vowel /i/ in three different F0 environments. The waveforms on the left are always superimposed by random and rapid fluctuations across time, regardless of whether the F0 contour is rising, level, or falling. Such glottal frication occurs as a consequence of the persistent leakage of airflow passing through the abducted glottis during vibration (Ladefoged, 1971, 2003; Laver, 1980, 1994; Lotto et al., 1997; Thurgood, 2000; Gordon & Ladefoged, 2001). In addition, the amplitude of the first harmonic (H1) is always higher than that of any other harmonics in the spectra. In other words, the amplitude descends as a function of the increase of frequency. This positive spectral tilt of H1-H2 reflects an increased open quotient and a less abrupt glottal closing gesture because the vocal folds are abducted with an incomplete closure during each vibration cycle, causing energy to be distributed dominantly in the fundamentals and to decrease in the higher frequency regions (Stevens, 1977; Huffman, 1987; Holmberg et al., 1995; Andruski & Ratliff, 2000; Ladefoged, 2003; Baart, 2010; Johnson, 2011). Therefore, the high vowel consistently presents typical acoustic characteristics of breathy voice: glottal frication and steeply positive spectral tilt of H1-H2 in the three tonal environments.

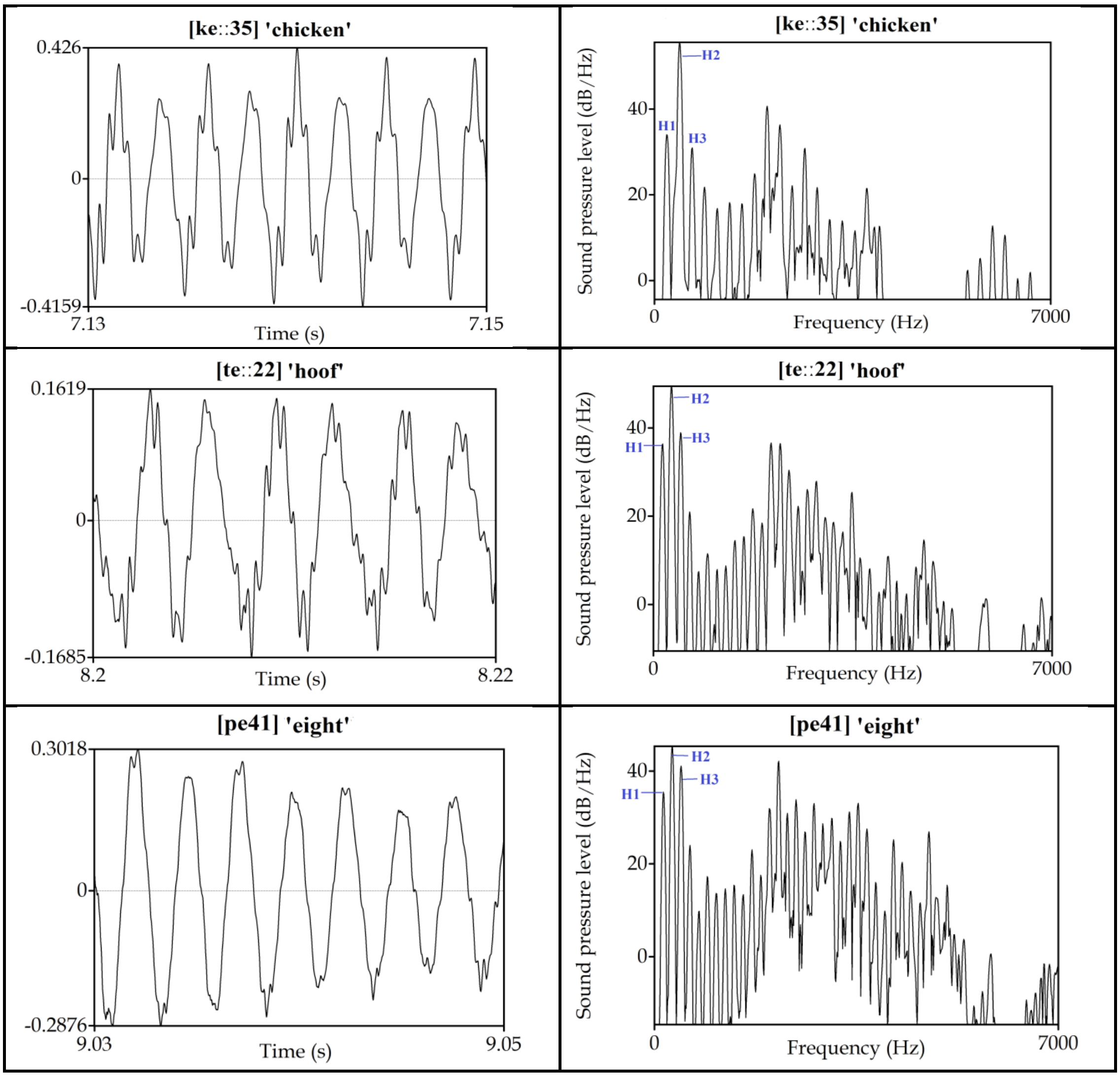


Figure 5-12. Phonation realisations of mid vowel /e/ in rising, level, and falling F0 contexts (WYF, male).

Figure 5-12 shows the voice quality realisation of the mid vowel /e/ across three different F0 contours. The intervals among glottal pulses are relatively evenly spaced, and the waveforms are not superimposed by intense and rapid frication across time or across different F0 contours. In the spectra, the amplitude of the second harmonic (H2) is always higher than that of its surrounding harmonics (e.g., H1 and H3). The acoustic manifestations indicate, during the production of the mid vowel, the vocal folds vibrate in a regular mode. Therefore, the mid vowel consistently presents acoustic characteristics of the modal phonation, regardless of the shape of F0 contours: regular glottal pulses and steeply negative spectral tilt of H1-H2.

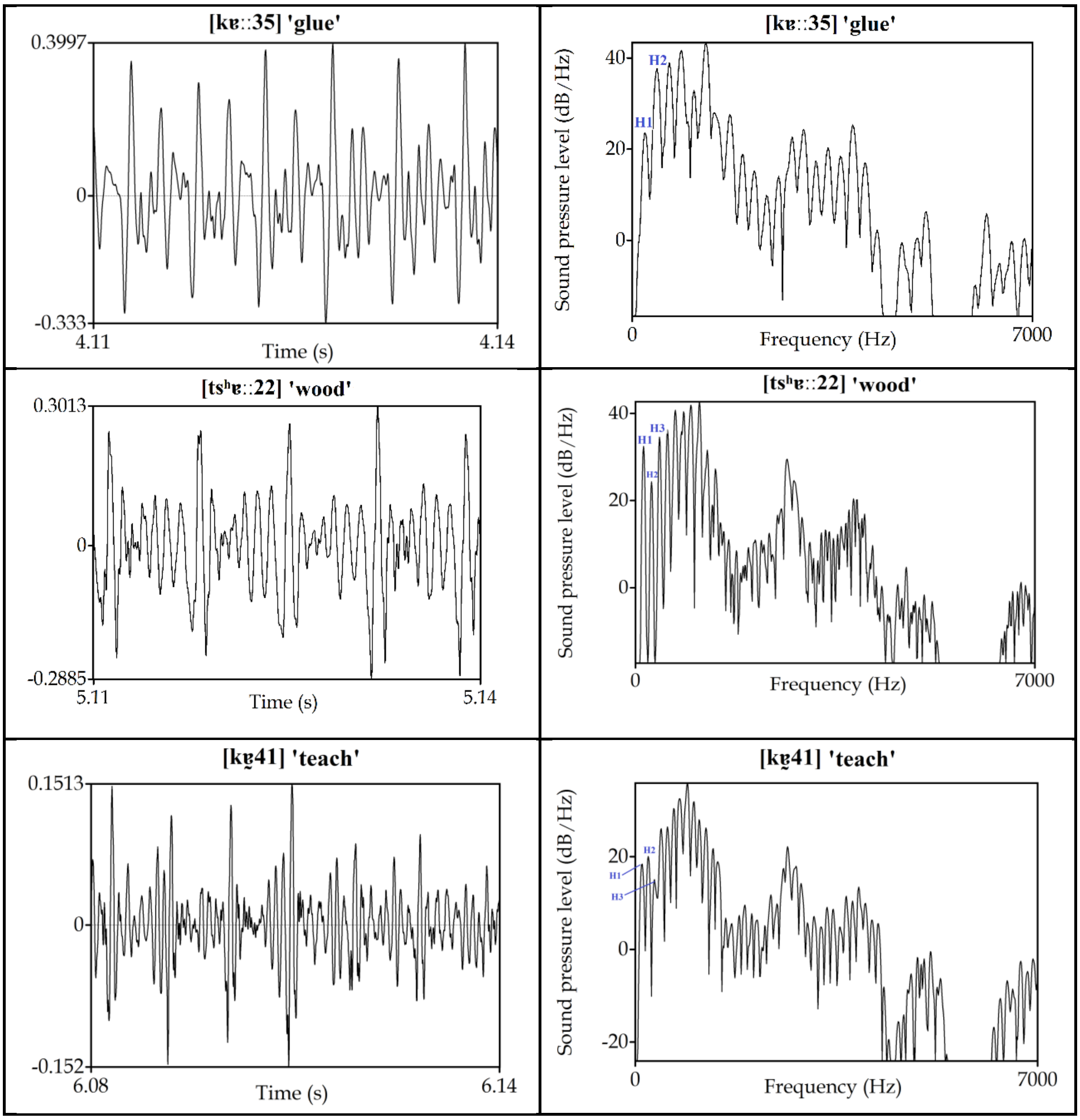


Figure 5-13. Phonation realisations of low vowel /ɐ/ in rising, level, and falling F0 contexts (WYF, male).

Figure 5-13 shows the voice quality realisation of the low vowel /ɐ/. In the rising contour [35], the glottal intervals in the waveform are regular and evenly spaced, and in the spectrum, the H2 is higher than the H1, indicating a modal voicing [ɐ] with regular vocal fold vibrations. In the level tonal environment [22], the glottal intervals are also regular, and no effect of glottal friction is seen in the waveform. In the spectrum, although the H1 tends to be higher than H2, unlike the breathy high vowel, whose amplitude descends as the frequency increases, the amplitude of this vowel basically increases across the low- and-mid frequency regions as shown in the rising contour environment. Thus, the low vowel in this context is articulated with a modal voice.

However, in the falling contour [41], both the waveform and spectrum vary differently. The vocal pulses are aperiodic and unevenly spaced, indicating the vocal folds vibrate at a relatively low and less frequent rate (Stevens, 1977; Huffman, 1987; Holmberg, et al., 1995). In the spectrum, the amplitude neither increases nor decreases dramatically during the first two harmonics but drops sharply in the third harmonic and increases across the low-mid frequency region. This spectrum indicates a decreased open quotient and a more precipitous closure gesture (Stevens, 1977; Huffman, 1987; Holmberg, et al., 1995). Therefore, the low vowel presents the acoustic manifestations of creaky voice in the falling contour: irregular glottal pulses and near-zero spectral tilt of H1-H2.

As indicated, the acoustic manifestations of waveform and spectral tilt appear able to distinguish three different types of phonation in Zhangzhou. The manifestations of spectral tilt for the non-modal phonation of breathy and creaky are consistent with previous findings in many other languages. For example, White Hmong (Esposito, 2012), Jalapa Mazatec (Garellek & Keating, 2011), Green Mong (Andruski & Ratliff, 2000) and San Lucas Quiavini Zapotec (Gordon & Ladefoged, 2001) all show a steeply positive value of H1-H2 for the breathy voice, and a small magnitude around zero for the creaky phonation, as found in Zhangzhou.

However, the magnitude of H1-H2 for the modal phonation tends to be variable across languages as a consequence of various constraints, including vowel quality and gender. For example, in Green Mong, the average H1-H2 amplitude for modal voice varies across vowels with a negative value for high vowel /i/, a negative value for back vowel /u/, while a value near zero for low vowel /a/ (Andruski & Ratliff, 2000). In the language of Yalálag Zapotec, female speakers tend to have a positive value for the modal voice while male speakers are more likely to present a negative value (Avelino, 2010). In Zhangzhou, the split tilt of modal sound also presents a language-specific property, for example, it consistently shows a negative value across mid vowel /e/ of different F0 contours, and across low vowel /ɐ/ of rising F0 contour in this male speaker (WYF)'s production. A negative H1-H2 value can thus be considered as a relatively reliable cue to distinguish modal from non-modal phonation.

In summary, the changes of voice quality with respect to vowel quality and tone in Zhangzhou are understandable from the perspectives of acoustic and auditory phonetics, and are reliable concomitants of realisations of Zhangzhou tones. However, conducting sophisticated acoustic quantifications of various parameters, including jitter, shimmer, cepstral peak prominence, H1-A1, and other dimensions, as well as articulatory experiments is imperative for a fuller understanding of the function of phonation in Zhangzhou tonal production, which is a promising direction for future studies.

5.3.4. Syllable coda

Chapter 4 indicated the obstruent codas tend not to be realised in the stopped tones in the monosyllabic setting, but the bilabial coda is occasionally perceivable. From the articulatory point of view, the non-realisation of obstruent codas indicates the appropriate active articulator is not manipulated to form a complete oral constriction with the passive articulator at a particular point along the vocal tract (Ladefoged, 1971; Laver, 1994; Bickford & Floyd, 2006; Gick et al., 2013).

From the acoustic point of view, it indicates the formant patterns are not dramatically transited to a particular frequency range as a consequence of the influence by the postvocalic obstruent production (Stevens & Klatt, 1974; Reetz & Jongman, 2009; Baart, 2010).

Figure 5-14 shows the acoustic realisation of Zhangzhou obstruent codas in the monosyllabic setting. The three morphemes have the same low vowel /ɐ/ and tone 7 at the underlying level, but differ in types of obstruent codas. As shown in the spectrograms, the formant patterns are largely steady without significant changes across the three morphemes. No formant transition can be seen during the last 5% portion of the curves, although a slight degree of fluctuation occurs at the beginning of F1 and F2 because of the perturbation effect of syllable onset production. The steady manifestations of acoustic formant patterns indicate the related vowels are realised as monophthongs and no obstruent coda is articulated following the vowels, supporting the auditory observation of non-realisation of obstruent coda in the monosyllabic setting of Zhangzhou.

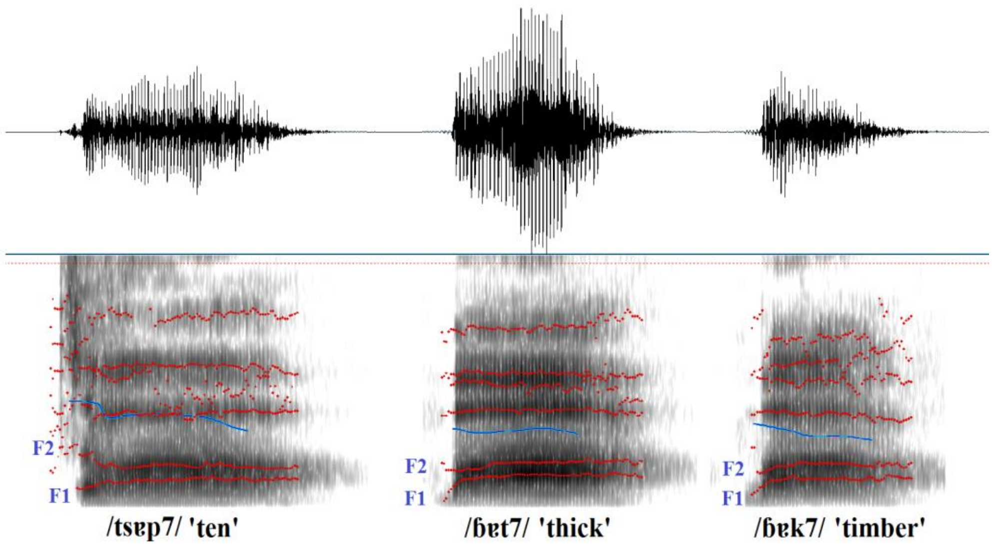


Figure 5-14. Acoustic realisations of obstruent codas in the monosyllabic setting (WYF, male).

On the contrary, if the obstruent coda were realised, the formant transitions would be expected to occur on the spectrogram (Reetz & Jongman, 2009; Baart, 2010). For example, in the /tsɐp7/ morpheme, the F2 must have had a falling tendency towards the frequency range around 800 Hz as an influence of the bilabial coda production. Similarly, in the /ɓɐt7/ morpheme, the F2 must have risen to somewhere around 1600 Hz to indicate an alveolar coda; however, no rising trend can be seen on the F2 curve.

The non-realisation of obstruent codas on the other hand can cause several phonetic effects on the whole syllables, which include syllable lengthening, high vowel diphthongisation, F0 contour depression, and vowel laryngealisation.

(1) The syllable lengthening effect can be seen in Figure 5-15. In comparison to its length realisation in the phrase-initial position of the disyllabic token /ɓɐk7.tsʰjɔ̃5/, the morpheme /ɓɐk7/ apparently has a much longer duration when it is produced in isolation.

(2) The high vowel diphthongisation effect can be seen in Figures 5-8 and 5-9 in Section 5.3.2, where both F1 and F2 present dynamic changes in the second half of the curves.

(3) The F0 contour depression effect can also be seen in Figure 5-14, where the blue curves representing the F0 are largely level during the first half but suddenly stop without any moving trend apparent during the second half of the three morphemes. Such a discontiguous F0 contour is generally regarded as one important acoustic manifestation of creaky voice because the vocal folds at this stage are tightly adducted and create a thick and compressed mass, causing an irregular and lower frequency vibration mode and a failure to extract the F0 contour.

(4) Vowel laryngealisation can be seen in Figure 5-16, in which the individual spectra were taken from the last 10% portion of each tokens as shown in Figure 5-14. The amplitude difference between the first two harmonics (H1-H2) is neither steeply positive, as shown for the breathy voice, nor steeply negative, as shown for the normal voice. Instead, the H1 value is either similar to the H2 value or slightly higher than H2, generally indicating the voice quality of creaky.

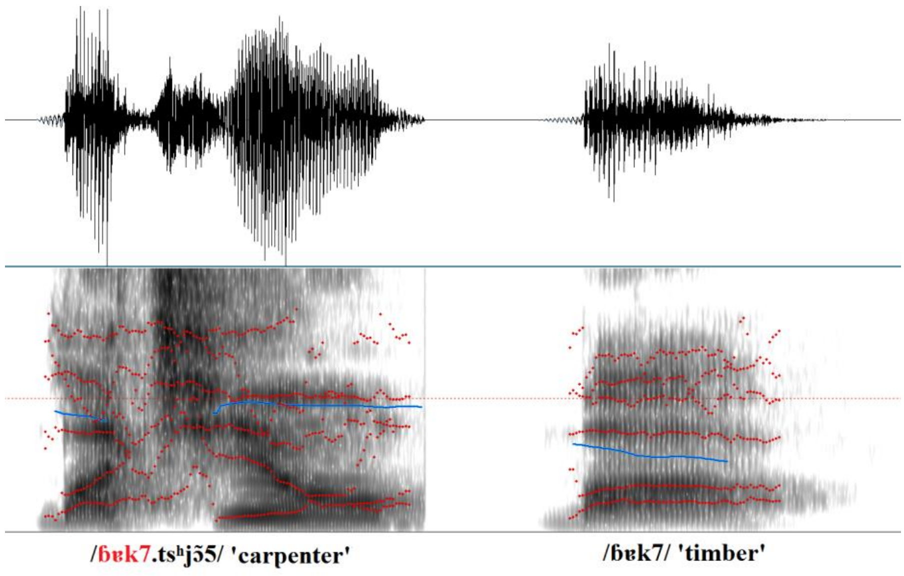


Figure 5-15. Acoustic demonstration of syllable lengthening in Zhangzhou (WYF, male).

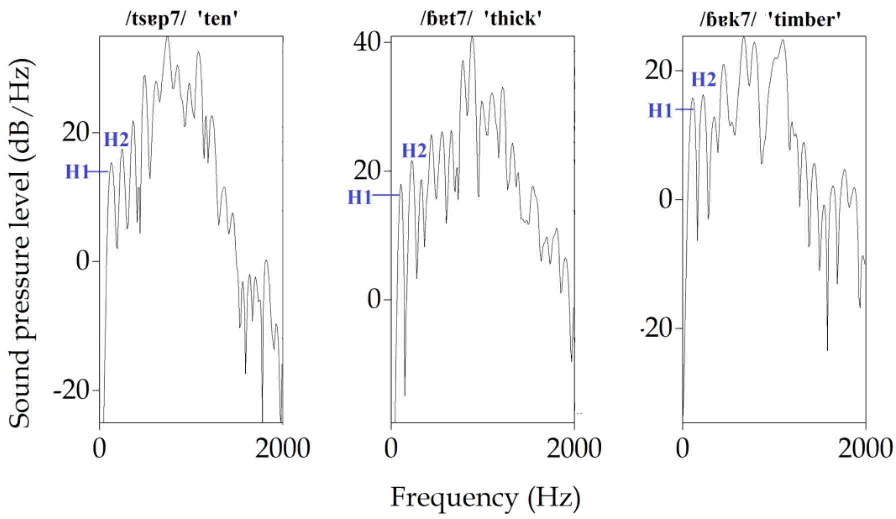


Figure 5-16. Acoustic manifestation of vowel laryngealisation in Zhangzhou (WYF, male).

5.3.5. Summary

According to the descriptions in this section, tone does not exclusively utilise F0 contour and height to distinguish otherwise segmentally similar lexical items in Zhangzhou. Rather, a bundle of other segmental (e.g., vowel quality and obstruent coda) and suprasegmental (e.g., duration and voice quality) parameters function together with F0 to realise the category that is called "tone". Table 5-8 shows the multidimensional realisations of Zhangzhou citation tones, including six dimensions. The mid vowels are largely produced with a modal voice across tones to most speakers, thus no column is particularly made to denote that. For example, tone 1 indicates a rising contour, extra long length, no diphthongisation, breahy high vowel, modal-voiced low vowel, and sonorant coda; while tone 6 suggests a mid-high falling contour, long length, high vowel diphthongisation, creaky vowel, and not-realised obstruent coda. Nevertheless, questions remain as to how these citation tones are related to each other from the phonological point of view.

To address the issue, the multidimensional properties for each citation tone were firstly transformed into a multiple sequence alignment to be computed in the SplitsTree software in order to generate a phylogram to visualise the internal relations among these citation tones. Figure 5-17 shows the generated result. The root of the tree represents the set of citation tones being investigated. The horizontal lines are branches representing the amount of similarity in terms of multidimensional realisations shared by the citation tones. The shorter the branch, the stronger the similarity that tones share in phonetics. The vertical lines indicate the relatedness between citation tones. The more divergent the lines, the higher the probability that the tones are not related categorically. For example, citation tones 1 and 7 are least related as they are distributed relatively far from each other. The code used to generate this phylogram is provided in the attached USB. Additionally, two notable aspects deserve further discussion.

- The tones that share similar F0/pitch realisation, for example, tones 4 and 6, as well as tones 2 and 7, are categorically different on the basis of the multidimensional framework. This fact further justifies the statement that the single dimension of F0/pitch is not sufficient to characterise Zhangzhou tones.
- Tonal contrast neutralisation can occur in the monosyllabic setting. Tones 2 and 8 are clustered together without significant differences in terms of the multidimensional parameters. The reason why they are considered two independent categories is primarily because they have different realisations in the multi-syllabic contexts, which will be discussed in later chapters.

Therefore, the multidimensional realisations of Zhangzhou citation tones are not only auditorily supported but also grounded in both acoustic and articulatory reality.

Table 5-8. Multidimensional realisations of Zhangzhou citation tones

Tone	F0	Length	Diphthong	High vowel	Low vowel	Syllable coda
1	[35]	extra long	-	breathy	modal	sonorant
2	[22]	extra long	-	breathy	modal	sonorant
3	[51]	medium	-	breathy	creaky	sonorant
4	[41]	medium	-	breathy	creaky	sonorant
5	[33]	extra long	-	breathy	modal	sonorant
6	[41]	short	+	creaky	creaky	nonrealised obstruent
7	[221]	long	+	creaky	creaky	nonrealised obstruent
8	[22]	extra long	-	breathy	modal	sonorant

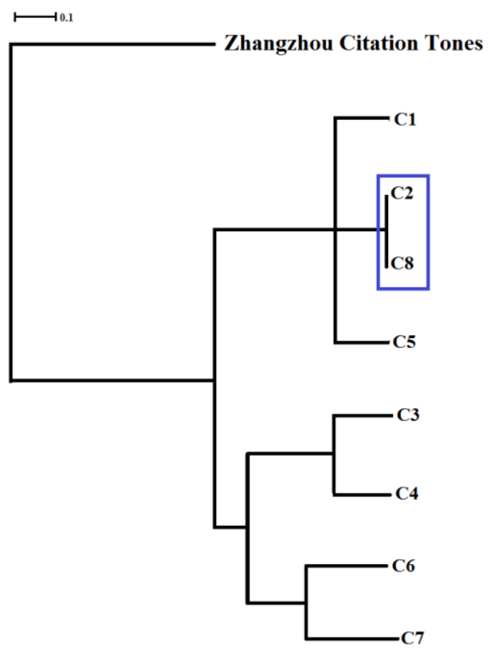


Figure 5-17. Phylogram representation of Zhangzhou citation tones (C=Citation).

5.4. Previous Studies

Previous descriptions of Zhangzhou citation tones have been largely impressionistic until, in the most recent decade, a few pioneering studies (Yang, 2008; 2014; Yin, 2009; Huang et al., 2016) started to explore the acoustic characteristics of Zhangzhou tones, dominantly investigating F0 realisation. Nevertheless, the studies exhibited several problematic aspects in terms of research design and analysis. They generally classified synchronic Zhangzhou citation tones in terms of Middle Chinese tonal categories; therefore, they all defaulted to a system of seven tones as the split and merger of MC Yang Shang tone. This study by contrast finds an eight-tone system by examining the realisations of diverse parameters across three different contexts—isolation, phrase-initial, and phrase-final.

In addition, Yang (2008; 2014) investigated the F0 system of Zhangzhou citation tones simply by comparing the raw F0 values from one male and one female speaker without quantification and normalisation. Yin (2009) took into account the normalisation approach, but the study addressed only the tonal F0 in unstopped syllables while neglecting the tones in stopped syllables and the interaction between F0 and other phonetic parameters. Huang et al. (2016) investigated the tonal F0 realisation of Zhangzhou citation tones using *z*-score normalisation of raw F0 values from 21 speakers, a stepforward in tonal studies of Zhangzhou. Nevertheless, that study did not address how other phonetic parameters contribute to make tonal distinctions in Zhangzhou, as discussed and described in this project, although the variation of F0 duration across different tones was briefly mentioned by Huang et al. (2016).

5.5. Conclusion

In this chapter, the acoustic and statistical descriptions, supplemented by articulatory explanations, support the auditory observations made in Chapter 4 concerning Zhangzhou tonal realisations being multidimensional, as well as the phonological tone in Zhangzhou being effectively a complex of phonetic features. The objective descriptions and analyses in this chapter help improve the accuracy of phonetic transcriptions of Zhangzhou tones and, to a large extent, help avoid many sources of human and machine errors. The quantified, normalised, statistical, and visual results provide a solid foundation for investigating the realisations of disyllabic tones and the mapping between citation and disyllabic tones in the rest of this thesis.

Part C: Zhangzhou Disyllabic Tones

Chapter 6: Auditory Properties of Zhangzhou Disyllabic Tones

Chapters 4 and 5 addressed the auditory and acoustic properties of Zhangzhou citation tones. This chapter is intended to encode and understand how these individual tones interact in constructions beyond monosyllables and how the segmental and suprasegmental parameters contribute to the tonal distinctions in polysyllabic settings from the perspective of auditory phonetics. The general properties of Zhangzhou tone sandhi, including the dominancy and specificification of sandhi domain are also discussed.

6.1. General Properties of Zhangzhou Tone Sandhi

This section addresses the following two questions. (1) What is the dominancy of Zhangzhou tone sandhi: right dominant, left dominant, or both? (2) What is the sandhi domain of Zhangzhou tonal alternation: metrical foot, word, phrase, or something else? (3) How do the realisations of disyllabic tone reflect the interaction among different linguistic levels?

6.1.1. Right-dominant

Zhangzhou presents a right-dominant tone sandhi system. The number of parametric contrasts is largely preserved on the rightmost tones (see Section 9.2), but is largely changed and reduced to some degree by processes of neutralisation on the non-rightmost tones (see Section 9.1). In addition, the realisations of the rightmost tones are relatively consistent with those of their corresponding citation forms in terms of categories but are not straightforwardly the same because of their phonetic sensitivity to surrounding contexts (see Sections 6.4 and 9.2). The realisations of the non-rightmost tones, on the other hand, are largely changed to forms different from their citation realisations at both the phonological and the phonetic levels (see Sections 6.3 and 9.1).

This right dominancy can be illustrated in Table 6-1 about the pitch realisation of eight individual tones before tone 2 (Pattern A) and after tone 2 (Pattern B). Segments are given their phonemic transcriptions. The citation column shows the pitch values of the eight tones when the related morphemes are produced in isolation. The non-right column presents the pitch values of the non-right morphemes in those disyllabic phrases, and the right column shows the pitch values of the rightmost morphemes in those disyllabic constructions.

As Pattern A shows, the pitches of the tones in the non-right dominant context are all different from their corresponding citation values. For example, the pitch of tone 3 in pattern 3+2 is changed to a mid-rising [25] from a high-falling contour [51] in citation. Further, processes of pitch neutralisation occur in this context. For example, tones 1 and 2 have similar realisation of mid level [33]; while tones 5, 7 and 8 have similar mid-falling contour [32] before tone 2.

Table 6-1. Examples of the right dominancy of Zhangzhou tone sandhi

A	Non-right	Citation	Example	B	Right	Citation	Example
1+2	[33]	[35]	/tsʰɛ̃1.tɛ2/ 'raw tea'	2+**1**	[34]	[35]	/tɛ2.hwɐ1/ 'camellia'
2+2	[33]	[22]	/ʔɐŋ2.tɛ2/ 'black tea'	2+**2**	[211]	[22]	/tɛ2.ɗɐw2/ 'tea
3+2	[25]	[51]	/tsɐ3.tɛ2/ 'morning tea'	2+**3**	[52]	[51]	/tɛ2.ɓi3/ 'dried tea'
4+2	[63]	[41]	/swɐ̃4.tɛ2/ 'unpacked	2+**4**	[41]	[41]	/tɛ2.tjɐm4/ 'tea store'
5+2	[32]	[33]	/ʔjɔŋ5.tɛ2/ 'have tea'	2+**5**	[33]	[33]	/tɛ2.tsʰju5/ 'tea tree'
6+2	[65]	[41]	/sip6.tɛ2/ 'moisten tea'	2+**6**	[41]	[41]	/tɛ2.sik6/ 'tea colour'
7+2	[32]	[221]	/sik7.tɛ2/ 'colourful tea'	2+**7**	[211]	[221]	/tɛ2.sit7/ 'tea dessert'
8+2	[32]	[22]	/pɛ8.tɛ2/ 'Bai tea'	2+**8**	[211]	[22]	/tɛ2.ɦjɵ8/ 'tea leaf'

In Pattern B, the pitches of the rightmost tones are very similar to their corresponding citation forms but present a certain degree of differences that are reasonably ascribed to the phonetic influence of their preceding pitch contours and the pitch declination effect on the rightmost position. For example, tone 2 is realised as [211] in the rightmost position but as [22] in citation, which is supposed to be affected by the effect of final-position declination. This finding comes into conflict with the general assumption that considers the rightmost tone is straightforwardly related to the citation tones, with the citation-tone pitch values preserved and unchanged (Wright, 1983; Shih, 1986; Ballard, 1988; Chen, 2000; Zhang, 2007; Rose, 2016).

As discussed with respect to the pitch realisation, Zhangzhou tones present typical properties of a right-dominant tone sandhi system, including paradigmatic pitch alternation and processes of neutralisation in the non-right context, as well as categorical preservation of pitch values in the right-most context. The pitch values of rightmost tones appear sensitive to the phonetic environments and present variable but predictable realisations; however, to which degree the realisation is affected needs to be further tested in statistics. Chapters 7 and 8 will present the parametric realisations of Zhangzhou citation tones from the acoustic and statistical points of views. It is hopped it will shed more light on the right-dominancy of Zhangzhou tone sandhi.

6.1.2. Syntactic boundary-relevant

Zhangzhou tone sandhi is syntactically relevant. The prosodic domain of tone sandhi in Zhangzhou appears not to be phonologically specified, but rather is syntactically determined. The rightmost boundary of the sandhi domain in Zhangzhou is found being aligned with the rightmost boundary of a syntactic phrase XP, where X is categorically variable, ranging over adjective (A), noun (N), verb (V), preposition (P), adverb (Adv), and other syntactic elements, as illustrated in Figure 6-1 (C=citation, S=sandhi). The sentences were elicited from conversations with the female speaker HYH with respect to her socio-background investigation in the field site.

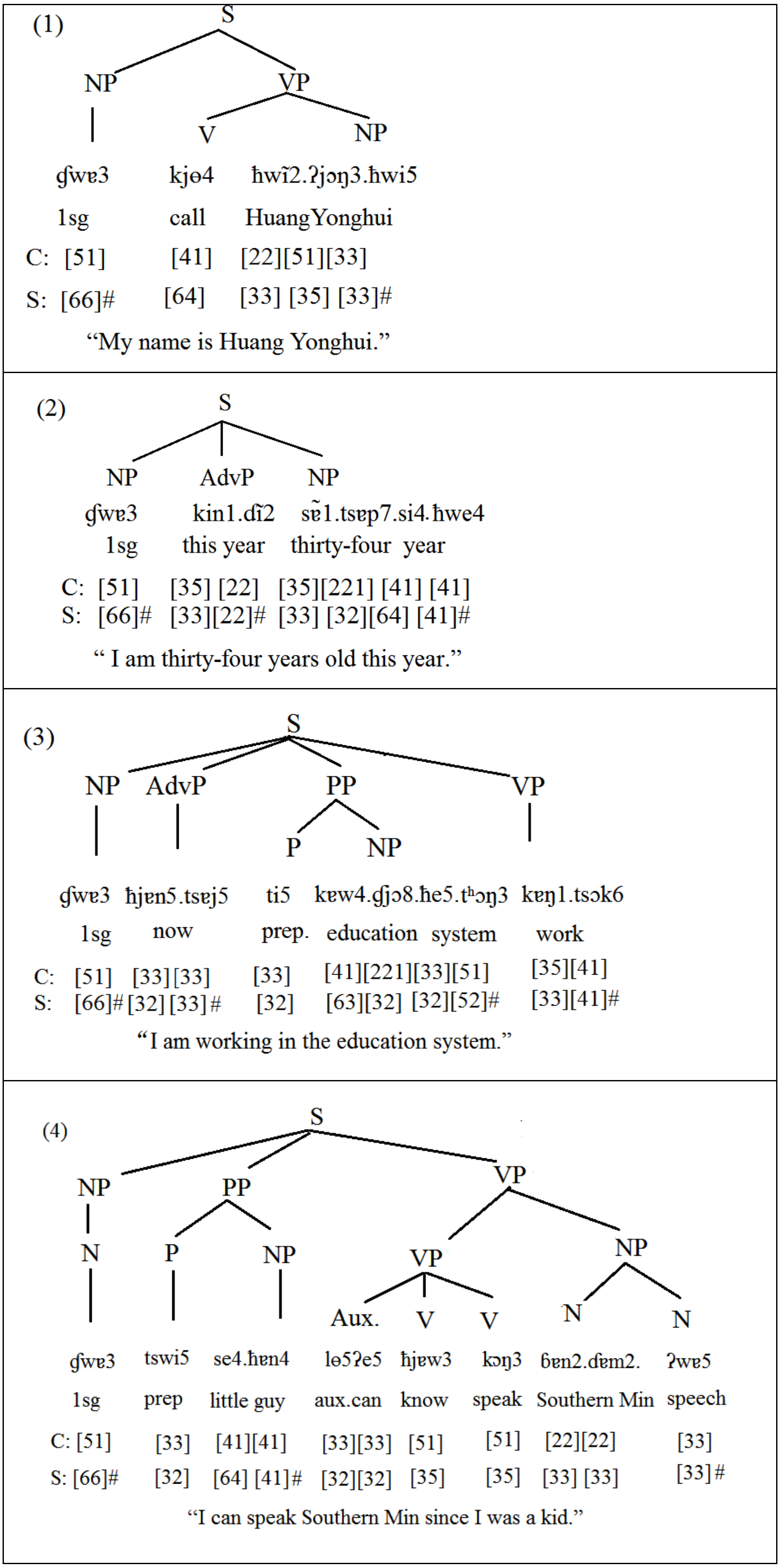


Figure 6-1. Examples of syntactically-determined tone sandhi domain in Zhangzhou.

As indicated in Figure 6-1, the pitches of the tones at the rightmost boundaries of syntactic phrases are largely the same as their corresponding citation values, while the values of non-phrase-final tones are all alternated to the forms significantly different from their citation forms regardless of the number of syllables within the phrases. The pitch value of the monosyllabic subject (/ɠwɐ3/ '1sg.') is raised up to be an extra high level although it is at the phrase-final boundary; however, it is reasonably assumed to result from the intonational effect in order to initiate a topic. It is thus plausible to propose that Zhangzhou tone sandhi domain is syntactically-relevant.

This finding supports the general assumption of a close interface between phonology and syntax for Southern Min tone sandhi as found in Xiamen and Taiwan (Chen, 1987; Lin, 1994; Zhang, 1994). Nevertheless, how the sandhi domain (known as tone group) can be defined has not yet reached a general consensus in the literature (see Section 1.6.1). For example, Chen (1987) and Zhang (1994) ascertained the tone sandhi process does not appear sensitive to the syntactic phrase XP that acts as an adjunct in natural speech. Lin (1994) claimed the prosodic domain for Southern Min tone sandhi is lexically governed. Exploring whether tone sandhi may occur in polysyllabic phrases that function as adjuncts in sentences in Zhangzhou, and how the sandhi domain should be better specified will be linguistically interesting and promising, but a fuller discussion has to be postponed until another occasion because of the scope limitation of this project.

6.1.3. Syntactic structure-irrelevant

While the tone sandhi domain in Zhangzhou tends to be syntactically relevant, it is only relevant to the boundaries of syntactic phrases, and appears irrelevant to the internal structure of syntactic phrases. Within a sandhi domain, the realisations of the tones are not sensitive to the categories of syntactic phrases, as shown in Table 6-2 about the pitch realisation of tone 3 in disyllabic phrases of 3+3 pattern, and in Table 6-3 about the pitch realisation of tone 3 in trisyllabic phrases of 3+3+3 pattern. The phrases differ from each other with respect to internal structures. As seen in Table 6-2, the pitches of phrase-initial morphemes are all alternated to a mid-rising contour [24] from a high-falling [51] citation contour across the four disyllabic phrases of different internal structures. Similarly, in Table 6-3, the pitches of the first morphemes are all alternated to a mid-rising [35] contour, while the pitches of the second morphemes are changed to a low-rising [24] contour regardless of the categorical differences of those trisyllabic phrases.

Table 6-2. Examples of the irrelevance to syntactic categories of Zhangzhou disyllabic tone sandhi

Pattern	Sandhi pitch	Citation pitch	Example	Syntactic category
3+3	[**24**+52]	[**51**+51]	/ħɐj3.tswi3/ 'seawater'	NP
3+3	[**24**+52]	[**51**+51]	/se3.ɓɛ3/ 'wash horse'	VP
3+3	[**24**+52]	[**51**+51]	/ħɵ3.ħɵ3/ 'very well'	AP
3+3	[**24**+52]	[**51**+51]	/kwɐ̃3.kin3/ 'hurry up'	AdvP

Table 6-3. Examples of syntactic irrelevance of Zhangzhou trisyllabic tone sandhi

Pattern	Sandhi	Citation	Example	Syntactic category
3+3+3	[**35**+**24**+51]	[**51**+**51**+51]	/ħɐj3.ɓɛ3.tsju3/ 'seahorse wine''	NP
3+3+3	[**35**+**24**+51]	[**51**+**51**+51]	/kɛ̃3.si3.si3/ 'cowardly'	AP
3+3+3	[**35**+**24**+51]	[**51**+**51**+51]	/kwɐ̃3.kin3.tsɐw3/ 'run away at once"	AdvP
3+3+3	[**35**+**24**+51]	[**51**+**51**+51]	/ɓe3.ħɵ3.tsju3/ 'buy good wine'	VP

As described, the pitch height of the first morphemes of trisyllabic phrases appears to be slightly higher from that of the initial morphemes of disyllabic phrases and that of the second morphemes of trisyllabic phrases although their contour shapes are similar; however, it is reasonably assumed to be a consequence of the raising effect of the utterance-initial context. The two patterns demonstrate that the change in the category of syntactic phrases does not cause any corresponding change in the tonal realisations within the domain. This impressionistic assertion can be justified by the acoustic quantified and normalised results from 21 speakers as shown in later chapter 7. For example, as seen in the Appendix B, the tonal pattern 1+7 involves four different syntactic phrases, including NP, AdvP, VP, and AP; however, no divergent F0 pattern emerges owing to internal differences of these syntactic structures, as reflected in Figure 7-1 (Section 7.1.1).

Thus, the property of syntactic-structure irrelevance of Zhangzhou tone sandhi has impressionistic and acoustic bases. In this regard, Zhangzhou is different from many Chinese dialects which are sensitive to the syntactic structures and present different tonal patterns (see Section 1.5.2). For example, in Pingyao of the Jin dialect, the verb-object phrases (VP) have a different sandhi pattern from the modifier-noun phrases (NP; Zhang, 2014).

6.1.4. Phonologically-inert

Zhangzhou tone sandhi is phonologically-inert. The parametric realisations tend not to be affected by the categories of their surrounding tones from the auditory point of view. They appear to have very similar realisations across tonal categories surrounding them. As indicated in Table 6-4, tone 3 consistently has a low-rising pitch contour in the phrase-initial position (Pattern A), regardless of whether its following tone has a rising, level, or falling pitch contour. Similarly, the pitch of tone 3 is consistently realised as a high-falling contour (Pattern B) in the phrase-final position in spite of the contour shape of preceding tones.

This impressionistic assertion can also be justified in acoustics and statistics. For example, as reflected in Figure 7-5 (Section 7.1.3), the normalised F0 contour of tone 3 in the phrase-initial context is consistently mid-low rising across different following tones. This phonologically-inert property of Zhangzhou tone sandhi is different from many Chinese dialects. For example, in Mandarin, the well-known third tone sandhi is typically phonologically-conditioned.

Table 6-4. Examples of phonologically-inert property of Zhangzhou tone sandhi

A	Sandhi pitch	Example	B	Sandhi pitch	Example
3+1	[**24**+34]	/ɓɛ3.kʰɐ1/ 'horse leg'	1+3	[33+**52**]	/ʔɔ1.ɓɛ3/ 'black horse'
3+2	[**25**+311]	/ɓɛ3.te2/ 'horse hoof'	2+3	[33+**52**]	/kʰjɐ2.ɓɛ3/ 'ride a horse'
3+3	[**24**+52]	/ɓɛ3.sɐj3/ 'horse shit'	3+3	[24+**52**]	/pɵ3.ɓɛ3/ 'BMW'
3+4	[**24**+51]	/ɓɛ3.tsʰwi4/ 'horse mouth'	4+3	[63+**52**]	/tsjɐn4.ɓɛ3/ 'battle horse'
3+5	[**24**+43]	/ɓɛ3.pɔ5/ 'horse step'	5+3	[32+**52**]	/tsʰi5.ɓɛ3/ 'feed horse'
3+6	[**24**+51]	/ɓɛ3.kut6/ 'horse bone'	6+3	[65+**52**]	/kɐk6.ɓɛ3/ 'gnu, horn horse'
3+7	[**25**+311]	/ɓɛ3.sut7/ 'equestrianism'	7+3	[32+**52**]	/ɓɔk7.ɓɛ3/ 'wooden horse'
3+8	[**25**+311]	/ɓɛ3.tsi8/ 'horse tongue'	8+3	[32+**52**]	/pɛ8.ɓɛ3/ 'white horse'

For a better understanding of the phonologically-inert property of Zhangzhou disyllabic tones, Chapters 7 and 8 will provide detailed descriptions and discussions with respect to two specific research questions. (1) What are the individual tones realised as a function of their surrounding tones, and (2) are their realisations affected by the tones nearby? If so, to what extent are they affected, and what conditions the variations?

6.1.5. Phonetically-sensitive

While the tonal realisations in multisyllabic settings tend not to be categorically affected by their surrounding tones, the realisations, especially the pitches of phrase-final tones, are phonetically sensitive to surrounding environments and show diverse variants. As shown in Table 6-4, the phrase-initial tone 3 has two variants of [24] and [25]. Its pitch offset appears slightly higher, When it precedes a tone having a low pitch contour (e.g., [311] of tone 2, 7 or 8). Similarly, as indicated in Table 6-5, the phrase-final tones 4 and 8 both present two variants. The pitch of phrase-final tone 4 is generally realised as a mid-high falling [41] contour but is realised as a high falling [51] after tones 3 and 6 that have a pitch offset higher than the mid point. The phrase-final tone 8 has a higher onset [311] occur when the preceding tone has a mid level contour or an extra-high and short falling contour.

The property of phonetical sensitivity of Zhangzhou tones in multisyllabic context also has a robust basis in acoustics and statistics. For example, as reflected in Figure 7-17 (Section 7.1.9), most of phrase-final tones have two variants that are statistically significantly different.

Table 6-5. Examples of phonetic sensitivity of Zhangzhou tones

A	Sandhi pitch	Example	B	Sandhi pitch	Example
1+4	[33+41]	/sɛ̃1.tʰi4/ 'raw iron'	1+8	[33+311]	/ɗiŋ1.pɛ8/ 'milk white'
2+4	[33+41]	/tɐŋ2.tʰi4/ 'copper iron'	2+8	[33+311]	/sun2.pɛ8/ 'pure white'
3+4	[24+51]	/ħɵ3.tʰi4/ 'good iron'	3+8	[24+311]	/ɓi3.pɛ8/ 'rice white'
4+4	[63+41]	/ħwi4.tʰi4/ 'scrap iron'	4+8	[63+211]	/pʰjɵ4.pɛ8/ 'whiten'
5+4	[32+41]	/te5.tʰi4/ 'subway'	5+8	[32+211]	/ɗwĩ5.pɛ8/ 'albumen'
6+4	[65+51]	/kʰip6.tʰi4/ 'magnet'	6+8	[65+311]	/kɔk6.ʔjɵ8/ 'national medicine'
7+4	[32+41]	/tɔk8.kʰi4/ 'toxic gas'	7+8	[32+211]	/ɗik7.pɛ8/ 'green in white'
8+4	[32+41]	/pɛ8.tʰi4/ 'white iron'	8+8	[32+211]	/pɛ8.pɛ8/ 'white'

The phonetic-sensitive property of Zhangzhou disyllabic tones also reflects the well-known across-linguistic phenomenon of tonal coarticulation (e.g., Scholz & Chen, 2014; Xu, 1994; Zhang & Liu, 2011), including carryover effect (i.e., the influence of the preceding tones on phrase-final tones) and anticipatory effect (i.e., the influence of the following tones on phrase-initial tones). The finding of this property not only situates Zhangzhou in the typology of tonal coarticulation for the first time, but also questions the default principle of right-dominance of tone sandhi, which conventionally considers that the right-dominant tones appear straightforwardly related to their corresponding citation tones without changing the forms (e.g., Shen & Rose, 2016; Zhang, 2007). Chapters 7 and 8 will conduct detailed acoustic and statistical analyses with respect to how the realisations of Zhangzhou disyllabic tones are affected by their surrounding tones, and to which degree they are affected, and what conditions the variants.

6.1.6. Summary

As discussed, Zhangzhou presents a right-dominant tone sandhi system. As a phonological phenomenon, its sandhi domain is not phonologically determined but syntactically relevant. Although the realisations of the non-rightmost tones undergo changes within a given syntactic phrase, they are not affected by the internal structure of syntactic phrases at either the phonological or the phonetic levels. In addition, the realisations of the tones in both prominent and non-prominent positions tend not to be affected by the categories of their surrounding tones, but appear phonetically-sensitive to pitch contour shapes of neighbouring tones. Therefore, the tonal realisations in multisyllabic contexts in Zhangzhou reflects an interface among linguistic levels, involving phonology, syntax, phonetics, and so on.

6.2. Multidimensional Realisations of Zhangzhou Phrase-Initial Tones

This section provides an auditory description of phrase-initial tones across 64 combinations. It addresses the following questions. (1) What are the phrase-initial tones realised with respect to multidimensional parameters identified in Chapters 4 and 5? (2) How do the segmental and suprasegmental parameters interact to shape the realisations of Zhangzhou phrase-initial tones contrastively? (3) Are the parametric realisations of phrase-initial tones influenced by the tones in phrase-final position at both the phonological and the phonetic levels?

6.2.1. Pitch

This subsection explores how the individual phrase-initial tones interact with following tones in terms of pitch and whether their pitch realisation may be influenced by following tones at both the phonological and the phonetic levels. Table 6-6 shows the results, which generalise the central tendency across 64 disyllabic combinations while ignoring the individual-dependent variations. The leftmost column shows the eight phrase-initial tones being investigated at this point with their corresponding citation pitch values for comparison purposes. Conversely, the top row shows the eight phrase-final tones with their citation pitch values, which serve as the conditioning environments to examine whether they exert any influence on the pitch realisation of their preceding tones. Various examples for each tonal combination appear in Appendix A (Table A2).

Table 6-6. Pitch realisation of Zhangzhou phrase-initial tones

Tone	T1 [35]	T2 [22]	T3 [51]	T4 [41]	T5 [33]	T6 [41]	T7 [221]	T8 [22]
T1 [35]	33	33	33	33	33	33	33	33
T2 [22]	33	33	33	33	33	33	33	33
T3 [51]	24	25	24	24	24	24	25	25
T4 [41]	63	63	63	63	63	63	63	63
T5 [33]	32	32	32	32	32	32	32	32
T6 [41]	65	65	65	65	65	65	65	65
T7 [221]	32	32	32	32	32	32	32	32
T8 [22]	32	32	32	32	32	32	32	32

As Table 6-6 indicates, several notable aspects exist in the pitch realisations of phrase-initial tones across 64 disyllabic combinations.

- All phrase-initial tones have pitch realisations different from their citation forms. For example, tone 1 is realised as a mid-level [33] in the phrase-initial position rather than a mid-rising [35] as occurs in citation. This observation indicates the right-dominant sandhi system in Zhangzhou because all pitches undergo changes at the non-rightmost position.
- All individual phrase-initial tones have similar pitch realisations across phrase-final tones. For example, tone 2 is consistently realised as a mid-level [33] across eight disyllabic constructions, regardless of whether the following tone has a rising, level, or

falling pitch contour. This observation indicates the pitch realisations of phrase-initial tones in Zhangzhou are not affected by their following tones at the phonological level.

- The pitch realisations of phrase-initial tones can be sensitive to following tones at the phonetic level. For example, phrase-initial tone 3 tends to have a higher pitch offset [25] before tones 2, 7, and 8, which perceptually have a lower pitch onset than other tones in the phrase-final position.
- The Chao system of five-scale pitch notation tends to be inadequate for characterising the phrase-initial pitch system in Zhangzhou. Both the highest and lowest pitch levels among phrase-initial tones appear perceptually higher than the ones occurring in phrase-final environment. For example, phrase-initial tone 6 has an extra high and extra short pitch contour, which is significantly higher than the highest pitch onset identified in phrase-final tone 3. Phrase-initial tones 5 and 8 have the lowest pitch offset, which is impressionistically higher than the lowest pitch offset found in phrase-final context. Therefore, this study claims a pitch range of 2 (lowest) to 6 (highest) for the phrase-initial tones.
- Processes of pitch contrast neutralisation occur in the phrase-initial position. Tones 1 and 2 neutralise their pitch contrasts to a mid-level [33] while tones 5, 7 and 8 neutralise the contrasts to a low-falling contour [32] in the phrase-initial context. Therefore, the non-right position creates a certain degree of pitch neutralisation in Zhangzhou.
- Tones that have similar pitch form in citation can be realised differently in the phrase-initial context. For example, in the monosyllabic setting, tones 2 and 8 are both realised as a mid-low-level contour [22], but in the phrase-initial position, they are realised differently. Tone 2 has a mid-level contour [33] while tone 8 has a mid-fall [32]. Thus, different tones can be distinguished in the sandhi position but are neutralised in citation.

Based on the description above, the pitch system of Zhangzhou phrase-initial tones can be summarised as in Table 6-7, including variations in both pitch contour and pitch height.

Table 6-7. Pitch system of Zhangzhou phrase-initial tones

Tone	Phrase-initial pitch	Example
1	mid level [33]	/ʔi1.siŋ1/ ‘doctor’
2	mid level [33]	/ʔi2.tɔŋ5/ ‘to move’
3	mid rise [24]/[25]	/ʔi3.kjɵ5/ ‘sedan chairs’
4	extra high fall [63]	/ʔi4.su4/ ‘meaning’
5	mid fall [32]	/ʔi5.ke4/ ‘anticipate’
6	stopped extra high fall [65]	/ʔik6.tsʰut6/ ‘overflood’
7	stopped mid fall [32]	/ʔik7.kin1/ ‘bath towel’
8	mid fall [32]	/tsi8.tʰɐj1/ ‘furred tongue’

6.2.2. Duration

Three aspects can be noted with respect to the central tendency of length realisation of phrase-initial tones across different following tones in Zhangzhou. The individual speakers reasonably may present a certain degree of variations in length realisation, but it is beyond the topic to conduct a sociophonetic investigation here.

- It is perceptually difficult to identify whether the length realisation of phrase-initial tones can be affected by tones in the phrase-final position. The individual phrase-initial tones tend to have very similar length across the following tones.
- It is perceptually difficult to identify whether the length realisation may vary among the unstopped tones in the phrase-initial position. The individual phrase-initial unstopped tones tend to have very similar length realisation regardless of the shapes of their pitch contour. For example, the length of the phrase-initial rising pitch contour (tone 3) appears very similar to that of the falling (e.g., tone 4) and level (e.g., tone 1) contours.
- The stopped tones 6 and 7 are both significantly shorter than any other tones in the phrase-initial context. They are much shorter than the shortest level found among the eight tones produced in isolation. Nevertheless, it is perceptually difficult to determine whether the two stopped tones may differ from each other in length phrase initially.

Therefore, the phrase-initial tones tend to have only two length levels that are primarily constrained by the syllable coda types. As shown in Table 6-8, a medium-length level [V] occurs for the unstopped tones whose bearing syllables end in sonorants while an extra-short level [V̆̆] only appears for the stopped tones whose associated syllables end in obstruents.

Table 6-8. Length system of Zhangzhou phrase-initial tones

Coda	Tone	Phrase-initial pitch	Length
Sonorant	1	mid level [33]	medium [V]
	2	mid level [33]	
	3	mid rise [25]/[24]	
	4	extra high fall [63]	
	5	mid fall [32]	
	8	mid fall [32]	
Obstruent	6	stopped extra high fall [65]	extra short [V̆̆]
	7	stopped mid-high fall [32]	

The two length levels can thus help enhance the perception and identification of Zhangzhou tones and segments to a certain degree. Additionally, if a phrase-initial syllable is characteristic of a very short length, it is likely to be in the stopped tone of either tone 6 or 7 and to have an obstruent coda at the underlying level. Therefore, duration and syllable coda types, to a certain degree, interact to characterise different tonal contrasts in the phrase-initial context in Zhangzhou.

6.2.3. Vowel quality

Unlike high vowels undergoing diphthongisation in stopped tones in the monosyllabic context, all vowels in Zhangzhou are realised consistently as monophthongs without changing their vowel qualities in the phrase-initial setting. As Table 6-9 indicates, the high vowels /i/ and /u/, as the low vowel /ɔ/, are realised as monophthongs across all phrase-initial tones.

Table 6-9. Vowel quality realisation of the phrase-initial syllables in Zhangzhou

Tone	High vowel /i/	High vowel /u/	Low vowel /ɔ/
1	[**ʔʲi̤33**.ɕiŋ::35] ‘doctor’	[**ʃṳ33**.ʔjẽ::22] ‘win or lose’	[**kɔ33**.pɵ::35] ‘grandpa’s sister’
2	[**ʔʲi̤3**3.tɔŋ::33] ‘move’	[**ʃṳ33**.t̠jɐm52] ‘dictionary’	[**kɔ33**.ʔjɵ::211] ‘medicinal plaster’
3	[**ʔʲi̤25**.kʲjɵ33] ‘sedan chairs’	[**ʃṳ25**.ʔjɔŋ::33] ‘use’	[**kɔ24**.ɗe::33] ‘encourage’
4	[**ʔʲi̤63**.ʃṳ41] ‘meaning’	[**ʃṳ63**.ħɔŋ::35] ‘square’	[**kɔ̰63**.ʔi̤41] ‘deliberate’
5	[ʔʲ**i̤32**.ke41] ‘predict’	[**ʃṳ32**.ɕiɛ̰211] ‘truth’	[**tɔ̰32**.sɔ̰41] ‘degree’
6	[ʔʲ**ĭ̆ᶠk̚65**.tʃwi51] ‘overflood’	[**kʷŭ̆ᶠt̚65**.tʰɐw::311] ‘bone’	[**kɔ̆̆ᶠʔ65**.kɛ::35] ‘country’
7	[ʔʲ**ĭ̤̆k̚**43.kʲin:35] ‘bath towel’	[**hʷṳ̆̆t̚32**.tsɔ̰52] ‘Buddha’	[**tɔ̰̆̆ʔ32**.kʰʲi̤41] ‘poison gas’
8	[**tɕi̤3**2.tʰɐj:35] ‘furred tongue’	[**t̥ʷṳ32**.t̠jɵ41] ‘poke to’	*

While high vowel diphthongisation is considered to be induced by the non-realisation of obstruent codas in the utterance-final position, the monophthong realisation of high vowels is considered a consequence of the preservation of obstruent codas in the non-utterance-final position, as discussed in Section 6.2.5. In other words, the realisation of vowel quality can change with respect to the realisation of obstruent codas across different linguistic contexts.

6.2.4. Voice quality

The realisations of voice quality are also dynamic and diverse in the phrase-initial context, including modal, breathy, creaky, and falsetto. Nevertheless, their occurrence restriction seems more complicated. The occurrence of breathy, creaky, and modal phonation is simultaneously conditioned by vowel quality, pitch and tonal category while the distribution of falsetto voice is motivated simply by tone. Table 6-10 shows the voice quality realisation of the phrase-initial morphemes across tones and vowels.

As Table 6-10 shows, the distributions of breathy, creaky, and modal phonation are largely constrained by vowel quality, pitch and tone. The breathy voice tends to occur on high vowels across phrase-initial tones, except for tone 6. The creaky voice is largely associated with low vowels but only in the tones with falling-pitch contours, except for tone 6. The modal phonation is largely seen in the mid vowels across different phrase-initial tones and in low vowels with non-falling pitch contour. The mid vowel [e] is occasionally perceived with a creaky voice in the falling pitch contour, especially among male speakers.

On the other hand, the falsetto voice occurs across different vowel heights but only if the vowel can occur in the phrase-initial stopped tone 6, which has an extra-high pitch contour and an extra-short duration. Both high and low vowels are perceived with a falsetto voice in the phrase-initial tone 6 while no mid vowels in this context are found in the data.

Table 6-10. Voice quality realisation of phrase-initial syllables in Zhangzhou

Tone	Pitch	High vowel /i/, /u/	Mid vowel /e/, /ɵ/	Low vowel /ɛ/, /ɐ/, /ɔ/
1	mid level [33]	breathy	modal	modal
2	mid level [33]	breathy	modal	modal
3	mid rise [24]/[25]	breathy	modal	modal
4	Extra-high falling [63]	breathy	modal/creaky	creaky
5	mid fall [32]	breathy	modal/creaky	creaky
6	stopped extra-high falling [65]	falsetto	*	falsetto
7	stopped mid fall [32]	breathy	*	creaky
8	mid fall [32]	breathy	modal/creaky	creaky

While the distribution of non-modal phonation is largely constrained by both vowel quality, pitch and tone and partly by tone, the occurrence of different types of phonation helps enhance the perception and identification of specific tones and vowels in the phrase-initial setting. For example, vowels being produced with the falsetto phonation indicate they are in the stopped tone 6. Vowels being produced with a creaky phonation indicate they are largely low vowels and in falling-pitch contours. Alternatively, vowels being produced with the breathy voice are supposed to be high vowels in the tones except tone 6.

6.2.5. Syllable coda

In contrast to the tendency toward non-realisation in the monosyllabic setting, the obstruent codas are largely perceivable but without audible release in the stopped tones 6 and 7 in the phrase-initial context. The velar obstruent codas are observed having two variants of [k] and [ʔ], depending on the backness of their preceding vowels. Table 6-11 shows the realisation of obstruent codas across vowels and across the two phrase-initial stopped tones. No mid vowels are observed in the stopped tonal environment in Zhangzhou. In addition, the preservation of obstruent codas can give rise to several phonetic effects across whole syllables:

- It significantly shortens the duration of phrase-initial syllables. Compared to their length realisation in the monosyllabic and phrase-final settings, the two stopped tones both present an extra-short property that makes them more distinguishable from the unstopped tones, especially from those sharing similar pitch contour and height. For example, both tones 4 and 6 have extra-high falling contours phrase initially, but the stopped tone 6 is perceptually much shorter than the unstopped tone 4.
- It means that diphthongisation is not motivated. For example, the high vowels of phrase-initial syllables are observed not undergoing diphthongisation in stopped tones as they do in the monosyllabic setting.
- It conditions the occurrence of falsetto voice in the stopped tone 6, which is perceived having extra-high pitch contour and extra-short length realisations.

Table 6-11. Obstruent coda realisation of the phrase-initial syllables in Zhangzhou

Coda	Vowel	Tone 6	Tone 7
/p/	/ɐ/	[**tsʰɐ̆̆ᶠp̚6**5.tɕʰju52] ‘interrupt’	[**tsɐ̰̆̆p̚32**.ɟi̤ːː33] ‘twelve’
	/i/	[**ɕĭ̆ᶠp̚65**.tɔːː33] ‘humidity’	[**tɕĭ̤̆p̚32**.tʰḛ52] ‘unit’
	/ɔ/	*	[**ħɔ̰̆̆p̚32**.t̪jɵ41] ‘have caught’
/t/	/ɐ/	[**tʰɐ̆̆ᶠt̚65**.kʲjuːː311] ‘play football’	[**pɐ̰̆̆t̚32**.ɟiɛ̰211] ‘another day’
	/i/	[**ɕĭ̆ᶠt̚65**. ɗɐj311] ‘indoor’	[**ɕĭ̤̆t̚32**.tsḛ41] ‘in reality’
	/u/	[**kʷŭ̆ᶠt̚65**.tʰɐwːː311] ‘bone’	[**hʷṳ̆̆t̚32**.tsɔ̰52] ‘Buddha’
/k/	/ɐ/	[**kɐ̆̆ᶠk̚65**.tɔːː33] ‘point of view’	[**ɗɐ̰̆̆k̚32**.tsɐ̰211] ‘sixty’
	/i/	[**t̪ĭ̆ᶠk̚65**.kɵːː35] ‘bamboo stick’	[**ɗ̪ĭ̤̆k̚32**.ɕiɛ̰̆41] ‘green colour’
	/ɔ/	[**ħɔ̆̆ᶠʔ65**.kʰʲi̤51] ‘good fortune’	[**tɔ̰̆̆ʔ32**.kʰʲi̤41] ‘poison gas’

Further, the preservation of obstruent codas arguably helps enhance the perception and identification of specific tones. If a phrase-initial syllable contains an obstruent coda at the underlying level, its related tone should be either tone 6 of extra high pitch, extra short length and falsetto voice at the surface, or tone 7 of mid-falling pitch, extra short length, breathy high vowel and creaky low vowel. Therefore, the parameters of syllable coda, pitch, length, and voice quality interact to shape the specific characteristics of the stopped tones in Zhangzhou.

6.2.6. Summary

As noted, the realisations of Zhangzhou phrase-initial tones are also multidimensional. They differ not only in terms of pitch height and contour but also in other segmental and suprasegmental parameters, which interact with each other in a complicated but systematic way to characterise the phrase-initial tones contrastively, as shown in Table 6-12. In addition, at the phonological level, the realisations of parameters for phrase-initial tones tend not to be influenced by the categories of the phrase-final tones. The variations for the realisations occur essentially with respect to the phrase-initial tones themselves and the interactions among parameters. Nevertheless, at the phonetic level, some of the realisations, especially pitch, can be sensitive to the phonetic contexts of their following tones.

Table 6-12. Summary of multidimensional realisations of phrase-initial tones in Zhangzhou

Parameter	Tone 1	Tone 2	Tone 3	Tone 4	Tone 5	Tone 6	Tone 7	Tone 8
Pitch	[33]	[33]	[24]/[25]	[63]	[32]	[65]	[32]	[32]
Length	medium	medium	medium	medium	medium	extra short	extra short	medium
High vowel	breathy	breathy	breathy	breathy	breathy	falsetto	breathy	breathy
Mid vowel	modal	modal	modal	modal/creaky	modal/creaky	*	*	modal/creaky
Low vowel	modal	modal	modal	creaky	creaky	falsetto	creaky	creaky
Diphthongisation	*	*	*	*	*	*	*	*
Coda type	sonorant	sonorant	sonorant	sonorant	sonorant	realised obstruent	realised obstruent	sonorant

6.3. Multidimensional Realisations of Zhangzhou Phrase-Final Tones

This section provides an auditory description of the multidimensional realisations of phrase-final tones across 64 disyllabic tonal combinations. It mainly addresses the following questions. (1) What the phrase-final tones are realised with respect to the multidimensional parameters? (2) How do the segmental and suprasegmental parameters interact to characterise phrase-final tones contrastively? (3) Are the realisations of phrase-final tones affected by the tones in phrase-initial position at both the phonological and the phonetic levels?

6.3.1. Pitch

This section has two purposes. One is to explore how the individual phrase-final tones interact with their preceding tones in terms of pitch height and contour. The other is to examine whether the pitch realisation of phrase-final tones may be influenced by their preceding tones.

Table 6-13 shows the findings for pitch realisation of phrase-final tones across 64 combinations by generalising the central tendency among 21 speakers while ignoring the individual speaker-dependent variation. The top row shows the eight phrase-final tones being investigated in this subsection with their corresponding citation pitch values for comparison purposes. The leftmost column shows the eight phrase-initial tones with their sandhi pitch values to examine whether they affect the pitch realisation of following tones. Various examples for each tonal combination appear in Table A2 in Appendix A.

Table 6-13. Pitch realisation of phrase-final tones in Zhangzhou

Tone	T1 [35]	T2 [22]	T3 [51]	T4 [41]	T5 [33]	T6 [41]	T7 [221]	T8 [22]
T1 [33]	34	211	52	41	33	41	211	211
T2 [33]	34	211	52	41	33	41	211	211
T3 [25]/[24]	23	311	52	51	33	51	311	311
T4 [63]	23	211	52	41	33	41	211	211
T5 [32]	24	211	52	41	33	41	211	211
T6 [65]	23	311	52	51	33	51	311	311
T7 [32]	24	211	52	41	33	41	211	211
T8 [32]	24	211	52	41	33	41	211	211

As Table 6-13 indicates, several notable aspects exist with respect to the pitch realisation of phrase-final tones across 64 disyllabic combinations and with respect to comparisons to the citation pitch values.

First, the phrase-final tones do not always have similar pitch realisation across their preceding tones. Most of them have two or more pitch variants because of their phonetic sensitivity to the pitch offset of their preceding tones. For examples, the phrase-final tone 4 generally have a higher pitch onset after tones 3 and 6 which end in a high pitch phrase-initially.

Second, the pitch realisation of phrase-final tones does not always resemble their corresponding citation form, although categorically they are very similar. The divergence can be ascribed to the following two factors. The first factor is phonetic sensitivity to the pitch offset height of preceding tones, causing them to have dynamic and diverse outputs. The second factor is the pitch declination effect in the phrase-final position (Lieberman, 1967; Pierrehumbert, 1987; Maeda, 1976; Cohen et al., 1982; Ladd, 1984; Yuan & Liberman, 2010; Rose, 2014). The phrase-final tones tend to have lower pitch heights than their corresponding citations produced in isolation. For example, although the phrase-final tones 2, 7, and 8 are largely realised as low level, perceptually they tend to have one degree lower than their citation forms ([11] as opposed to [22]). Similarly, the pitch onset of phrase-final tone 1 is generally lower than its citation form ([2] as opposed to [3]).

Third, the pitches of some tones are neutralised in the phrase-final position. For example, tones 2 and 8 neutralise their pitch contours to a low falling while they present somewhat different contours in the phrase-initial context.

As noted, the pitch realisations of most phrase-final tones are phonetically sensitive to the environments of both preceding pitch contour and phrase-final position, rendering their pitch values as not straightforwardly related to their citation forms. Table 6-14 shows a summary of the pitch system of phrase-final tones in Zhangzhou.

Table 6-14. Pitch system of Zhangzhou phrase-final tones

Tone	Phrase-final pitch	Example
1	low/mid rise [23]/[24]/[34]	/ɓjɐ̃2.ʔi3/ ‘famous doctor’
2	mid-low level [211]/[311]	/tswɐn3.ʔi2/ ‘transfer’
3	high fall [52]	/kɐw1.ʔi3/ ‘arm chair’
4	mid-high/high fall [41]/[51]	/ħɵ3.ʔi4/ ‘good intention’
5	mid level [33]	/zjɔŋ2.ʔi5/ ‘easy’
6	stopped mid-high/high fall [41]/[51]	/ħwe2.ʔik6/ ‘recall’
7	stopped mid-low level [211]/[311]	/se3.ʔik7/ ‘take a bath’
8	mid-low level [211]/[311]	/ti1.tsi8/ ‘pig tongue’

6.3.2. Duration

Several points are noted in relation to the length realisation of phrase-final tones. These findings focus on the central tendency among 21 speakers while ignoring the speaker-dependent variations.

- Determining whether the length realisations of phrase-final tones may be affected by the tones preceding them is perceptually difficult. The individual phrase-final tones tend to have very similar realisations in length across different tonal combinations and regardless of whether the preceding pitch contours are rising, level, or falling.
- The phrase-final tones vary considerably in terms of length. Perceptually, three length levels can be generalised from the data, including extra-long [Vːː], long [Vː], and medium [V]. As shown in Table 6-15, the levels are ranked with *1* indicating the extra-long length and *3* the medium.
- The two stopped tones 6 and 7 tend to have the shortest length among the phrase-final tones. Nevertheless, their duration is perceptually considerably longer than the shortest length found among the citation tones and phrase-initial tones. Therefore, the length of a tone can vary across different contexts.

The various length levels may arguably help enhance the perception and identification of specific tones in the phrase-final position. For example, if a phrase-final syllable has the longest duration, its tone is expected to be an unstopped tone having either a rising or a level contour. Therefore, pitch and duration also interact to characterise different tonal contrasts in the phrase-final context in Zhangzhou.

Table 6-15. Length system of Zhangzhou phrase-final tones

Tone	Pitch	Length	Length ranking
1	low/mid rise [24]/[23]/[34]	extra long	1
5	mid level [33]		
8	low fall [211]/[311]		
2	low fall [211]/[311]	long	2
3	high fall [52]		
4	mid-high/high fall [41]/[51]	medium	3
6	stopped mid-high/high fall [41]/[51]		
7	stopped low fall [211]/[311]		

6.3.3. Vowel quality

Perceptually, the realisation of vowel quality of phrase-final syllables tends not to be affected by the phrase-initial tones. Nevertheless, the two high vowels are observed undergoing an active alternation between monophthongs in the unstopped tones and diphthongs in the stopped tones phrase finally. On the other hand, other vowels are realised consistently as monophthongs across different tonal combinations. Table 6-16 shows the realisations of Zhangzhou high vowels /i/ and /u/ and low vowel /ɔ/ across phrase-final tones. The symbol * indicates no token was found in the related context.

Table 6-16. Vowel quality realisations across Zhangzhou phrase-final tones

Tone	High vowel /i/	High vowel /u/	Low vowel /ɔ/
1	[mjẽ̆33.**ʔi̤ːː34**] ‘famous doctor’	[ɗɵ25.**ʃṳːː24**] ‘teacher’	[ɦjɔ̃33.**kɔːː24**] ‘mushroom’
2	[tʃwɐn25.**ʔi̤ːː311**] ‘transfer’	[tɔŋ32.**ʃṳːː211**] ‘verb’	[ɓi̤33.**kɔ̰ːː211**] ‘rice pasta’
3	[kɐw33.**ʔi̤ː52**] ‘armed chair’	[d̪̤i̤t̚32.**ʃṳː52**] ‘history’	[pʰɐ63.**kɔ̰ː52**] ‘play the drum’
4	[ɦɵ35.**ʔi̤ː41**] ‘good intention’	[ɦɔŋ63.**ʃṳː41**] ‘unbridled’	[tɕjɐw63.**kɔ̰ː41**] ‘look after’
5	[j̰jɔŋ33.**ʔi̤ːː33**] ‘easy’	[ɦɵ25.**ʃṳːː33**] ‘good deed’	[kɛ̰̆ᶠk̚65.**tɔːː33**] ‘point of view’
6	[hʷwe33.**ʔiɛ̰41**] ‘recall’	[pɐ̰j33.**kʷuɛ̰41**] ‘ribs’	[t̪̰jɔŋ33.**kɔ̰41**] ‘China’
7	[se25.**ʔiɛ̰̆311**] ‘take a bath’	[pɐ̰j33.**hʷuɛ̰211**] ‘worship Buddha’	[t̪̰jɔŋ63.**kɔ̰211**] ‘poisoning’
8	[t̪̤i̤33.**tɕi̤ːː211**] ‘pig tongue’	*	*

As Table 6-16 shows, the high front vowel [i] is diphthongised to [iɛ̰] while the high back vowel [u] shifts to a diphthong [uɛ̰] in the phrase-final stopped tones 6 and 7. In contrast, the low vowel /ɔ/ is realised consistently as [ɔ] without changing the vowel quality. The high vowel diphthongisation is suggested to be motivated by the non-realisation of obstruent codas at the phrase-final position (see Section 6.3.5).

Vowel diphthongisation does help enhance the perception and identification of specific tone, vowel, and syllable coda type. For example, if the vowel of a phrase-final syllable undergoes diphthongisation, it indicates its related syllable has a stopped tone and an obstruent coda underlyingly. Alternatively, if a high vowel is realised as a monophthong in the phrase-final position, it indicates its related syllable has an unstopped tone and ends in a sonorant coda.

6.3.4. Voice quality

As in the monosyllabic setting, three different phonation types—modal, breathy, and creaky—can be identified in phrase-final syllables, and the realisation of non-modal phonation is also conditioned by both vowel quality, pitch and tone. As shown in Table 6-17, the breathy voice is more likely to occur on the high vowels. The creaky voice largely occurs on the low vowels but occasionally occurs on the mid vowel [e], especially among male speakers. The modal phonation is largely found on the mid vowels and the low vowels with non-falling pitch contour.

Table 6-17. Voice quality realisations across Zhangzhou phrase-final vowels and tones

Tone	Pitch	High vowel /i/, /u/	Mid vowel /e/, /ɵ/	Low vowel /ɛ/, /ɐ/, /ɔ/
1	low/mid rise [24]/[23]/[34]	breathy	modal	modal
2	low fall [211]/[311]	breathy	modal/creaky	creaky
3	high fall [52]	breathy	modal/creaky	creaky
4	mid-high/high fall [41]/[51]	breathy	modal/creaky	creaky
5	mid level [33]	breathy	modal	modal
6	stopped mid-high/high fall [41]/[51]	creaky	*	creaky
7	stopped low fall [211]/[311]	creaky	*	creaky
8	low fall [211]/[311]	breathy	modal/creaky	creaky

Further, the realisation of voice quality changes in certain tonal circumstances. For example, the voice quality of low vowels is changed from creaky to modal in rising and mid-level contexts. The voice quality of high vowels shifts to the creaky phonation from breathy in the stopped tonal contexts. As for the mid vowels, they are perceived largely with the modal phonation regardless of pitch contour, but occasionally the mid vowel [e] is perceived with a creaky voice in the falling and low pitch environments. Examples of voice quality realisations in the phrase-final position appeared in Table 6-16.

While vowel quality, pitch and tone simultaneously condition the realisation of voice quality, voice quality arguably helps enhance the perception and identification of specific tones and vowels in Zhangzhou. For example, if a vowel is produced with the breathy voice, it is expected to be a high vowel in a non-stopped tone. Alternatively, if a vowel is produced with a modal phonation, it is considered to be a mid vowel or low vowel in a non-falling pitched tone.

6.3.5. Syllable coda

In contrast to the preservation of obstruent codas in the phrase-initial syllable, but similar to the non-realisation in citation, the obstruent codas—/p/, /t/, and /k/—tend not to be realised in the phrase-final context, although the bilabial coda [p] is occasionally perceivable (Table 6-18).

Similar to the situation in the monosyllabic setting, the non-realisation of obstruent codas in the phrase-final position can also give rise to several phonetic effects on whole syllables, including syllable lengthening, high vowel diphthongisation, vowel laryngealisation, and pitch contour depression, as those discussed in Section 4.2.4.

Table 6-18. Obstruent coda realisation with respect to Zhangzhou phrase-final tones

Coda	Vowel	Tone 6	Tone 7
/p/	/ɐ/	[hʷwe33.**tɐ̰41**] ‘answer’	[ɠɔ̰32.**tsɐ̰**ːː**211**] ‘fifty’
	/i/	[t̪jɵ32.**kʲiɛ̰41**] ‘anxious’	[tʃʰŭ̆ᶠt̚65.**ʝiɛ̰**ːː**311**] ‘go out and come in’
	/ɔ/	*	*
/t/	/ɐ/	[kʷwɐn33.**tsʰɐ̰41**] ‘observe’	[tʃʰŭ̆ᶠt̚65.**ɗɐ̰**ːː**311**] ‘exert oneself’
	/i/	[kʰɵ64.**ɕiɛ̰41**] ‘classroom’	[ɗɐw25.**ɕiɛ̰**ːː**311**] ‘honest’
	/u/	[pɐj33.**kʷuɤ̰4**1] ‘ribs’	[pɐj63.**hʷuɤ̰**ːː**211**] ‘serve Buddha’
/k/	/ɐ/	[ɣʷṳ33.**kɐ̰41**] ‘cattle horn’	[tsɐ̰p̚32.**ɗɐ̰**ːː**211**] ‘sixteen’
	/i/	[ɓɔ̰ʔ32.**t̪iɛ̰41**] ‘purpose’	[se25.**ʔiɛ̰**ːː**211**] ‘take a bath’
	/ɔ/	[t̪jɔŋ33.**kɔ̰41**] ‘China’	[t̪jɔŋ63.**kɔ̰**ːː**211**] ‘poisoning’

6.3.6. Summary

The above descriptions have shown the realisations of phrase-final tones are multidimensional, involving diverse phonetic parameters that interact in a complicated but systematic way, as shown in Table 6-19. In addition, the realisations of various parameters for phrase-final tones tend not to be categorically influenced by the phrase-initial tones. Nevertheless, at the phonetic level, the realisations can be considerably affected by the tones preceding them. For example, the pitches of phrase-final tones are easily sensitive to the pitch offsets of the tones in the phrase-initial position and present dynamic and diverse variants.

6.4. Previous Studies

Essentially, no empirical study has explored what individual tones are realised and how they interact in multisyllables. No descriptive study has specified the sandhi domain of Zhangzhou. No theoretical explanation has been developed to account for how tones in multisyllables, especially the sandhi tones, are related to their citation forms. What has been consistently documented in previous studies of Zhangzhou tone sandhi only involves a brief piece of information that one tone is changed into another tone for the non-final characters (Dong, 1959; Lin, 1992; Ma, 1994; ZCCEC, 1999; Gao, 1999; Yang, 2008), which can be expressed in terms of Middle Chinese tonal categories, as summarised in Table 6-20. For example, the Yinping tone is reported changing to the Yangqu tone while the Yangqu tone alternating with the Yinqu tone in non-final characters. Yang (2008) also broadly mentioned that the tone sandhi of Zhangzhou also forms a tone circle, as found in other Southern Min varieties, starting with Yinping and ending with Yinping tone, but she did not interpret why tonal altenations form a circle, and how sandhi tones are related to the citation tones.

Table 6-19. Summary of multidimensional realisations of Zhangzhou phrase-final tones

Parameter	Tone 1	Tone 2	Tone 3	Tone 4	Tone 5	Tone 6	Tone 7	Tone8
Pitch	[23]/[24]/[34]	[211]/[311]	[52]	[41]/[51]	[33]	[41]/[51]	[211]/[311]	[211]/[311]
Length	extra long	long	long	medium	extra long	medium	medium	long
High vowel	breathy	breathy	breathy	breathy	breathy	creaky	creaky	breathy
Mid vowel	modal	modal/creaky	modal/creaky	modal/creaky	modal	*	*	modal/creaky
Low vowel	modal	creaky	creaky	creaky	modal	creaky	creaky	creaky
Diphthongisation	-	-	-	-	-	+	+	-
Coda type	sonorant	sonorant	sonorant	sonorant	sonorant	not realised obstruent	not realised obstruent	sonorant

139

Table 6-20. Summary of previous description of Zhangzhou tone sandhi

Tone	Citation	Sandhi
1	Yinping	Yangqu
2	Yangping	Yangqu
4	Shangsheng	Yinping
4	Yinqu	Shangsheng
5	Yangqu	Yinqu
6	Yinru [ʔ]	Yinqu
	Yinru [p], [t], [k]	high level [55]/[44]/[5]/[53]
7	Yangru	Yinqu

Therefore, although previous studies did briefly mention the tone sandhi phenomenon in Zhangzhou, their characterisation of sandhi environments in terms of Chinese characters apparently is not a statement of linguistic significance to encode the prosodic or the grammatical relation between morphemes in polysyllabic constructions in which tone sandhi occurs. Their interpretation of tonal alternation in terms of Middle Chinese tonal categories does not reveal the mechanism behind the tone sandhi. In this regard, Zhangzhou can be seen as a hitherto under-described variety with respect to tonal study beyond the monosyllables. This chapter fills a research gap with a systematic auditory description.

6.5. Conclusion

This chapter has shown that Zhangzhou presents a right-dominant tone sandhi system, involving a complicated interface among phonology, syntax, phonetics, and other linguistic elements. Its tone sandhi domain appears aligned with syntactic phrases, but the process of tone sandhi is only sensitive to the position instead of being sensitive to the internal structures of given syntactic phrases. The parametric realisations of a given tone in both the dominant and the non-dominant positions tend not to be affected by surrounding tones at the phonological level but can be phonetically sensitive to neighbouring tones, especially the pitch realisation of phrase-final tones.

In addition, this chapter had included a systematic description of the auditory properties of Zhangzhou disyllabic tones, showing the realisations of the tones at both phrase-initial and phrase-final contexts are multidimensional, involving a variety of phonetic parameters. The auditory description provides a framework for subsequent acoustic quantification and statistical testing of phrase-initial tones in Chapter 7 and of phrase-final tones in Chapter 8. It also provides a solid foundation for the theoretical discussion of mapping between citation tones and disyllabic tones in Chapter 9.

Chapter 7: Acoustic Properties of Zhangzhou Phrase-Initial Tones

This chapter explores the acoustic properties of Zhangzhou phrase-initial tones. It largely examines what phrase-initial tones are realised in acoustics with respect to the multidimensional parameters across 64 tonal combinations. It applies the techniques of pairwise *t* testing and the hierarchical clustering algorithm to determine whether the F0 and duration realisations of individual phrase-initial tones may be affected by their following tones. It employs the SplitsTree software to assess the relatedness between phrase-initial tones from the phonological perspective.

7.1. *Z*-Score Normalised F0 Model

This section largely addresses three issues. (1) What are the F0 realisations of individual phrase-initial tones? (2) Are the F0 realisations of phrase-initial tones affected by the tones in the phrase-final position? If so, to what extent are they affected, and what conditions the variations? (3) How many normalised F0 levels are contrastive among the phrase-initial tones?

The F0 contours presented in this section were obtained by *z*-score normalising the raw F0 values of the related contours from 21 speakers' utterances. The techniques of pairwise *t* testing and the hierarchical clustering algorithm were applied to examine the statistical relation among the normalised F0 values of each phrase-initial tone across its eight following tones.

Logically, each phrase-initial tone would have 28 (=8*7/2) paired normalised F0 differences to be tested at either the 10% or the 100% sampling point, assuming that the normalised values are all independent and identically normally distributed. The Bonferroni corrected alpha of 0.00186 (=0.05/28) was consistently performed to control for the Type I Error for the values at the sampling point under consideration. The threshold of 1 was consistently employed to cluster the values and examine how many normalised F0 levels are contrastive phrase initially from the statistical point of view.

The normalised F0 values and the values of the pairwise *t*-testing result for individual phrase-initial tones are tabulated in Appendix B (Tables B6-B9). The methodologies of acoustic and statistical processing are described in Chapter 2.

7.1.1. Phrase-initial tone 1

The phrase-initial tone 1 was indicated in Chapter 6 having very similar pitch realisation of mid-level [33] across eight tonal combinations. The acoustically quantified results basically support this impressionistic description. As Figure 7-1 shows, the F0 contours largely level around the midpoint. Nevertheless, slight variation occurs in terms of the contour distribution. For example, the contours before tones 2, 7, and 8 are higher than they are before other tones.

The statistical testing and clustering results based on pairwise *t* tests revealed significant differences among the normalised F0 values, but only at the 100% sampling point, as shown in Figure 7-2. The left panel visualises the normalised F0 values of phrase-initial tone 1 across eight phrase-final tones at the 10% sampling point, while the right panel represents the values at the 100% sampling point. As indicated, the normalised F0 values before tones 2, 7, and 8 are clustered together and significantly higher than they are before other tones at the 100% sampling point. On the contrary, no significant differences exist among the values at the 10% sampling point.

Therefore, although the F0 realisations of the phrase-initial tone 1 are very similar as a function of the phrase-final tones, the offset values are phonetically sensitive to the tones in the phrase-final position and present statistically significantly different variants. The conditioning factor for the variation can be generalised as [+low onset], which is shared in common by the phrase-final tones 2, 7, and 8. If the phrase-final tone has a lower onset, the phrase-initial tone 1 is supposed to have a statistically significantly higher offset.

7.1.2. Phrase-initial tone 2

In Chapter 6, the phrase-initial tone 2 was perceived as having similar realisation of a mid-level [33] in spite of the categories of phrase-final tones. The auditory impressions are largely supported by the acoustically quantified results, as presented in Figure 7-3, with the contours largely level around the midpoint. Similar to the realisations of phrase-initial tone 1, the normalised F0 contours appear numerically higher before tones 2, 7, and 8 than they are before other tones.

Statistically, the pairwise *t*-test comparisons revealed significant differences among the normalised F0 values of phrase-initial tone 2 but only at the offset position. As indicated in Figure 7-4, the eight phrase-final tones in the right panel have been clustered into two levels at the threshold of 1, indicating that the normalised F0 values of the phrase-initial tone 2 are statistically significantly different as a function of phrase-final tones at the 100% sampling point. The results also indicate that the offset values before tones 2, 7, and 8 are significantly higher than the values before other tones. Conversely, the eight phrase-final tones are clustered into one single class at the 10% sampling point (left panel), indicating no statistically significant differences at the onset realisations of this phrase-initial tone 2.

Therefore, the F0 realisations of phrase-initial tone 2 are not quite the same with respect to phrase-final tones from both acoustic and statistical points of view. The variations of its F0 offset are also conditioned by the feature [+low onset] shared by the phrase-final tones 2, 7 and 8. If the tone at the phrase-final position has a low onset value, the offset value of the phrase-initial tone 2 is supposed to be statistically significantly higher.

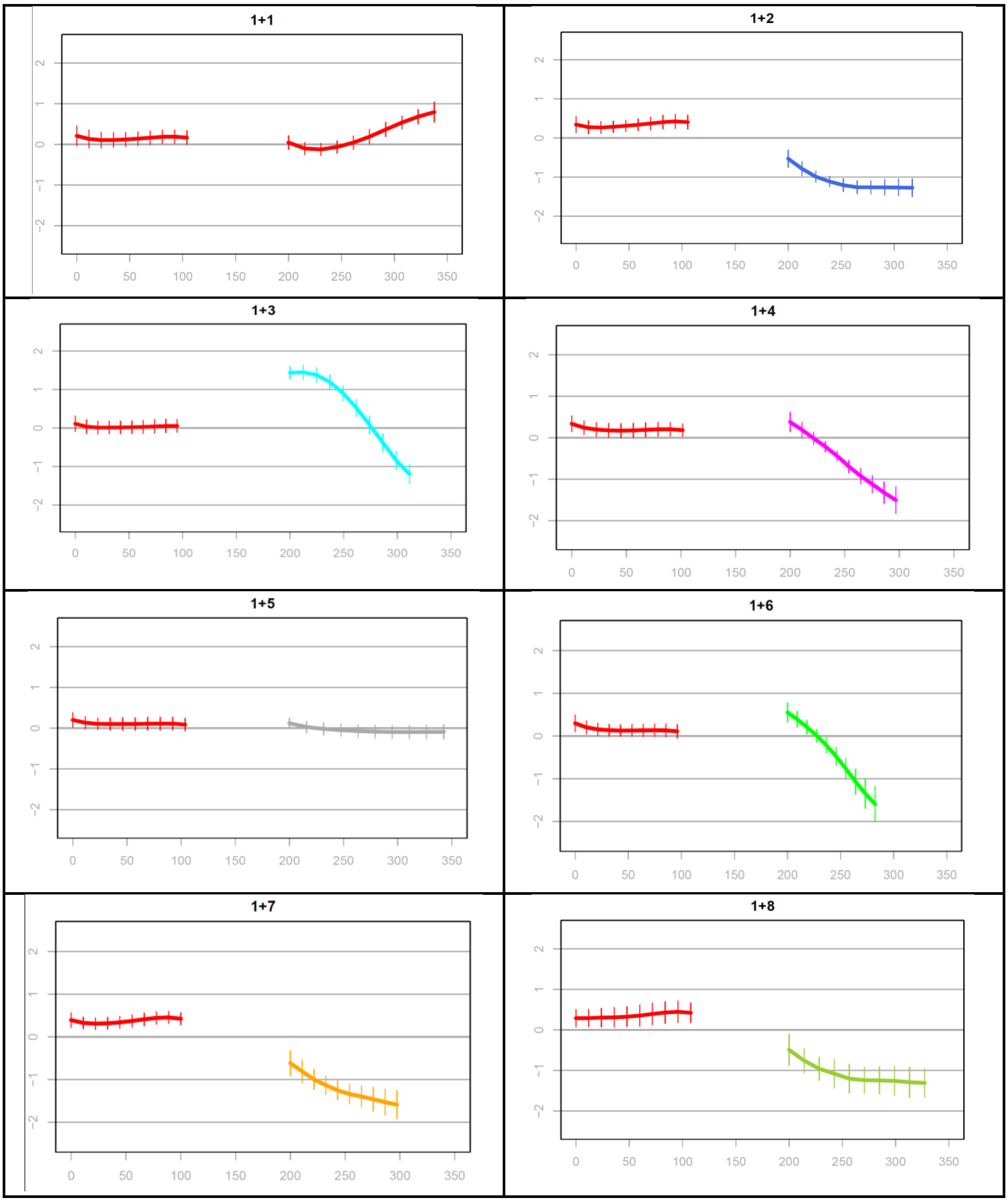


Figure 7-1. *Z*-score normalised F0 shapes of phrase-initial tone 1 across phrase-final tones from 21 speakers.

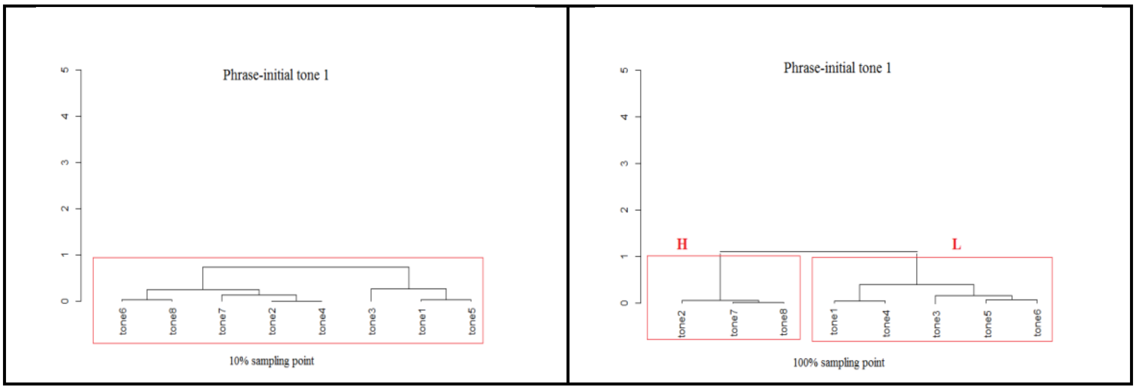


Figure 7-2. Clustering of normalised F0 levels of phrase-initial tone 1 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

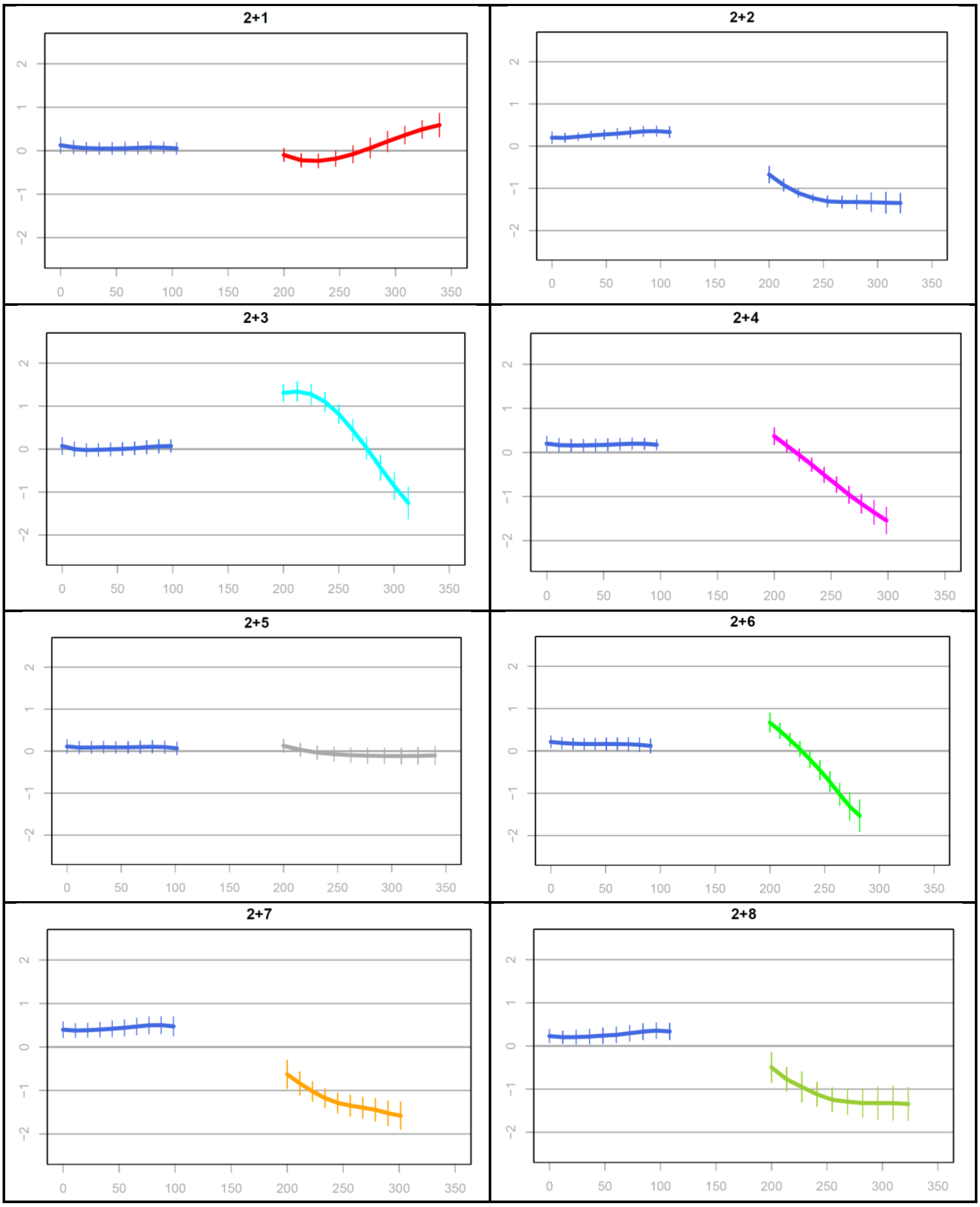


Figure 7-3. *Z*-score normalised F0 shapes of phrase-initial tone 2 across phrase-final tones from 21 speakers.

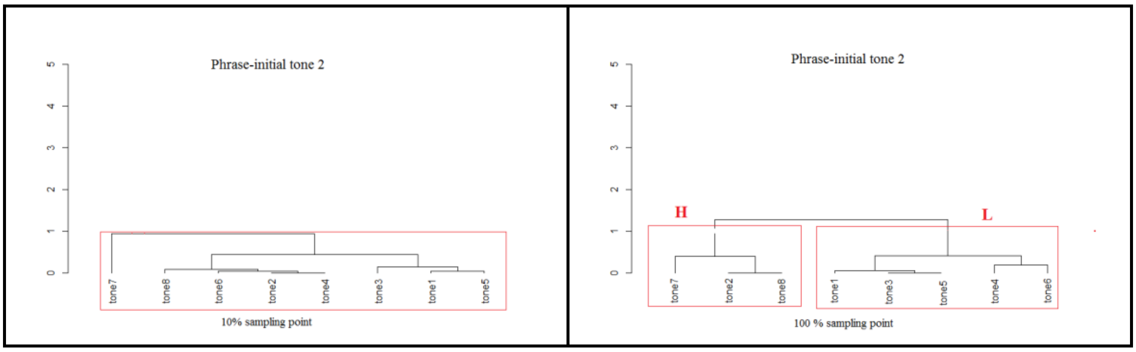


Figure 7-4. Clustering of normalised F0 levels of phrase-initial tone 2 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

7.1.3. Phrase-initial tone 3

The pitch of phrase-initial tone 3 was indicated in Chapter 6 as being a rising contour, regardless of the categories of its following tones, but variations ([24] and [25]) exist in the pitch offset height. This auditory description is justified both acoustically and statistically. In Figure 7-5, the normalised F0 contours of this tone present a rising trend in all combinations but with apparent variations. For example, this tone has a considerably higher F0 offset before tone 2 than it has before tone 1.

Statistically, the results of pairwise *t*-test comparison revealed significant differences among the normalised F0 offset values of the phrase-initial tone 3 but no significant differences among the onset values across different combinations. As Figure 7-6 indicates, the eight phrase-final tones in the right panel have been clustered into two levels at the threshold of 1 at the 100% sampling point, indicating that the normalised F0 offset values of the phrase-initial tone 3 are statistically significantly different as a function of phrase-final tones. The results also indicate that the offset values before tones 2, 7, and 8 are significantly higher than the values before other tones. Conversely, the eight phrase-final tones are clustered into one single class at the 10% sampling point (left panel), indicating no statistically significant differences between the onset realisations of this phrase-initial tone 3 across its following tones.

Therefore, the realisations of this tone are also not quite the same across different tonal combinations. The offset realisations are phonetically sensitive to the onset values of the tones at the phrase-final position. The low F0 onset feature of tones 2, 7, and 8 conditions the phrase-initial tone 3 to have a higher offset height.

7.1.4. Phrase-initial tone 4

Phrase-initial tone 4 has been described in Chapter 6 as having an extra-high falling contour [63] across all phrase-final tones. The pitch onset is perceived considerably higher than the highest pitch onset among the phrase-final tones. This auditory observation is justified acoustically. As shown in Figure 7-7, all contours of this tone present a falling tendency from 2 s.d. to somewhere slightly over the midpoint. The F0 onset is significantly higher than the onset of phrase-final tone 3, which is the highest among the phrase-final tones.

Statistically, the results of pairwise *t*-test comparison by effect size revealed no significant differences among the normalised F0 values of phrase-initial tone 4 at both the 10% and 100% sampling points as a function of the phrase-final tones, as shown in Figure 7-8. The eight phrase-final tones are clustered together as a single group at either the 10% (left) or the 100% (right) sampling point at the threshold of 1, indicating that the F0 realisations of both onset and offset of phrase-initial tone 4 are statistically undifferentiated across their following tones.

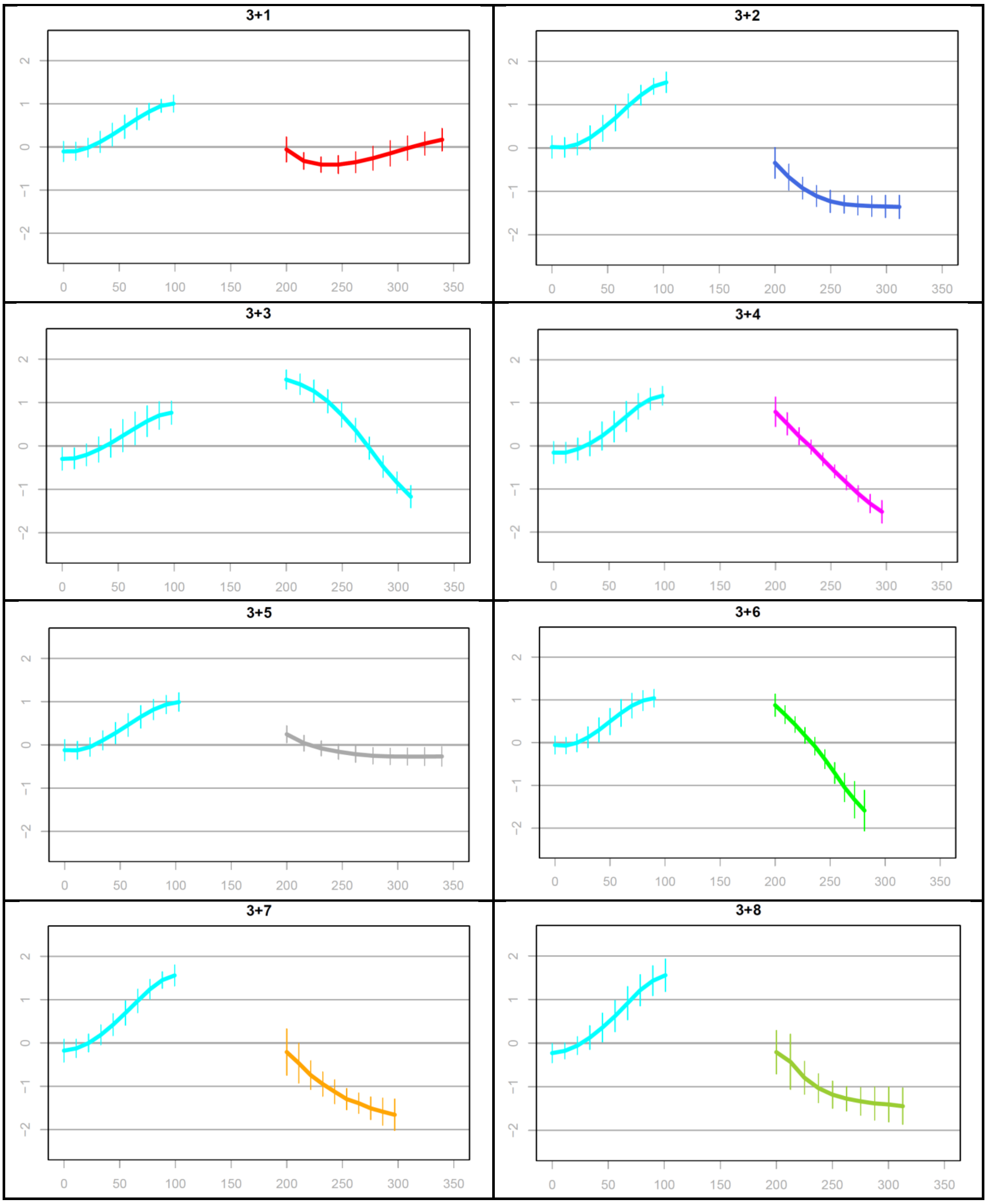


Figure 7-5. *Z*-score normalised F0 shapes of phrase-initial tone 3 across phrase-final tones from 21 speakers.

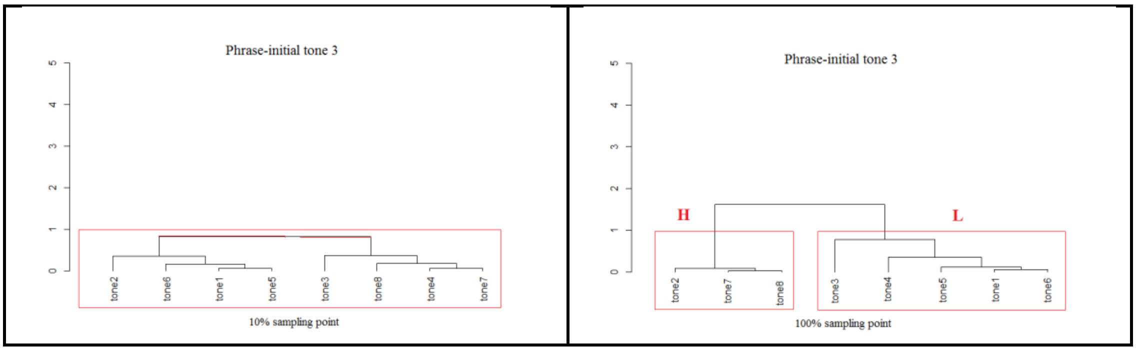


Figure 7-6. Clustering of normalised F0 levels of phrase-initial tone 3 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

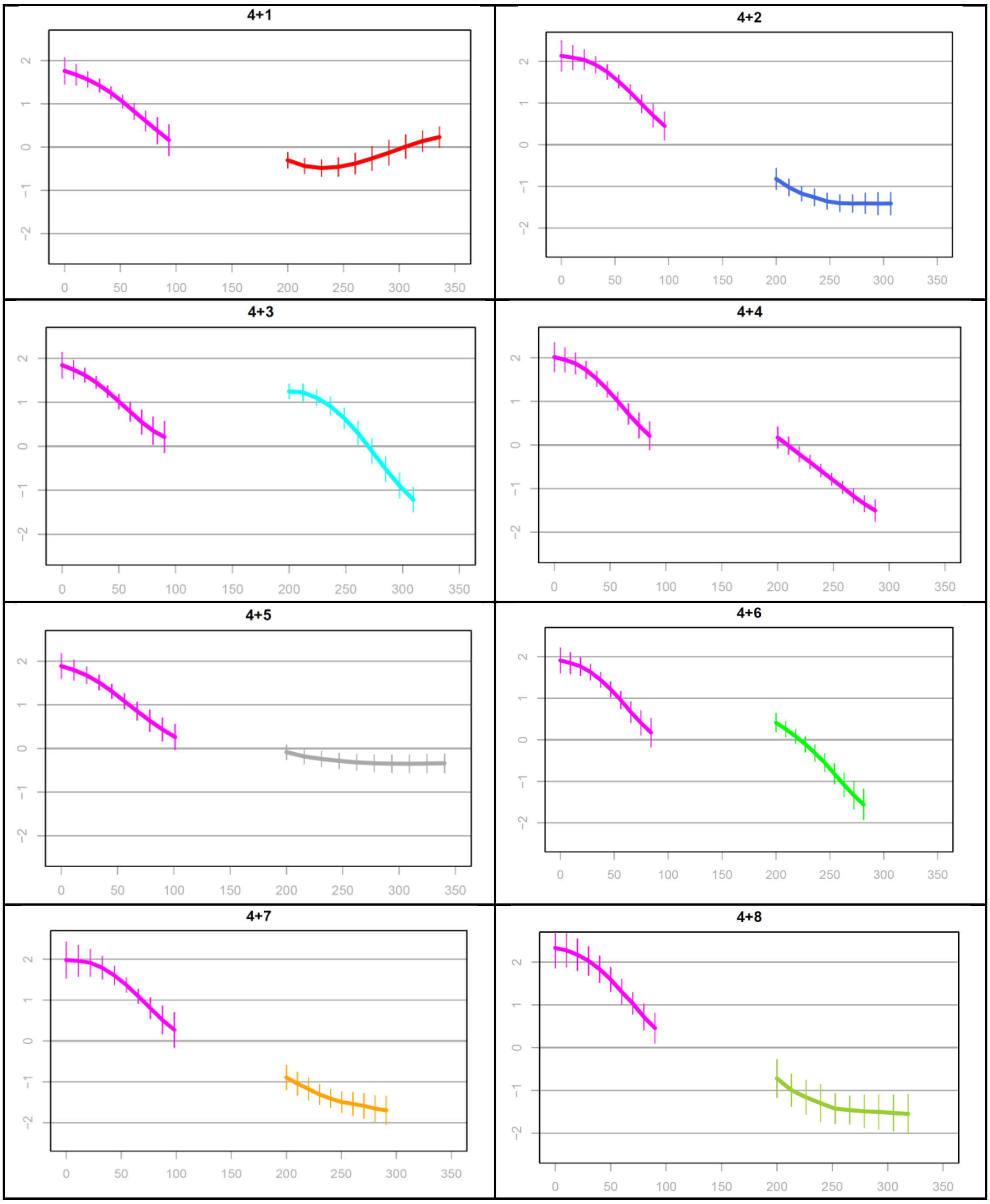


Figure 7-7. ***Z*-score normalised F0 shapes of phrase-initial tone 4 across phrase-final tones from 21 speakers.**

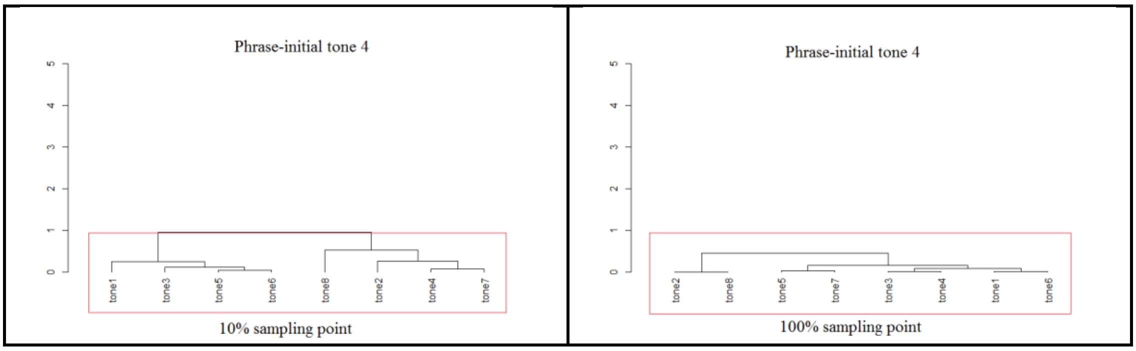


Figure 7-8. Clustering of normalised F0 levels of phrase-initial tone 4 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

7.1.5. Phrase-initial tone 5

The pitch of phrase-initial tone 5 was indicated as a mid-falling contour [32] across eight disyllabic combinations in Chapter 6. This auditory impression is somewhat justified by the acoustically quantified results. As shown in Figure 7-9, all contours of this tone present a falling tendency from around the midpoint to 1 s.d. below. Nevertheless, variations can also be observed at both contour onset and offset. For example, the contour onset is slightly higher before tone 2 than it is before tone 1 while the offset is slightly higher before tone 7 than it is before tone 5.

Statistically, the normalised F0 values of the phrase-initial tone 5 are not significantly different as a function of the phrase-final tones at both onset and offset positions on the basis of pairwise t test comparison by effect size. As shown in Figure 7-10, the eight phrase-final tones at either the 10% sampling point (left) or the 100% sampling point (right) are clustered into one single class at the threshold of 1, indicating that the F0 realisations of the phrase-initial tone 5 are statistically undifferentiated across their following tones.

7.1.6. Phrase-initial tone 6

The phrase-initial tone 6 was indicated as having an extra-short and high-falling contour [65] across different tonal combinations in Chapter 6. Similar to phrase-initial tone 4, its pitch onset is also perceptually higher than the highest pitch onset of the tones in phrase-final position. This auditory impression is justified acoustically. As shown in Figure 7-11, all contours show a falling tendency between 2 s.d. and 1 s.d above the midpoint, and the onset and offset of phrase-initial tone 6 are generally higher than the highest onset of phrase-final tone 3. Nevertheless, small variations can be seen in the F0 distribution. For example, the contour has a higher offset before tone 8 than before tone 6.

Statistically, the results of pairwise *t*-test comparison by effect size revealed no significant differences among the normalised F0 values of the phrase-initial tone 6 at both the 10% and 100% sampling points as a function of the phrase-final tones, as shown in Figure 7-12. The eight phrase-final tones at either the 10% sampling point (left) or the 100% sampling point (right) are clustered into one single class at the threshold of 1, indicating that the values are statistically undifferentiated. In other words, the F0 realisations of phrase-initial tone 6 are not categorically affected by its following tones.

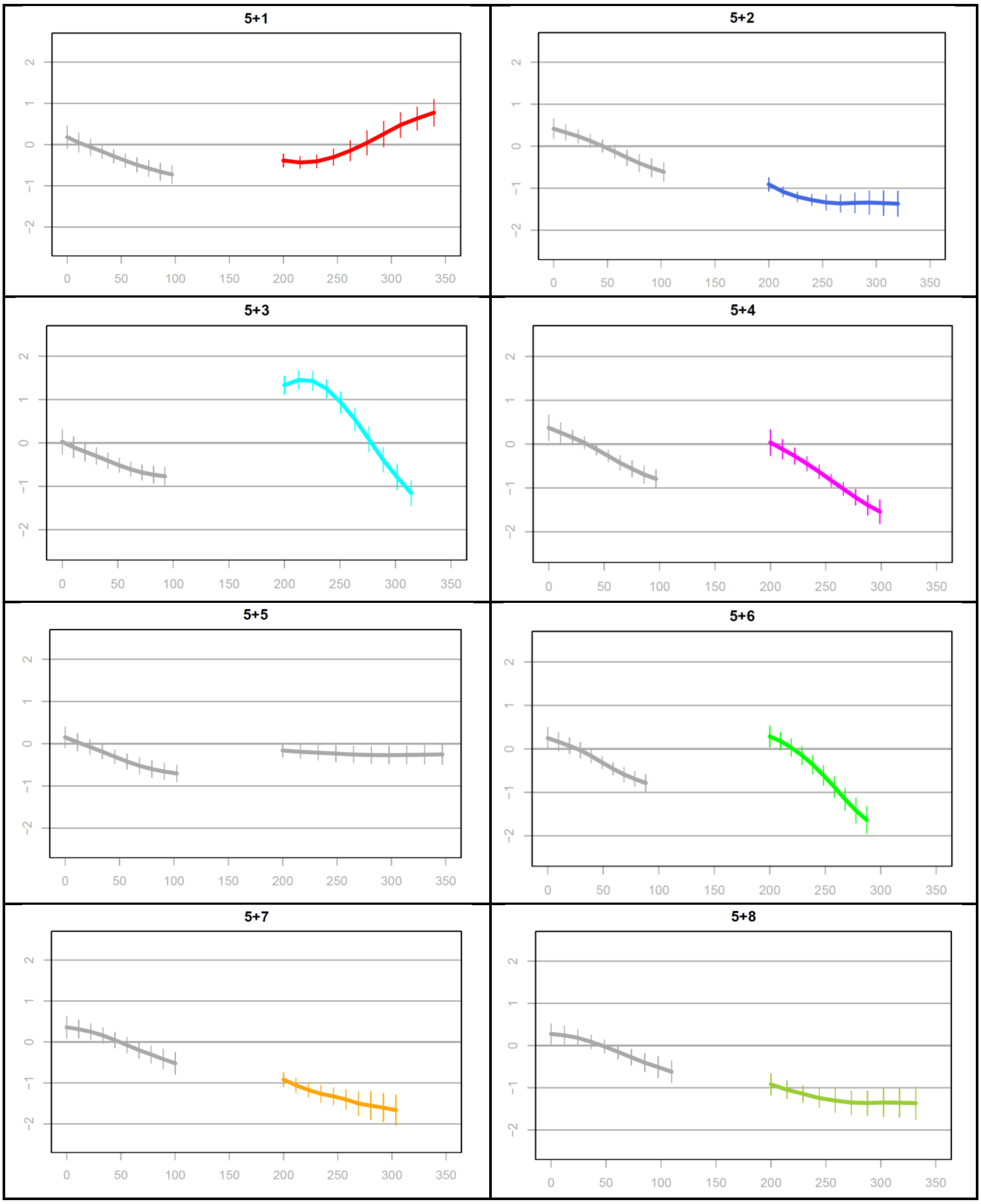


Figure 7-9. *Z*-score normalised F0 shapes of phrase-initial tone 5 across phrase-final tones from 21 speakers.

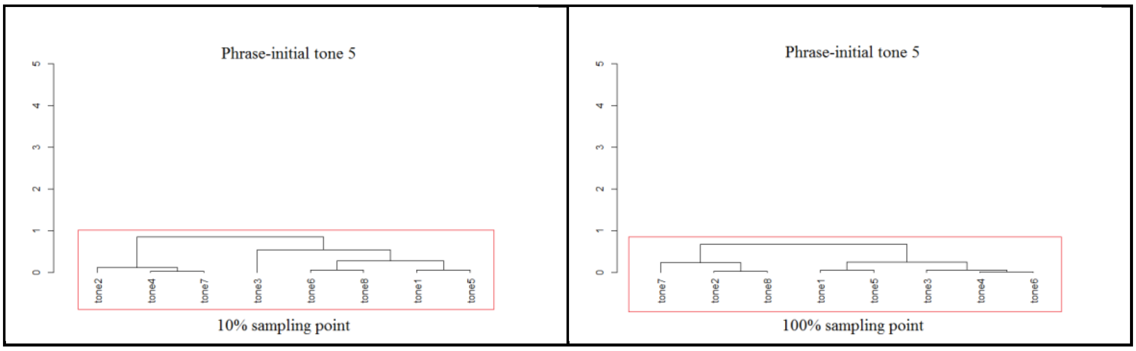


Figure 7-10. Clustering of normalised F0 levels of phrase-initial tone 5 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

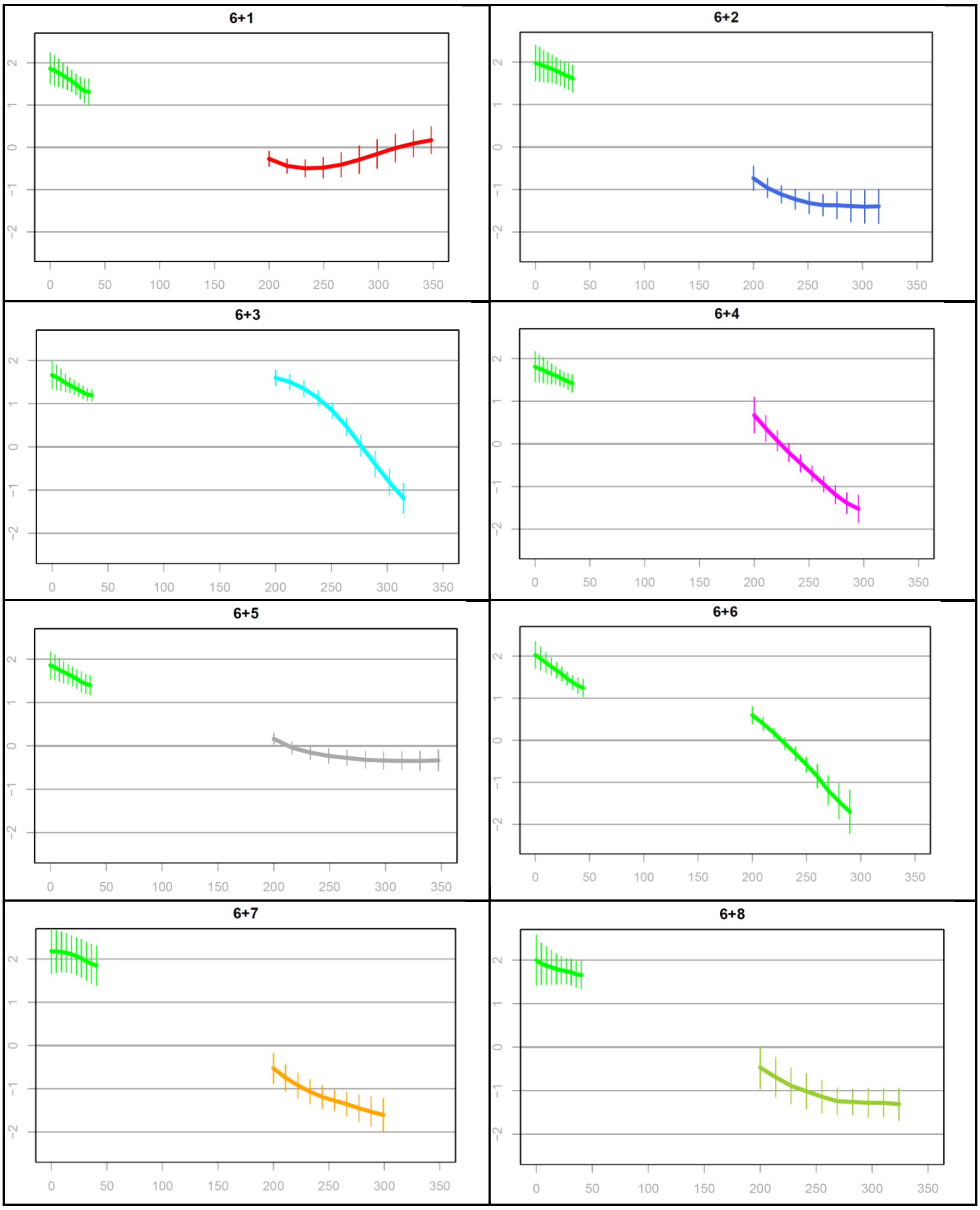


Figure 7-11. *Z*-score normalised F0 shapes of phrase-initial tone 6 across phrase-final tones from 21 speakers.

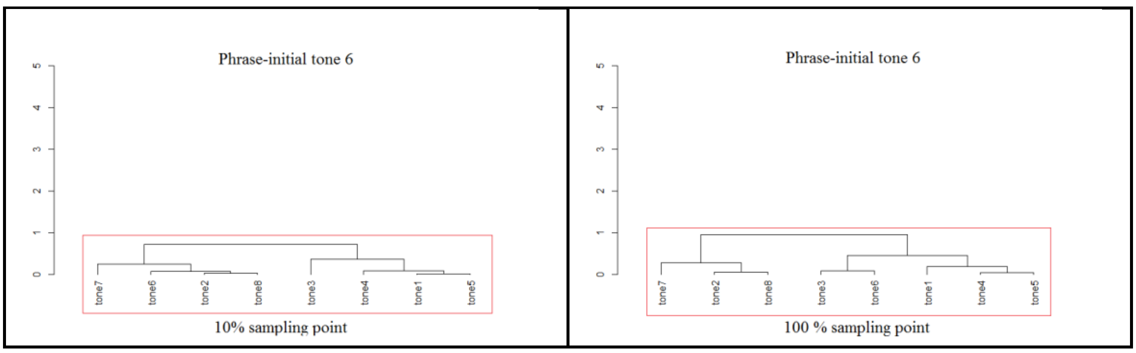


Figure 7-12. Clustering of normalised F0 levels of phrase-initial tone 6 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

7.1.7. Phrase-initial tone 7

The pitch of phrase-initial tone 7 has been perceived as an extra-short mid-falling contour [32] across most disyllabic combinations. The acoustically quantified results basically justify the auditory impression. As shown in Figure 7-13, the contours all present a falling tendency but with a very trivial slope around 0.5 s.d. Small variation can also be seen in the contour offset. For example, the contour before tone 2 has a slightly higher offset than the contour before tone 1.

Statistically, the results of pairwise *t*-test comparison revealed no significant differences among the normalised F0 values of phrase-initial tone 7 at both the 10% and 100% sampling points as a function of the phrase-final tones, as shown in Figure 7-14. The eight phrase-final tones at either the 10% sampling point (left) or the 100% sampling point (right) are clustered into one class at the threshold of 1, indicating that the values are statistically undifferentiated. In other words, the F0 realisations of phrase-initial tone 7 are not categorically affected by the tones at the phrase-final position.

7.1.8. Phrase-initial tone 8

The pitch of phrase-initial tone 8 was described as a mid-falling contour [32] as a function of phrase-final tones in Chapter 6 because it is perceptually difficult to identify whether the phrase-final tones may exert influence on its F0 realisations. The acoustically quantified results effectively justify this auditory observation. As shown in Figure 7-15, all contours present a falling tendency from 0.5 s.d. above the midpoint to 1 s.d. below the midpoint; however, small variation can also be seen in the normalised values. For example, the contour before tone 3 tends to have a slightly lower offset.

Statistically, no significant differences were revealed by the pairwise *t*-test comparison by effect size among the normalised F0 values of the phrase-initial tone 8 as a function of its following tone at both 10% and 100% sampling points, as shown in Figure 7-16. The eight phrase-final tones at either the 10% sampling point (left) or the 100% sampling point (right) are clustered into one single class at the threshold of 1, indicating that the normalised F0 values of phrase-initial tone 8 are statistically undifferentiated across different tonal combinations. In other words, the F0 realisations of phrase-initial tone 8 are not categorically affected by its following tones.

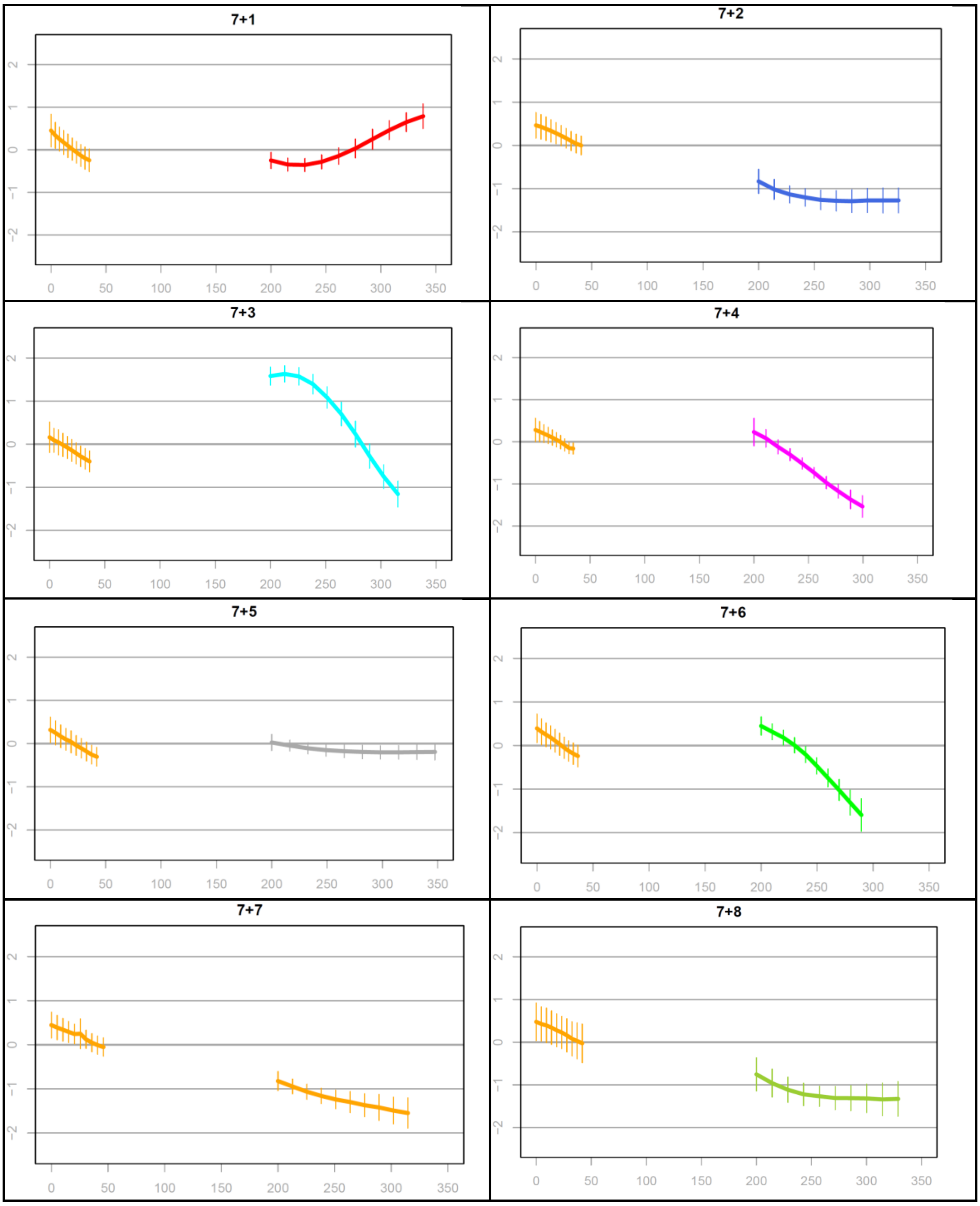


Figure 7-13. *Z*-score normalised F0 shapes of phrase-initial tone 7 across phrase-final tones from 21 speakers.

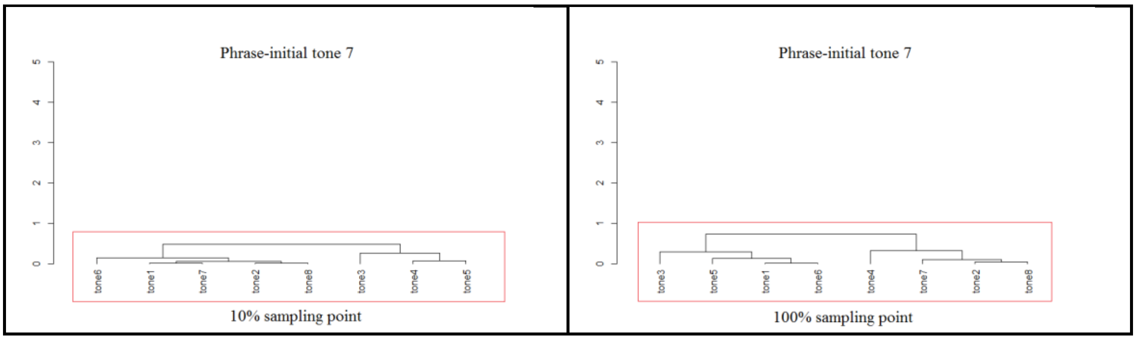


Figure 7-14. Clustering of normalised F0 levels of phrase-initial tone 7 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

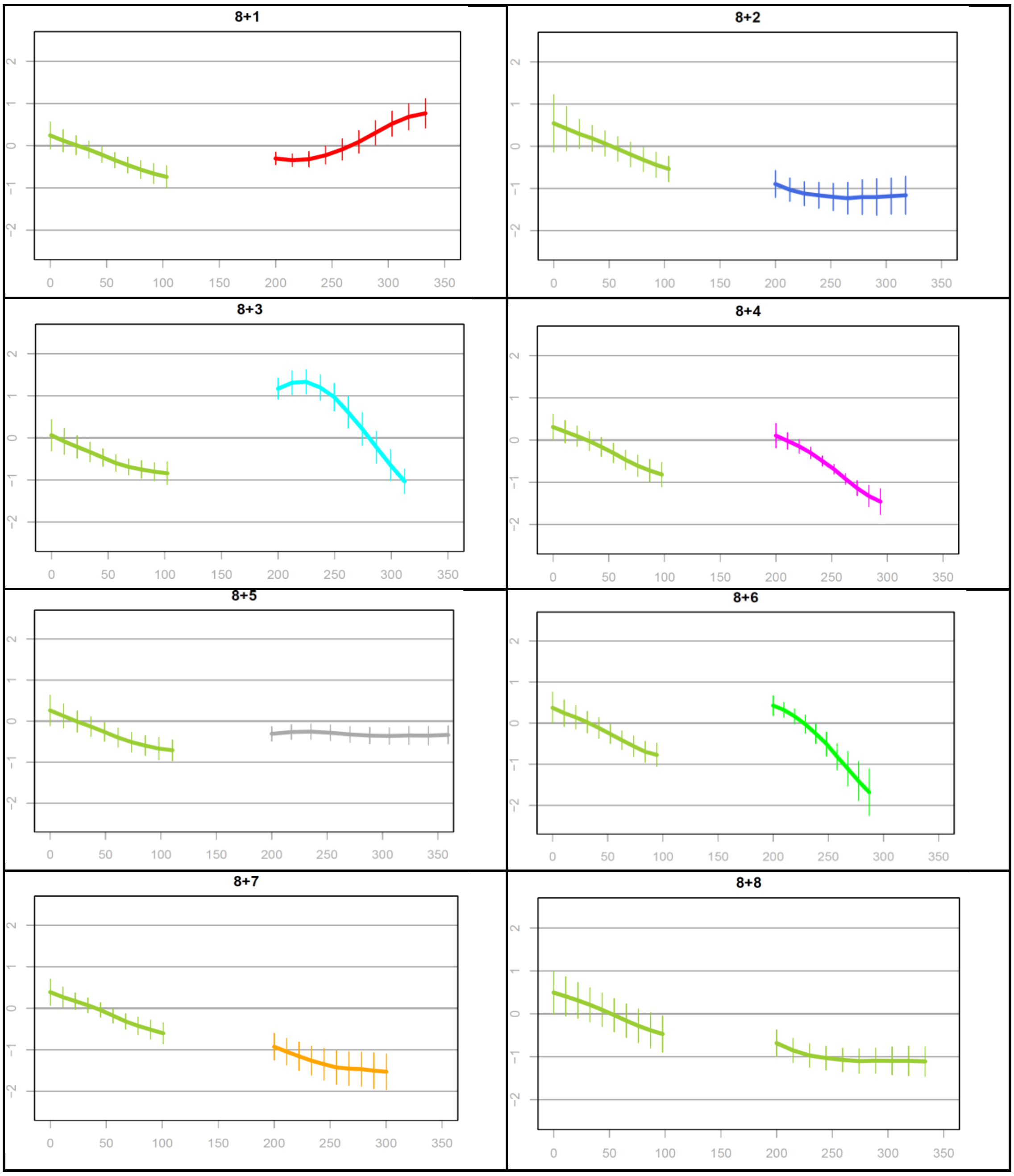


Figure 7-15. *Z*-score normalised F0 shapes of phrase-initial tone 8 across phrase-final tones from 21 speakers.

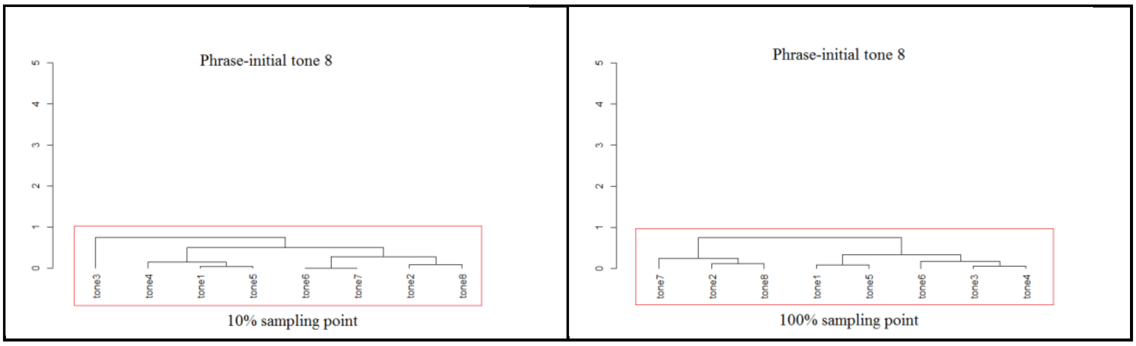


Figure 7-16. Clustering of normalised F0 levels of phrase-initial tone 8 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

7.1.9. Summary

As discussed, the acoustic F0 realisation of individual phrase-initial tones tends not to be categorically affected by tones in phrase-final position, but their offset F0 values appear to be phonetically sensitive to the following F0 contours. In general, the lower the F0 onset of phrase-final tones, the higher the F0 offset of the tones in the phrase-initial position. This phonetic sensitivity is statistically supported by the pairwise *t* test, as visualised in Figure 7-17 using the hierarchical clustering algorithm. The phrase-initial tones 1, 2 and 3 have two statistically significantly different offset variants with the value significantly higher preceding tones 2, 7, and 8 which present a low F0 onset.

According to the results of acoustic quantification and normalisation, and statistical testing across 64 tonal combinations, the whole linguistically phonetic F0 system of Zhangzhou phrase-initial tones is thus plotted in Figure 7-18, representing the central tendency of Zhangzhou as an independent variety. As shown in this figure, Zhangzhou has a relatively simple F0 inventory in terms of contour shapes—falling, level and rising—among the tones at the phrase-initial context. Nevertheless, questions remain with respect to how many normalised F0 levels are categorically distinguished, and how can they be determined scientifically?

For example, each tone, except tones 1, 2 and 3, has two putative targets of onset and offset, while tones 1, 2, and 3 have three putative targets of onset and offset; thus there logically would be 19 putative levels giving rise to 171 (=19*18/2) paired differences at either the 10% or 100% sampling point to be tested under the assumption that all levels are independent and identically normally distributed. The Bonferroni corrected alpha of 0.000292 has to be performed to achieve significance. Figure 7-19 visualises the result of pairwise *t* tests using the hierarchical clustering algorithm, in which the 19 putative levels are clustered into five classes at the threshold of 1. Nevertheless, several aspects need to be further noted.

- The diverse levels of normalised F0 contours, especially at the offset position, prose a challenge to both statistical testing and the Chao pitch notation system with respect to how many F0 levels are categorically distinguished. The pairwise *t*-test results do not satisfactorily cluster levels into reasonable categories. For example, the four contours from tones 1 and 2 are indicated differently, including [23], [22], [33], and [32], as indicated in Figure 7-19. The five levels of the Chao system appear insufficient to characterise different F0 levels in Zhangzhou.
- The F0 range of phrase-initial tones tends to be raised, ranging from above one standard deviation below the mid to two standard deviations above the mid. Correspondingly, the highest F0 level appears higher than the highest level found in phrase-final position, and the lowest F0 level is higher than the lowest level of phrase-final tones. Such an F0-raising phenomenon is considered bying induced by position in this study. In other word, the utterance-initial position has a raising effect on F0 realisation.

Incorporating the auditory observations, acoustic distributions, and reasonable portions of the statistical testing results, the F0 system of Zhangzhou phrase-initial tones is indicated in Table 7-

1, with *6* indicating the highest F0 level and *2* the lowest. The number *6* is used to denote the extra-high pitch/F0 level among the phrase-initial tones. As indicated in this table, the single dimension of F0 appears to be inadequate to differentiate the contours between phrase-initial tones 1 and 2, and the contours among phrase-initial tones 5, 7 and 8. The pitches of tones 1 and 2 both present a mid level contour [33]; similarly, the normalised F0 contours of tones 5, 7, 8 all present a mid falling contour [32]. It is thus imperative to investigate other phonetic parameters which may function with F0 to construct the phrase-initial tonal distinctions.

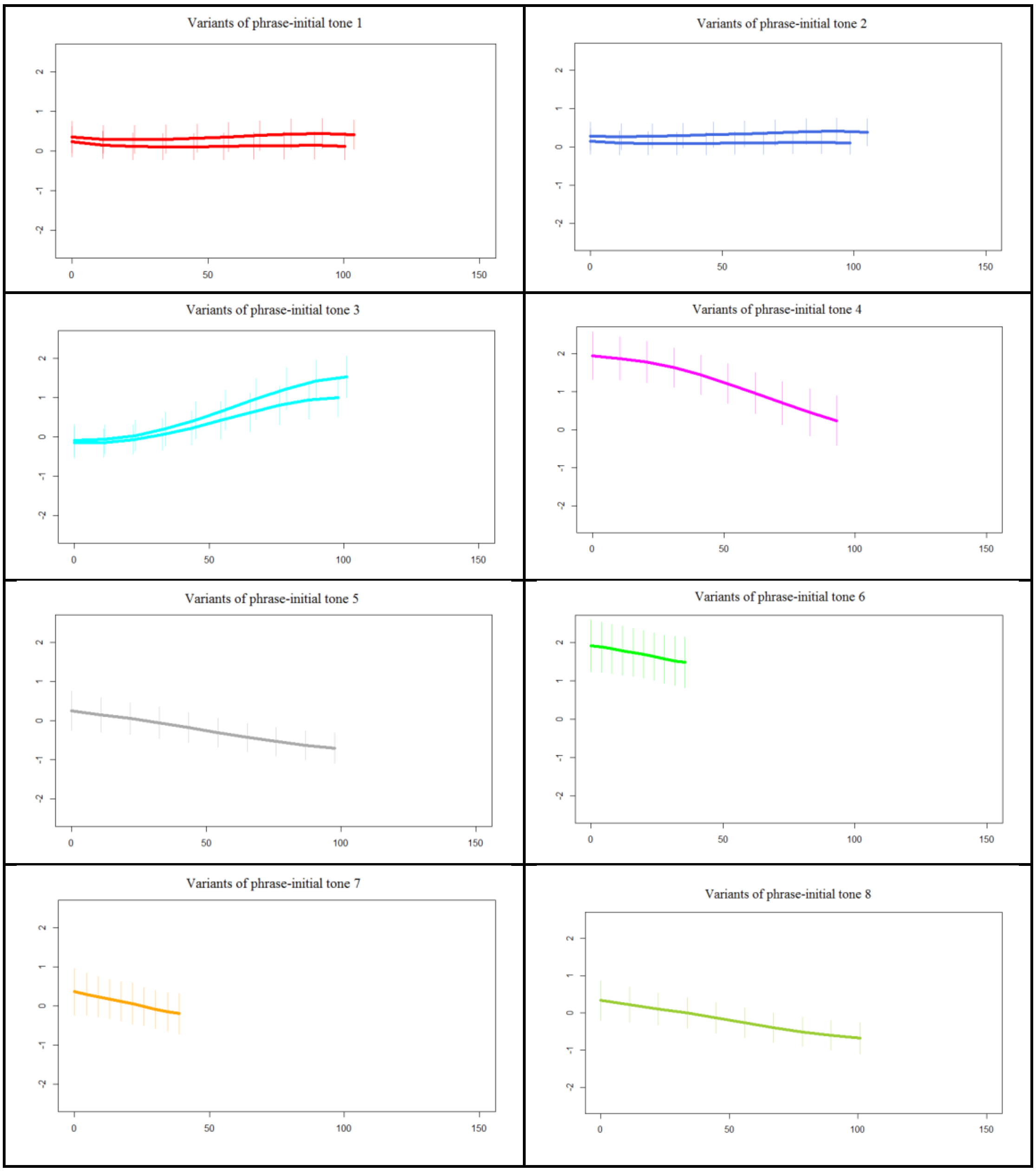


Figure 7-17. Summary of the F0 realisations of individual phrase-initial tones in Zhangzhou.

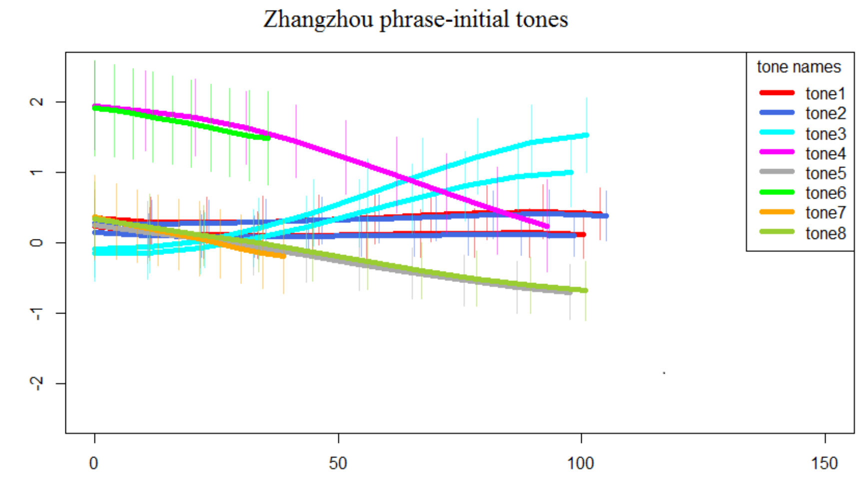


Figure 7-18. *Z*-score normalised F0 system of Zhangzhou phrase-initial tones.

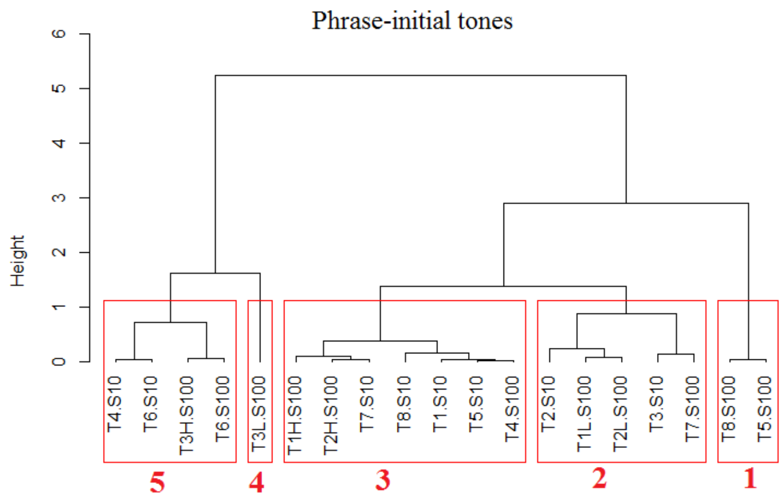


Figure 7-19. Clustering of normalised F0 levels of phrase-initial tones at both 10% and 100% sampling points based on pairwise *t*-tests.

Table 7-1. The numerical F0 representations of Zhangzhou phrase-initial tones

Tone	F0 contour description	Numerical representation
1	mid level	[33]
2	mid level	[33]
3	low rising	[24]/[25]
4	extra-high-falling	[63]
5	mid falling	[32]
6	stopped extra-high falling	[65]
7	stopped mid falling	[32]
8	mid falling	[32]

7.2. Interactions Among Acoustic Parameters

This section mainly addressed how various segmental and suprasegmental parameters interact to make the distinctions in acoustics among phrase-initial tones. The techniques of pairwise *t* testing and hierarchical clustering algorithm were applied to determine whether the duration realisations of individual phrase-initial tones may be affected by their following tones. The SplitsTree software was used to assess the relatedness between phrase-initial tones from the phonological perspective.

7.2.1. Duration

This subsection addresses three issues. (1) What are the length realisations of phrase-initial tones across 64 disyllabic combinations? (2) Are their length realisations affected by phrase-final tones? If so, to what extent are they affected, and what conditions the variations? (3) How many length levels are statistically contrastive among the phrase-initial tones?

The duration contours presented in this section have been normalised and expressed as a percentage of the average duration of all phrase-initial tones. Each phrase-initial tone also logically would have 28 (=8*7/2) paired length differences to be tested across eight phrase-final tones, assuming the length values are all independent and identically normally distributed. The Bonferroni corrected alpha of 0.00186 (=0.05/28) was consistently performed to control for the Type I Error and to achieve significance. The threshold of 1 was consistently employed to cluster the values and examine how many length levels are contrastive from the statistical point of view.

The normalised duration values and the values of pairwise *t*-testing result for individual phrase-initial tones are tabulated in Appendix B (Tables B10-B12). The methodologies of acoustic and statistical processing are described in Chapter 2.

7.2.1.1. Phrase-initial tone 1

In Chapter 6, the phrase-initial tone 1 was perceived as having very similar length across different disyllabic combinations. The acoustically quantified results basically justify the auditory description but show small variation in the normalised values across the phrase-final tones. For example, as plotted in Figure 7-20, the contour duration is slightly shorter before tone 3 than before tone 8. Nevertheless, the clustering results based on the pairwise *t* tests by effect sizes revealed no significant differences in the normalised duration values of this tone across eight phrase-final tones. As shown in Figure 7-21, the terminal nodes at the bottom represent the eight phrase-final tones, while the horizontal lines indicate the amount of difference in the normalised F0 of this phrase-initial tone 1. All the values in this figure are clustered into one level at the threshold of 1, indicating no statistically significant differences among them. Therefore, the duration realisations of phrase-initial tone 1 can be considered statistically undifferentiated across the eight combinations. In other words, the length realisations of phrase-initial tone 1 are not affected by their following tones either phonologically or phonetically.

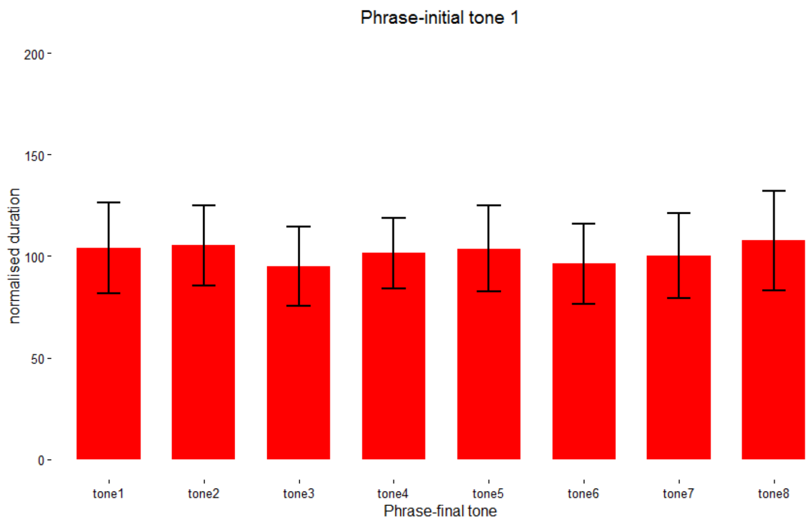


Figure 7-20. Normalised duration of phrase-initial tone 1 across phrase-final tones from 21 speakers.

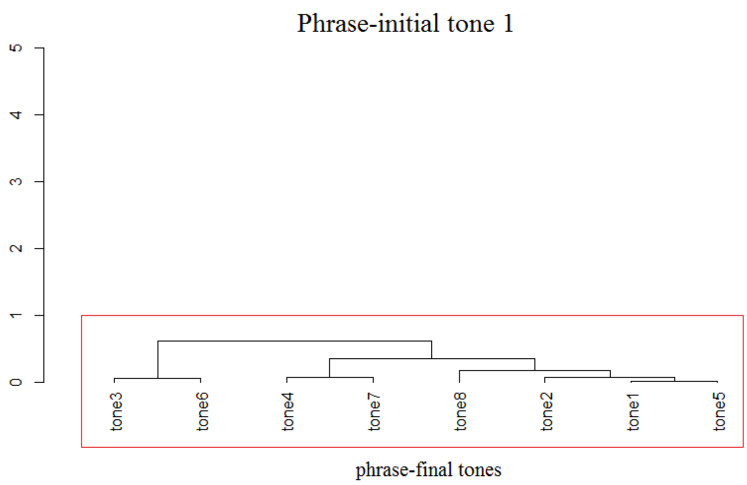


Figure 7-21. Clustering of normalised length levels of phrase-initial tone 1 across eight phrase-final tones based on pairwise *t* tests.

7.2.1.2. Phrase-initial tone 2

Phrase-initial tone 2 was described in Chapter 6 as having very similar length realisations across the phrase-final tones. The acoustically normalised results show a small degree of variation in the normalised values among eight combinations, as shown in Figure 7-22. For example, the duration before tone 6 is slightly shorter than it is before tone 5. Nevertheless, the acoustic differences are not statistically supported by the pairwise *t* tests. As indicated in Figure 7-23, the eight terminal nodes, representing the eight phrase-final tones, have been clustered into one class at the threshold of 1, indicating that the differences in the normalised F0 of this phrase-initial tone 2 are not statistically significant across the eight phrase-final tones.

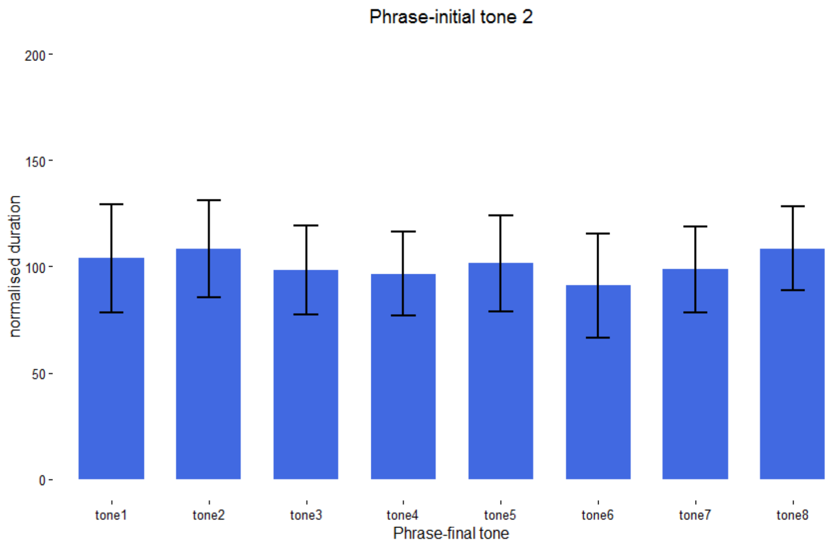


Figure 7-22. Normalised duration of phrase-initial tone 2 across phrase-final tones from 21 speakers.

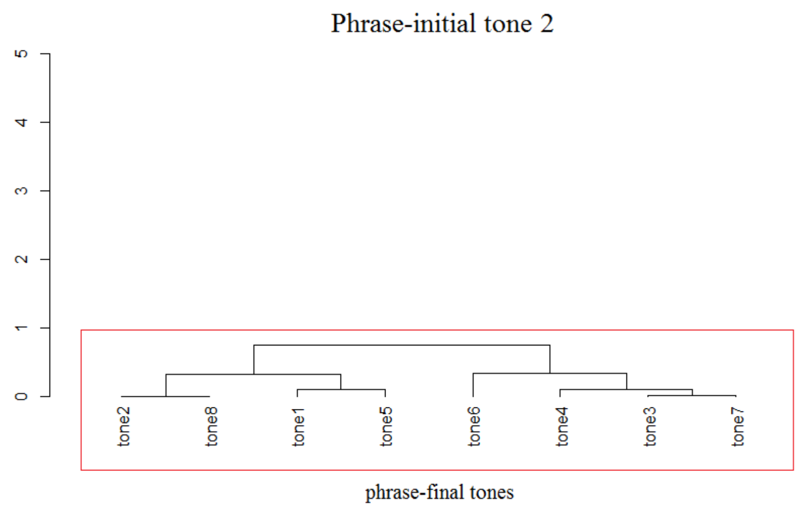


Figure 7-23. Clustering of normalised length levels of phrase-initial tone 2 across eight phrase-final tones based on pairwise *t* tests.

7.2.1.3. Phrase-initial tone 3

Phrase-initial tone 3 was perceptually evaluated as having very similar length across eight phrase-final tones. The results of the acoustic quantification and normalisation show some small variation in the normalised duration values across combinations. For example, as shown in Figure 7-24, the duration before tone 6 is slightly shorter than it is before any other tones. Nevertheless, no statistically significant differences are discovered by the pairwise *t* tests among the eight normalised duration values of the phrase-initial tone 3. As shown in Figure 7-25, the terminal nodes at the bottom, representing the eight phrase-final tones, have been clustered together at the threshold of 1, indicating that the differences in the normalised F0 of this phrase-initial tone 3 are not statistically significant across their following tones. In other words, its duration realisations are not affected by their following tones phonologically.

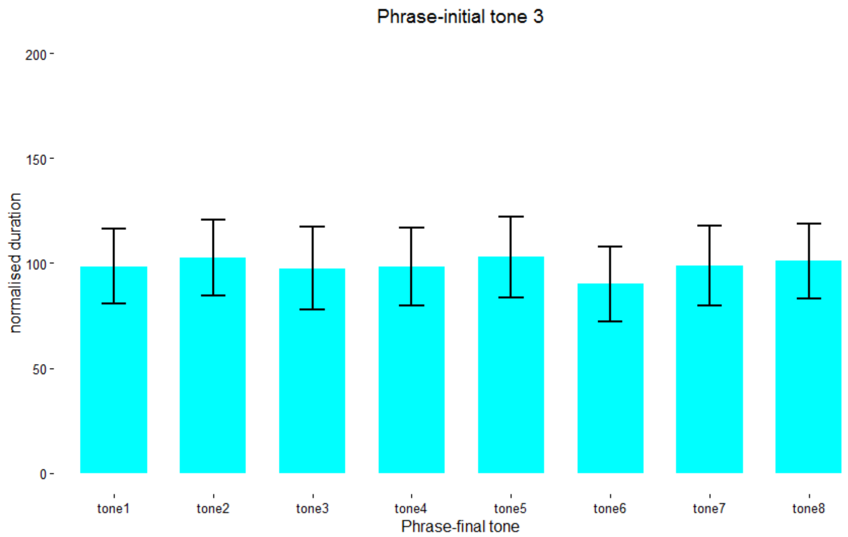


Figure 7-24. Normalised duration of phrase-initial tone 3 across phrase-final tones from 21 speakers.

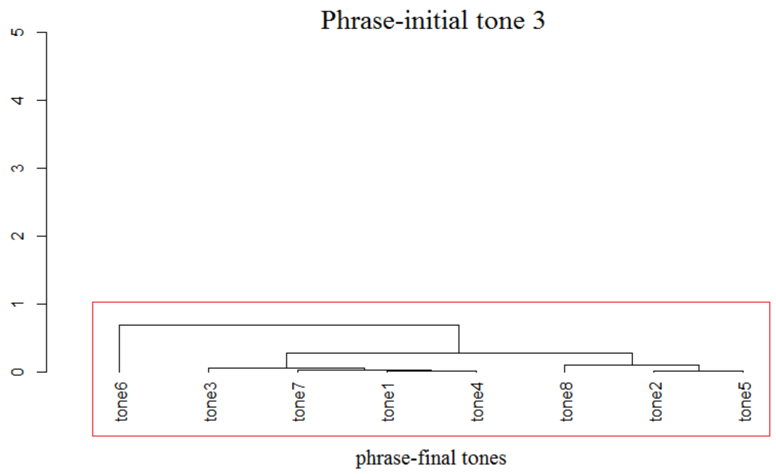


Figure 7-25. Clustering of normalised length levels of phrase-initial tone 3 across eight phrase-final tones based on pairwise *t* tests.

7.2.1.4. Phrase-initial tone 4

Perceptually, phrase-initial tone 4 has very similar length realisations across different phrase-final tones. Acoustically, a considerable amount of variation occurs among the normalised duration values. For example, as Figure 7-26 shows, the duration is apparently longer before tone 5 than before tone 6. Nevertheless, the pairwise *t*-testing results reveal no significant difference among the eight normalised values for the phrase-initial tone 4. As indicated in Figure 7-27, the eight terminal nodes, representing the normalised duration values of the phrase-initial tone 4 across eight phrase-final tones, have been clustered into one class at the threshold of 1, indicating that the durational differences are not statistically significant. Therefore, the auditory observation is statistically justified, although some acoustic variation exists.

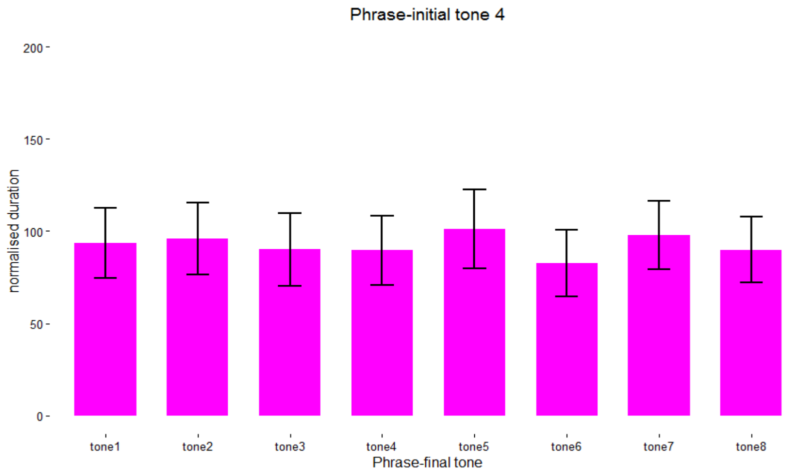


Figure 7-26. Normalised duration of phrase-initial tone 4 across phrase-final tones from 21 speakers.

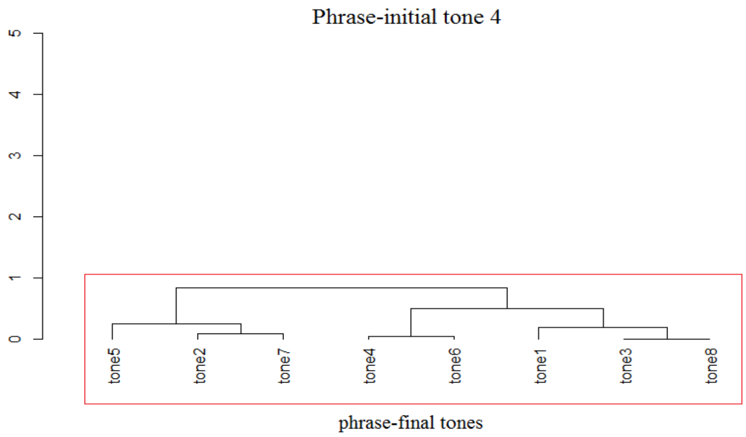


Figure 7-27. Clustering of normalised length levels of phrase-initial tone 4 across eight phrase-final tones based on pairwise *t* tests.

7.2.1.5. Phrase-initial tone 5

Phrase-initial tone 5 was perceptually observed as having very similar realisations across eight tonal combinations. Acoustically, as shown in Figure 7-28, the normalised duration values are not quite the same across the phrase-final tones. For example, the duration is shorter before tone 3 than it is before tone 2. Statistically, the pairwise *t*-test comparison results show significant differences among the eight normalised duration values. As Figure 7-29 indicates, the eight terminal nodes at the bottom, which represent the normalised duration values of this phrase-initial tone 5 across eight phrase-final tones, have been clustered into two levels at the threshold of 1. This clustering indicates that the two levels are statistically significantly different, with the values before tones 3 and 6 being significantly shorter than the values before other tones.

Therefore, the auditory observation is not quite justified either statistically or acoustically. In other words, the phrase-initial tone 5 can be marginally considered having two levels of length realisation, since the effect is not great, it is dependent on the selection of the threshold for significance.

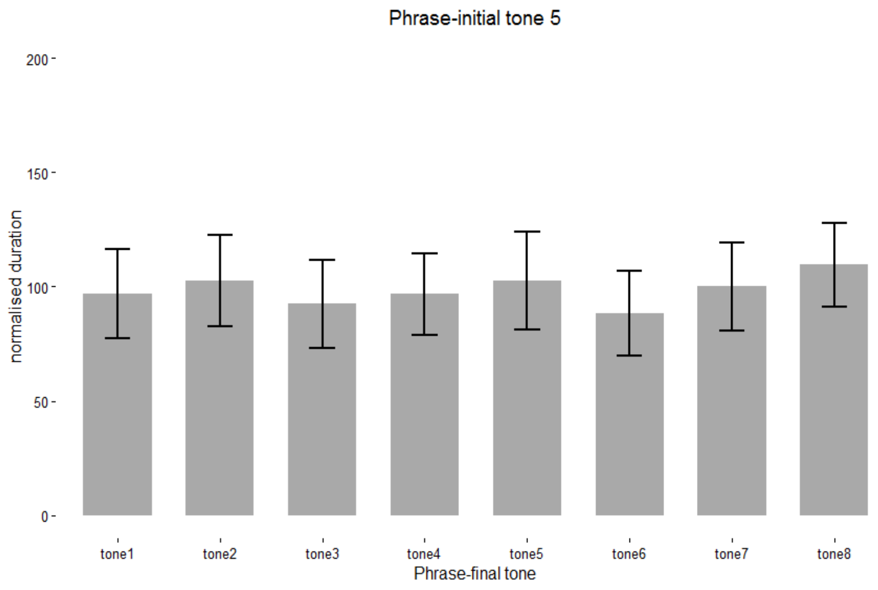


Figure 7-28. Normalised duration of phrase-initial tone 5 across phrase-final tones from 21 speakers.

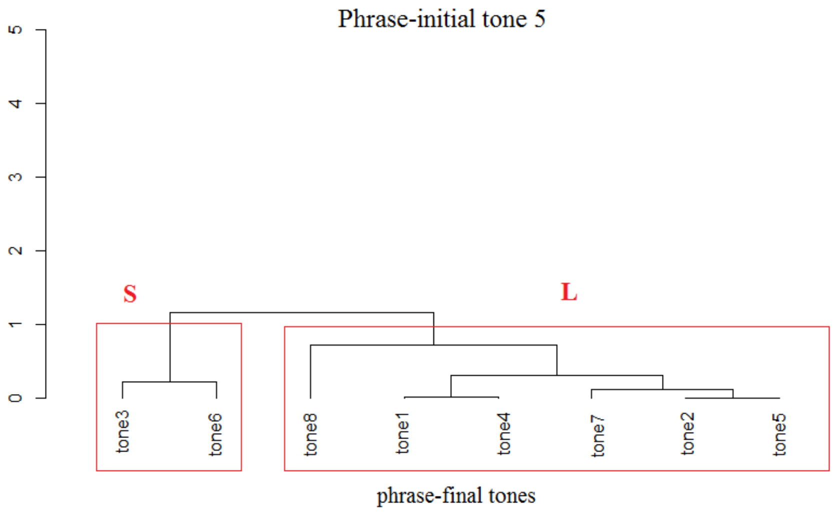


Figure 7-29. Clustering of normalised length levels of phrase-initial tone 5 across eight phrase-final tones based on pairwise *t* tests.

7.2.1.6. Phrase-initial tone 6

Phrase-initial tone 6 was indicated in Chapter 6 as being perceived as having an extra-short duration, regardless of its following tones. The acoustically quantified results basically justify the auditory description, but the normalised duration values still show a small amount of variation. As indicated in Figure 7-30, the duration before tone 7 is slightly longer than it is before tone 6. Nevertheless, the differences in the normalised values of this tone across the phrase-final tones are not statistically significant according to the pairwise *t*-test comparisons. As shown in Figure 7-31, the eight terminal nodes at the bottom, representing the duration values of this tone across eight tonal combinations, are clustered into one group at the threshold of 1, which was essentially expected, since very short durations have less scope for variation in absolute terms.

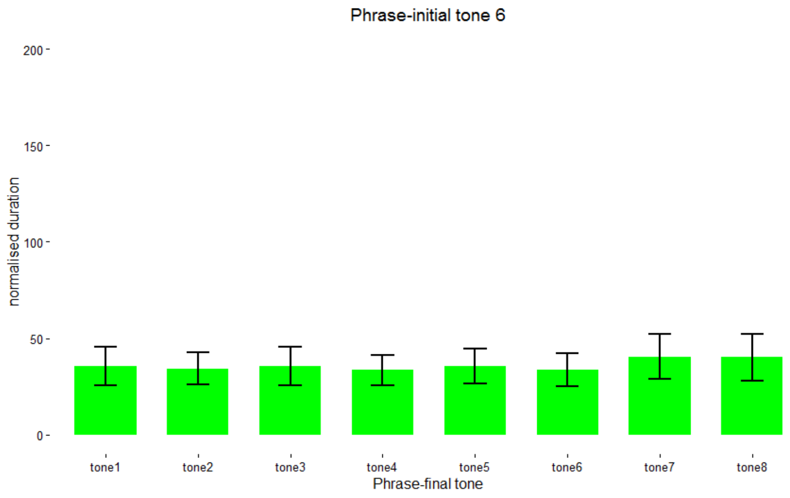


Figure 7-30. Normalised duration of phrase-initial tone 6 across phrase-final tones from 21 speakers.

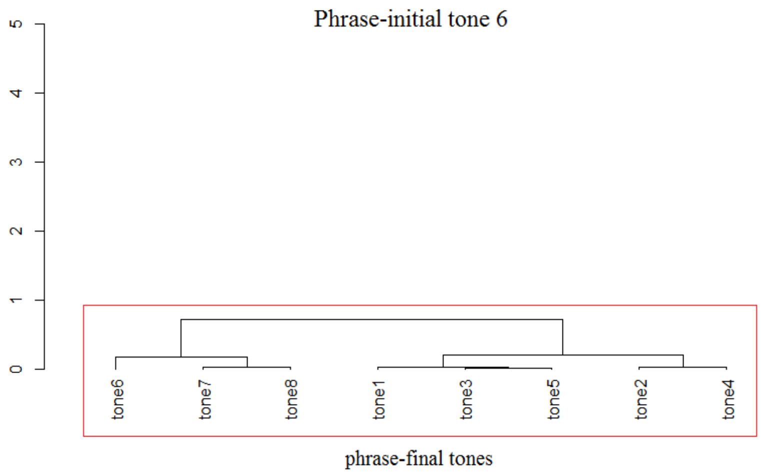


Figure 7-31. Clustering of normalised length levels of phrase-initial tone 6 across eight phrase-final tones based on pairwise *t* tests.

7.2.1.7. Phrase-initial tone 7

Determining whether the length realisations of phrase-initial tone 7 are affected by following tones is perceptually difficult because its length is consistently realised as extra-short across different tonal combinations. The auditory impression is effectively justified by the acoustically quantified results. As shown in Figure 7-32, the normalised duration values are all around 50% of the average duration of all tones from all speakers, and the variations among the values are very small. The statistical testing results also support the auditory observation because no significant difference is found among eight duration values. As indicated in Figure 7-33, the eight terminal nodes at the bottom are clustered into one group at the threshold of 1, showing that the duration values of this phrase-initial tone 7 are statistically indistinguishable across the eight tonal combinations. In other words, its realisations are not affected by the tones at the phrase-final context at both the phonological and phonetic levels.

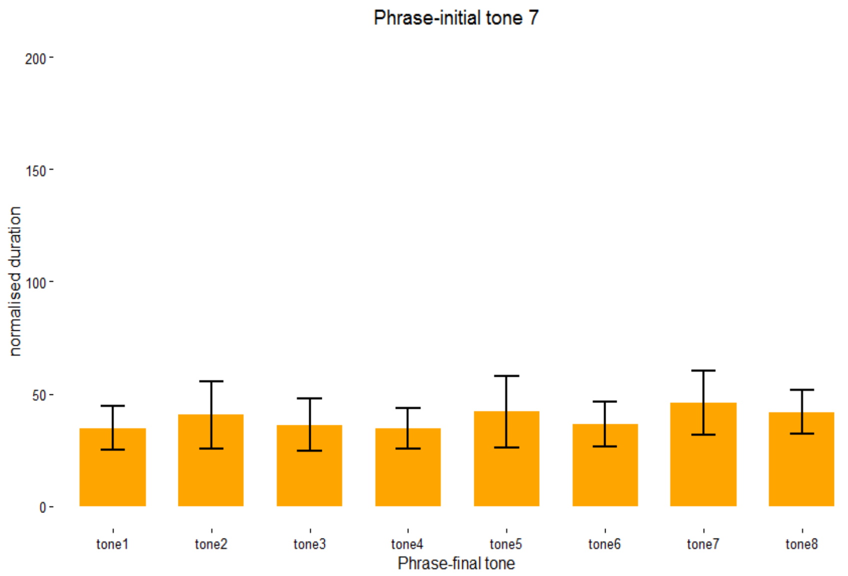


Figure 7-32. Normalised duration of phrase-initial tone 7 across phrase-final tones from 21 speakers.

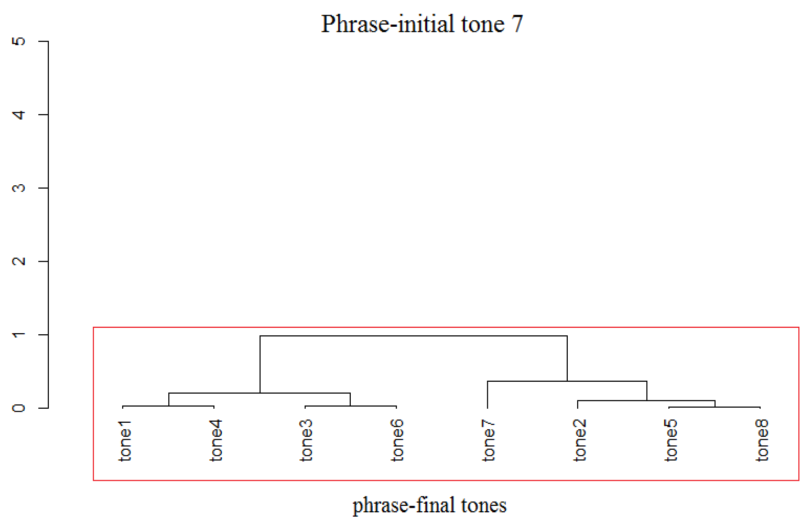


Figure 7-33. Clustering of normalised length levels of phrase-initial tone 7 across eight phrase-final tones based on pairwise *t* tests.

7.2.1.8. Phrase-initial tone 8

Phrase-initial tone 8 was indicated in Chapter 6 as being perceived as having very similar realisations in terms of length across the phrase-final tones. Acoustically, a small amount of variation occurs in the acoustically normalised duration values with respect to different tonal combinations. For example, as shown in Figure 7-34, the duration is apparently longer before tone 5 than before tone 6. Nevertheless, the results of pairwise *t* test comparisons justify the auditory observation because no significant difference is found among the eight normalised duration values. As clustered in Figure 7-35, the eight terminal nodes are grouped together at the threshold of 1, indicating that the normalised duration values of this phrase-initial tone 8 are statistically indistinguishable across its following tones.

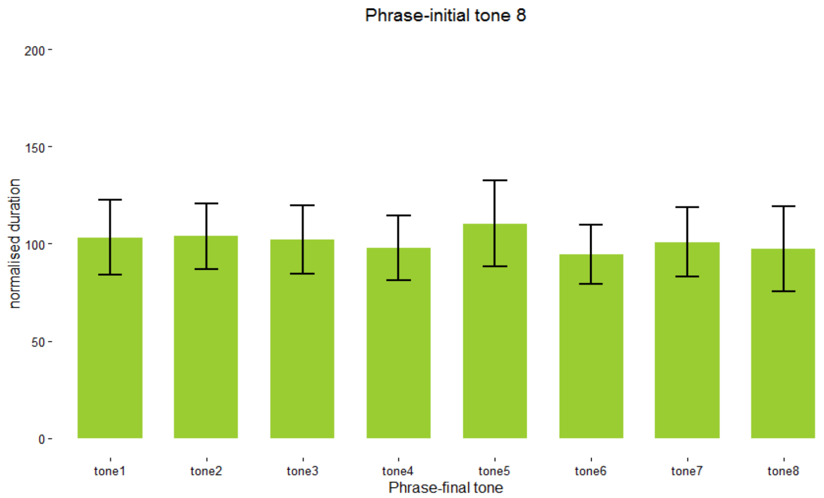


Figure 7-34. Normalised duration of phrase-initial tone 8 across phrase-final tones from 21 speakers.

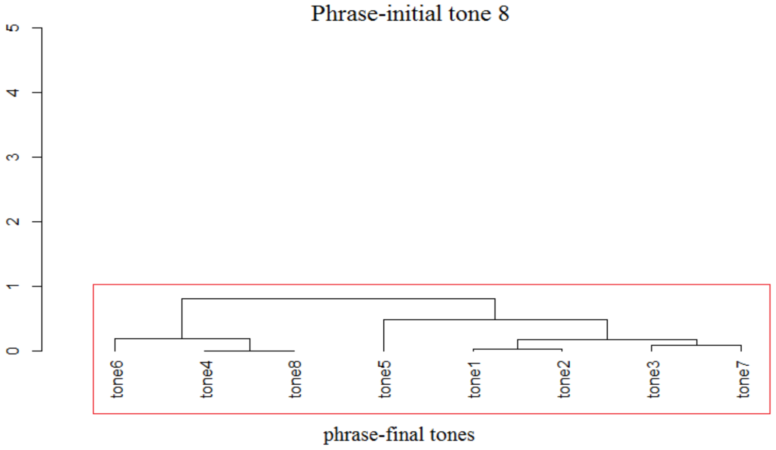


Figure 7-35. Clustering of normalised length levels of phrase-initial tone 8 across eight phrase-final tones based on pairwise *t* tests.

7.2.1.9. Summary

As described above, the length realisations of phrase-initial tones are not quite the same across different tonal combinations. In general, the duration tends to be measurably shorter before falling contours (e.g., tones 3, 4, and 6), especially before the stopped tone 6, while they are more likely to be longer before dominantly level contours (e.g., tones 2, 5, and 8), especially before the mid-level pitched tone 5. Nevertheless, the acoustic differences in the normalised duration values are generally not statistically significant, according to the pairwise *t*-test comparisons by effect sizes across 64 combinations, except for phrase-initial tone 5 showing two marginally significant levels, as shown in Figure 7-36.

In the figure, the nine phonetic variants of length realisation of phrase-initial tones are expressed as a percentage of the average normalised duration value of all phrase-initial tones over 64 tonal combinations and over 21 speakers. As shown, tonal length does vary across different phrase-initial tones. The two stopped tones 6 and 7 tend to have the shortest length. Nevertheless, whether the length differences among the stopped tones and among the unstopped tones are statistically significant, and as well as how many length levels are contrastive among the tones at the phrase-initial context, need to be further tested using the pairwise *t* test comparison.

To address this issue, an exhaustive pairwise *t*-test comparison was conducted on 36 (=9*8/2) paired differences in the normalised duration values with the Bonferroni corrected alpha of 0.001389 (=0.05/36) to control for the Type I Error and to achieve significance. The statistical result was further visualised using the hierarchical clustering algorithms as shown in Figure 7-37. The nine phonetic variants of length realisations of phrase-initial tones across 64 tonal combinations are clustered into two levels, with the values before tones 6 and 7 being statistically significantly shorter than the values before other tones.

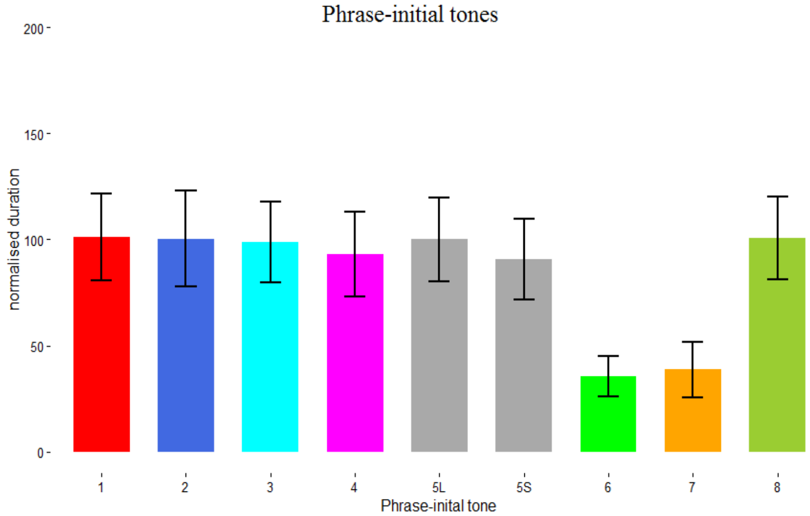


Figure 7-36. Summary of normalised F0 duration of individual Zhangzhou phrase-initial tones

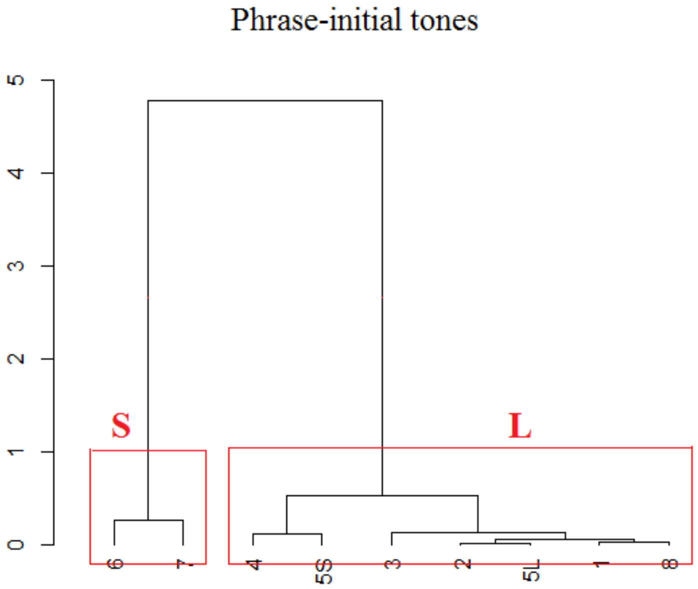


Figure 7-37. Clustering of normalised tonal length levels of phrase-initial tones across eight phrase-final tones based on pairwise *t* tests.

Based on the acoustic quantification and statistical clustering results, a linguistic-tonetic length representation of Zhangzhou phrase-initial tones is shown in Table 7-2. The number of length contrasts is significantly reduced to only two among the phrase-initial tones. The length contrasts among the unstopped tones have been neutralised to medium level while the contrasts among the stopped tones have been neutralised to extra-short level. The extra-short property of the two stopped tones are considered to be motivated by the preservation of obstruent codas in the phrase-initial syllables (see Section 7.2.4).

Table 7-2. Length system of Zhangzhou phrase-initial tones based on acoustics and statistics

Tone	F0 contour description	Length	Notation
1	mid level [33]	medium	[V]
2	mid level [33]		
3	low rising [24]/[25]		
4	extra-high-falling [63]		
5	mid falling [32]		
8	mid falling [32]		
6	stopped extra-high falling [65]	extra short	[V̆̆]
7	stopped mid falling [32]		

7.2.2. Vowel quality

In Chapter 6, all vowels were perceived consistently as monophthongs without changing their qualities in the phrase-initial syllables. From the articulatory point of view, the monophthong realisation indicates only one articulatory trajectory because the tongue generally remains in the same position of the oral cavity throughout the articulation of related sounds.

From the acoustic point of view, the monophthongs indicate relatively stable manifestations of the first and second formants not changing across time. Figures 7-38 and 7-39 show the acoustic realisations of Zhangzhou vowels with respect to phrase-initial stopped and unstopped tones. Because only high and low vowels are found in the stopped tones, no examples of the mid vowels are provided.

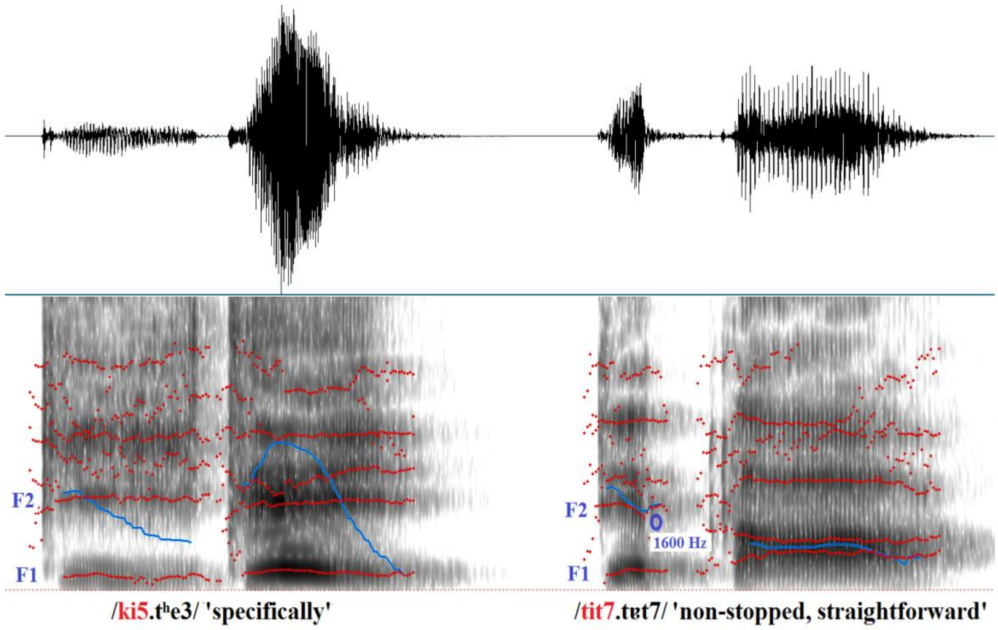


Figure 7-38. Acoustic realisations of high vowel /i/ in phrase-initial unstopped tone 5 and stopped tone 7 (WYF, male).

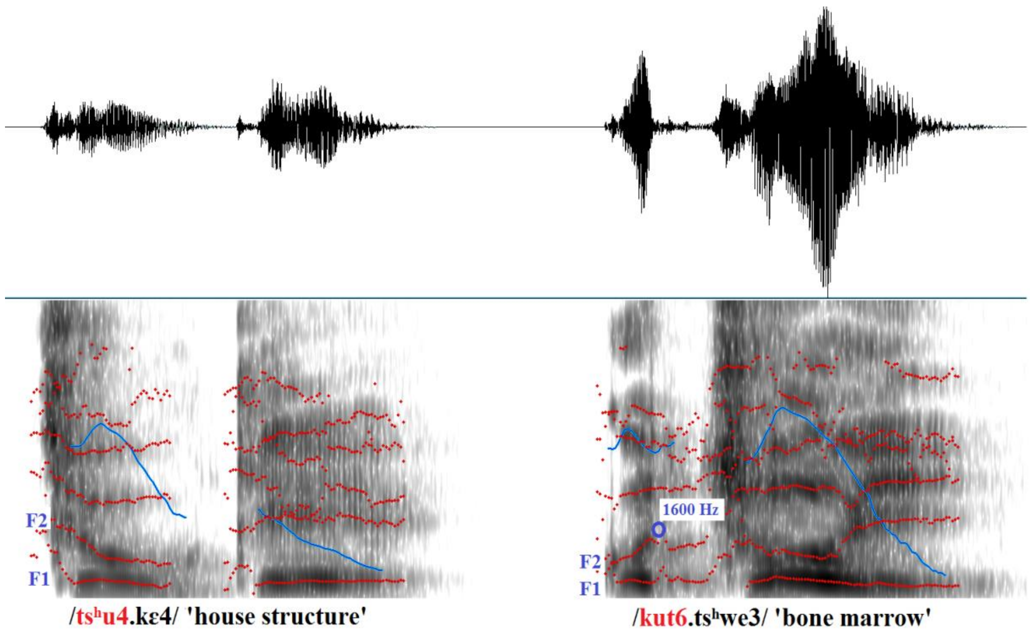


Figure 7-39. Acoustic realisations of high vowel /u/ in phrase-initial unstopped tone 4 and stopped tone 6 (WYF, male).

In Figure 7-38, the first morphemes of the two disyllabic tokens have the same high vowel /i/ at the underlying level and similar mid-falling F0 contours at the surface. The formant patterns of the /ki5/ morpheme in the unstopped tone 5 are relatively steady and horizontal without significant change across time, indicating a monophthong realisation of this high vowel.

In the second token /tit7.tɐt7/ 'non-stopped', the F1 curve of the /tit7/ morpheme is relatively stable across time while the F2 is largely steady but presents a slightly falling trend toward a frequency around 1600 Hz during the last 5%. Nevertheless, the falling F2 curve does not indicate a backward movement of the tongue for a diphthong production in this context but rather indicates a formant transition to the obstruent coda at the alveolar ridge. Therefore, the acoustics justify a monophthong realisation of the high vowel in the stopped tone.

In Figure 7-39, the first morphemes of the two tokens have the same underlying high vowel /u/ and similar high-falling F0 contours. The F1 curve of the morpheme /tsʰu4/ in the unstopped tone 4 is relatively stable across time while its F2 curve is largely stable, except for the first 10% showing a falling tendency. However, it is reasonable to assume it is perturbed by the aspirated onset [tsʰ]. The acoustic manifestations essentially indicate a monophthong realisation of this high vowel /u/.

In the second token /kut6.kɛ4/ 'bone skeleton', the F1 curve of the morpheme /kut6/ is relatively stable across time while the F2 presents a gradual rising tendency toward a frequency range around 1600 Hz. This rising F2 contour does not indicate a fronting movement of the tongue within the oral cavity; instead, it indicates the formant transition to the alveolar obstruent coda. Therefore, the production is characteristic of monophthong in this stopped tone.

Therefore, although the formant patterns of each token are not exactly the same and both present a different dynamic portion influenced by the production of adjacent segments, the diphthongisation manifestations as noted for equivalent vowels in the citation tones are not observed in this context.

7.2.3. Voice quality

Four different phonation types—modal, breathy, creaky, and falsetto—in phrase-initial syllables were discussed in Chapter 6 from the auditory point of view. The distributions of modal, breathy, and creaky phonation are simultaneously conditioned by the vowel quality, pitch and tone while the occurrence of falsetto voice is purely constrained by tone, because vowels that can occur in the phrase-initial stopped tone 6 are all perceived as having a falsetto voice.

The articulatory correlations between vowel height, pitch, and the non-modal phonation of creaky and breathy were discussed in Chapter 5. For example, the production of high vowel and breathy voice share common articulatory gestures of a lowered larynx and distended supraglottal cavity with a raised tongue body and advanced root. The correlation between falsetto voice and the phrase-initial tone 6 is also understandable. The pitch of the phrase-initial tone 6 is extra short and high, indicating a highly rapid vibration of the vocal folds within a very short time frame. Similarly, the falsetto voice involves particularly thin-edged glottal margins of the vocal folds, which are stretched longitudinally from front and back during the production and which contribute to generating sounds with abnormally high pitch (Laver, 1994; Baart, 2010). Therefore, the falsetto voice and tone 6 share similar laryngeal settings of vocal folds that rapidly vibrate and create the impression of high pitch.

These differences in articulatory configurations are expected to generate differences in acoustic manifestations. Figures 7-40, 7-41, and 7-42 show how the acoustic realisations of voice quality of the phrase-initial syllables vary with respect to vowel quality, pitch and/or tone in Zhangzhou. The illustrated vowels include high, mid, and low. The F0 contours include rising [35], level [33], mid fall [32], extra-high fall [63], and extra-short and high fall [65]. The waveforms and spectral slices are extracted from the last 10% of the related sounds.

Figure 7-40 concerns the voice quality realisations of the high vowel /i/ in the phrase-initial syllables across five different F0 contours. In the rising [35], level [33], mid-falling [32], and extra-high falling [63] contexts, the waveforms are all superimposed with very dense and random glottal fluctuations over time while the amplitude differences of H1-H2 are consistently positive and the H1 value is always higher than any other amplitude value across frequency ranges. The acoustic manifestations of dense glottal fluctuation and steeply positive H1-H2 values justify the auditory observation concerning the correlation between high vowel and breathy voice.

In the extra-short and high-falling contour [65], the manifestations are neither breathy nor creaky. The glottal pulses are neither superimposed with dense fluctuations nor irregularly distributed. The amplitude value of the first harmonic is neither the highest across the frequency ranges nor similar to that value of the second harmonic. Instead, the manifestations seem characteristic of modal phonation. Nevertheless, the individual harmonics in the spectrum are more clearly identified and separated with more rounding and smoothing shapes. The number of glottal pulses within a given time (e.g., 0.02 second) is greater than the number in other F0 contexts, indicating much faster vibrations of vocal folds. Therefore, this high vowel appears to present a falsetto voice rather than breathy, creaky, or modal phonation in this extra-short and high-falling F0 environment (tone 6).

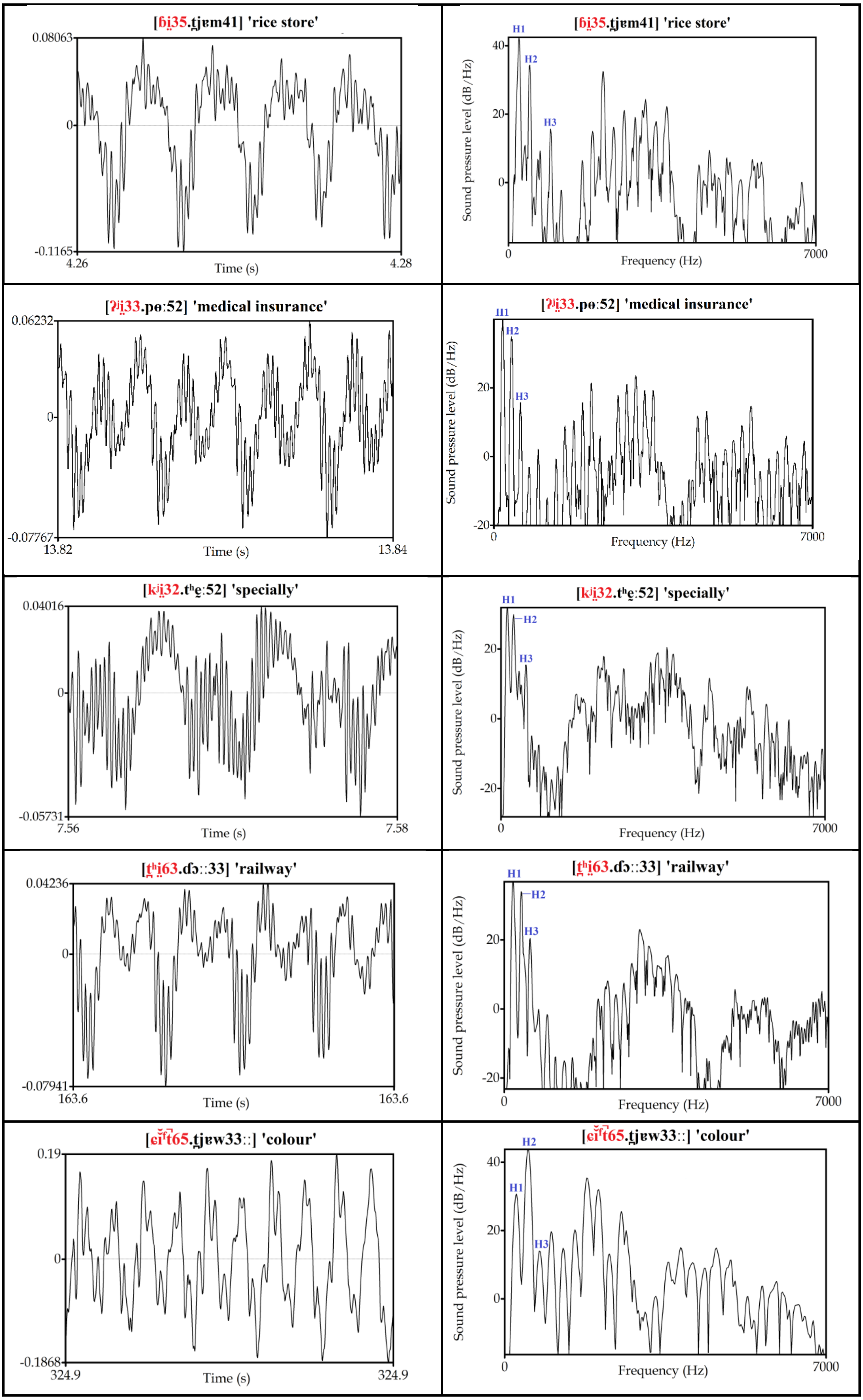


Figure 7-40. Phonation realisations of phrase-initial high vowel /i/ in five different F0 contexts (WYF, male).

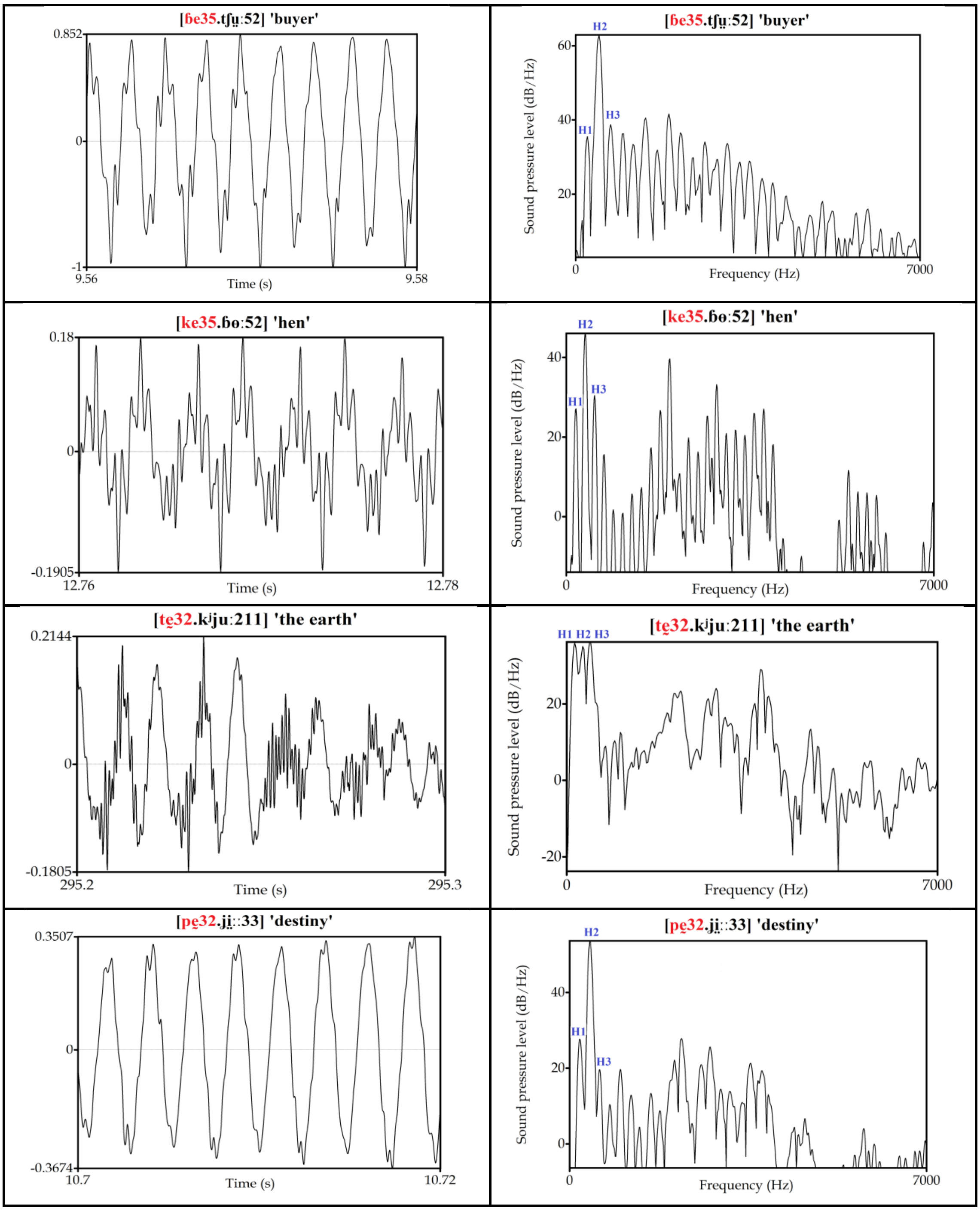


Figure 7-41. Phonation realisations of phrase-initial mid vowel /e/ in four different F0 contexts (WYF, male).

Figure 7-41 indicates the voice quality realisations of mid vowel /e/ in four F0 contours. No example is provided with the extra-short and high-falling F0 because the mid vowels are not found in the stopped tones in Zhangzhou. In the rising [35], level [33], and extra-high falling [63] contours, the glottal pulses are consistently periodic and regular without widespread fluctuations. The spectral tilt of H1-H2 is steeply negative. The two manifestations indicate a modal voice of the mid vowel in these three tonal environments. Nevertheless, in the mid-falling context [32], the glottal pulses become aperiodic and irregularly spaced, and the amplitude difference between H1 and H2 is around zero, indicating a creaky voice accompanying the production of the mid vowel in this mid-falling context.

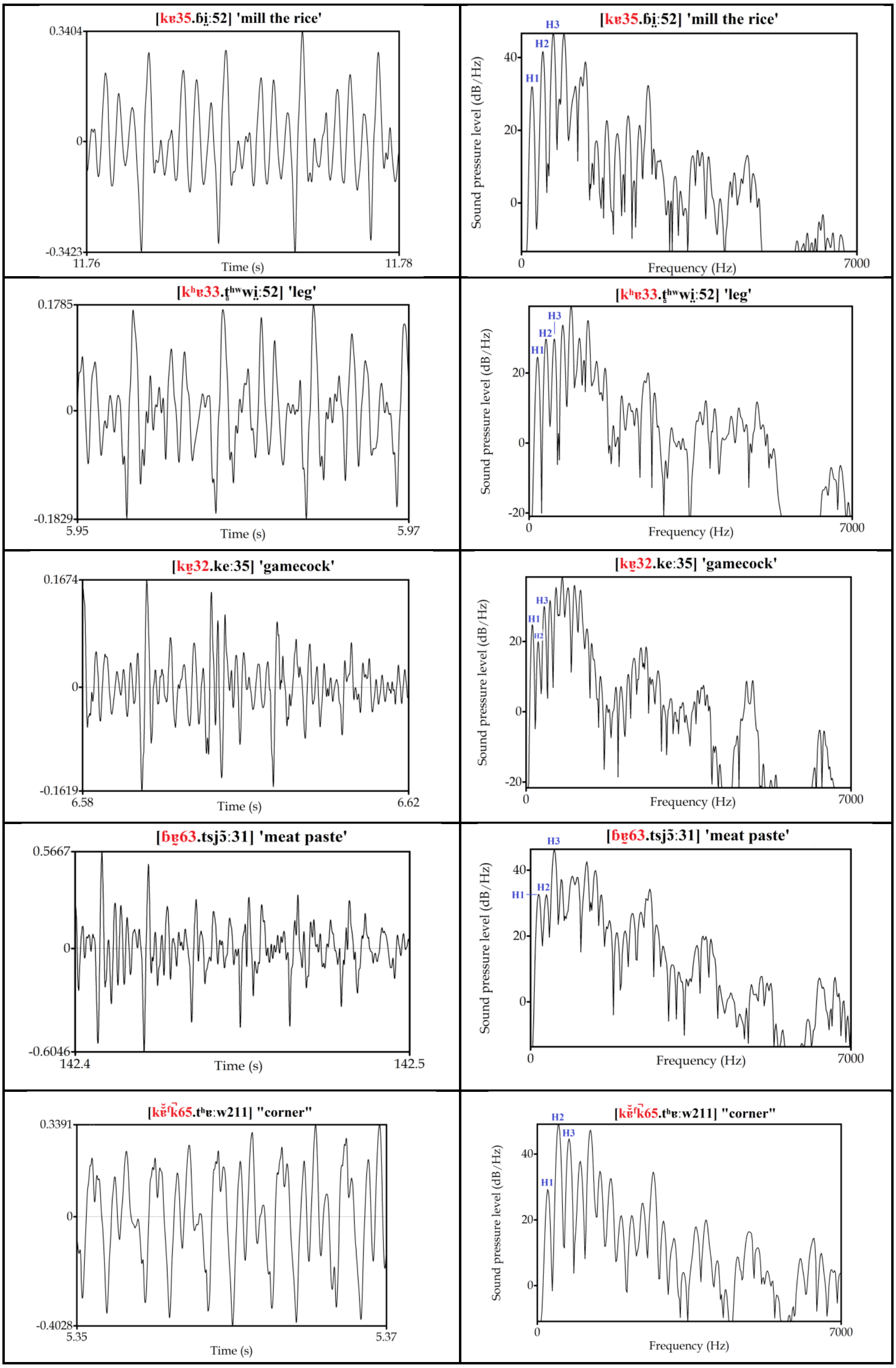


Figure 7-42. Phonation realisations of phrase-initial low vowel /ɐ/ in five different F0 contexts (WYF, male).

Figure 7-42 shows three different types of acoustic manifestations for the voice quality realisation of low vowel /ɐ/. (1) In the rising and level F0 contours, the intervals between glottal pulses are periodic and evenly spaced while the amplitude of the second harmonic is consistently higher than that of the first harmonic. The two manifestations justify the auditory impression of the modal voice realisation of the low vowels in non-falling tonal environments. (2) In the extra-high falling [63] and mid-falling [32] contours, the intervals of glottal pulses tend to be irregular and less frequent during the last 10% of the utterance. The H1 has a similar amplitude value as the H2 in the [63] contour, a typical property of creaky phonation. The H1 value is slightly higher than that of the H2 in the [32] contour but considerably lower in other harmonics, falsifying a breathy voice manifestation. Combined with its irregular realisation on the waveform, this vowel is acoustically creaky in the [32] F0 context.

(3) In the extra-short and high-falling [65] context, the acoustic manifestations of the low vowel are neither breathy nor creaky voice for two reasons. First, the glottal pulses are neither irregularly distributed nor superimposed with random fluctuations. Second, the amplitude value of H1 is neither steeper than that of H2 nor similar to that of H2. On the contrary, the manifestations seem similar to that of modal phonation, but with some differences. For example, the number of glottal pulses within 0.02 second is much greater than the number occurring in modal phonation. The H2 value is the highest among all harmonics, rather than lower or closer to the H2 as shown in modal phonation in the rising and level contexts. Therefore, this vowel acoustically presents characteristics of another type of phonation—falsetto, which is associated with extra high and short F0/pitch level.

7.2.4. Syllable coda

In contrast to the non-realisation of obstruent codas in the monosyllabic setting, the obstruent codas were described largely as perceivable in phrase-initial syllables in Chapter 6. From the articulatory point of view, the preservation of obstruent codas indicates an active formation of oral constriction at a particular point of the oral cavity, whereby the airstream is momentarily blocked. From the acoustic point of view, the temporary oral constriction for obstruent coda production indicates the formant frequency transitions to a particular frequency range at the end of the preceding vowel. Figure 7-43 shows the acoustic realisations of three different types of obstruent codas in the phrase-initial syllables with the same nucleus and tone at the underlying level.

As indicated in Figure 7-43, the formant patterns of the low vowel [ɐ] are considerably different across the last 5% of the three phrase-initial morphemes being investigated. In the /tsɐp7/ morpheme, the F2 curve shows an apparently falling tendency to a frequency range around 850 Hz, indicating a postvocalic-bilabial obstruent coda. In the /pɐt7/ morpheme, the F2 curve is slightly raised towards a frequency range around 1600 Hz, indicating an alveolar place of articulation. In the /ɓɐk7/ morpheme, the F2 curve noticeably increases while the F3 curve abruptly drops, but the two curves converge at a frequency range of about 2100 Hz, indicating a velar obstruent coda.

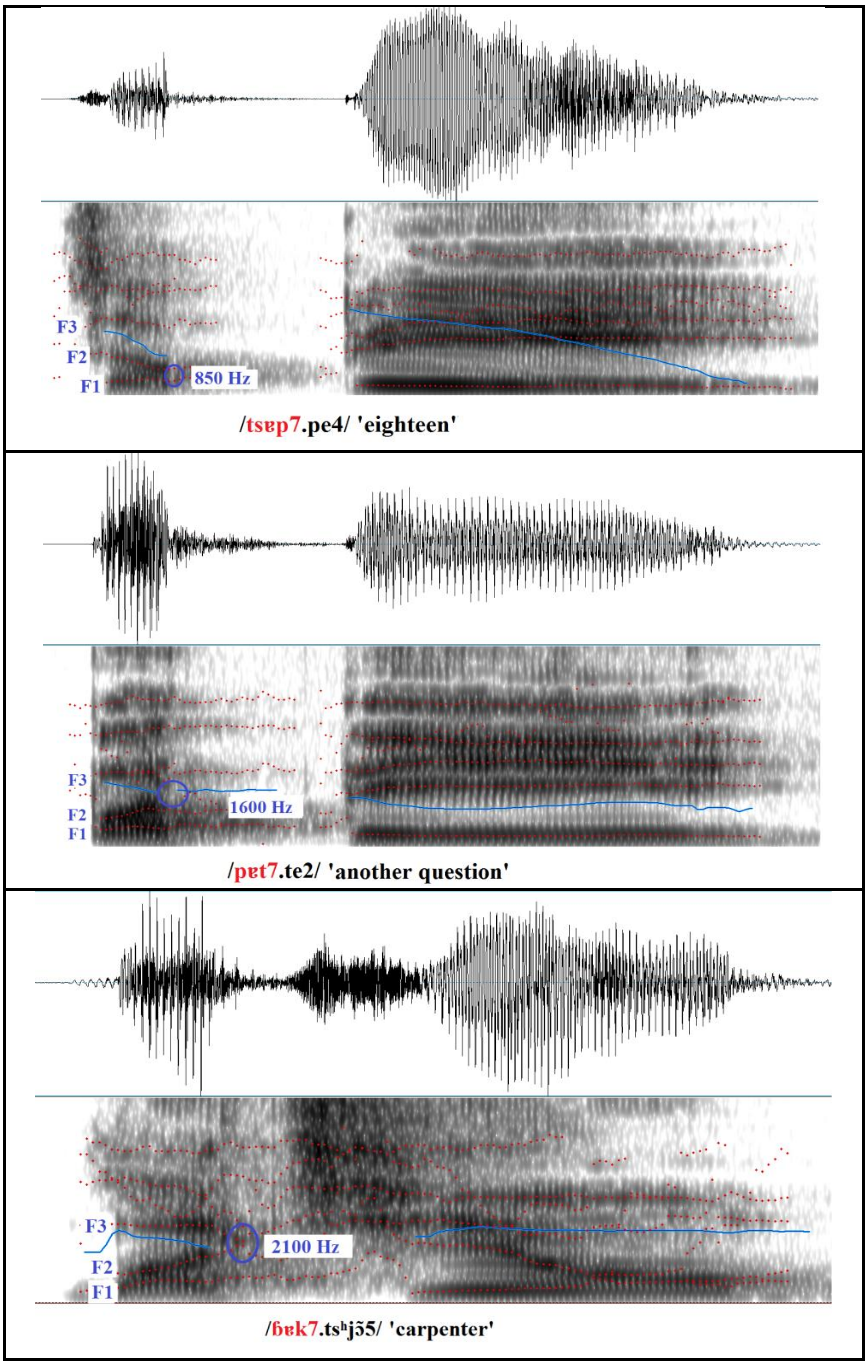


Figure 7-43. Acoustic realisations of obstruent codas in phrase-initial morphemes (WYF, male).

Therefore, the preservation of obstruent codas in phrase-initial morphemes is acoustically justified by the manifestations of formant transitions, and the target frequency ranges the formants transit to. In addition, the preservation of obstruent codas in phrase-initial morphemes also gives rise to several phonetic effects on the whole syllables, including syllable shortening, monophthong realisation of high vowels, and falsetto voicing of vowels in the stopped tone 6.

- The syllable shortening effect was addressed in Section 7.2.1. The stopped tones 6 and 7 both present an extra-short property in terms of length, especially tone 7, whose duration is statistically longer than many other tones in citation but shortest in the phrase-initial position.
- The monophthong realisation of high vowels in the stopped tones was addressed in Section 7.2.2. Contrary to high-vowel diphthongisation in the stopped tones in the monosyllabic setting, the high vowels are all realised as monophthongs across all phrase-initial tones.
- The effect of falsetto voicing of vowels was discussed in Section 7.2.3, in which vowels that can occur in the stopped 6 are shown in falsetto voice with an extra-short and high-falling F0 contour.

7.2.5. Summary

According to the descriptions in this section, a bundle of segmental and suprasegmental parameters function together to code the tonal realisations and distinctions among the phrase-initial syllables. Table 7-3 shows multidimensional realisations of Zhangzhou phrase-initial tones. Each tone consists of a bundle of phonetic features that interact in a complicated but also systematic way. For example, the phrase-initial tone 7 indicates a mid falling contour [32], extra short duration, no diphthongisation, breathy high vowel, creaky low vowel, and realised obstruent coda; while phrase-initial tone 1 suggests a mid-level contour [33], medium length, no diphthongisation, breathy high vowel, modal-voiced low vowel, and sonorant coda. Nevertheless, questions remain as to how these phrase-initial tones are related to each other in terms of their multidimensional realisations from the phonological point of view, as well as how the patterns hidden within this framework can be revealed and interpreted.

Table 7-3. Multidimensional realisations of Zhangzhou phrase-initial tones

Tone	F0/Pitch	Length	Diphthong	High vowel	Low vowel	Syllable coda
1	[33]	medium	-	breathy	modal	sonorant
2	[33]	medium	-	breathy	modal	sonorant
3	[24]/[25]	medium	-	breathy	modal	sonorant
4	[63]	medium	-	breathy	creaky	sonorant
5	[32]	medium	-	breathy	creaky	sonorant
6	[65]	extra short	-	falsetto	falsetto	realised obstruent
7	[32]	extra short	-	breathy	creaky	realised obstruent
8	[32]	medium	-	breathy	creaky	sonorant

To address these issues, the SplitsTree software was employed to generate a phylogram in order to visualise the internal relations among the eight phrase-initial tones. Their multidimensional properties were firstly transformed into a multiple sequence alignment to be computed in the SplitsTree software. Figure 7-44 shows the calculated result of the phonological mapping between Zhangzhou phrase-initial tones. The root of the tree represents the set of phrase-initial tones being investigated. The horizontal lines are branches representing the amount of similarity in terms of multidimensional parameters across eight phrase-initial tones. The longer the branch, the lesser the amount of similarity that tones share in phonetics. The vertical lines indicate the relation between tones. The closer the lines, the higher the probability that the tones belong to the same categories, in other words, the higher the probability of tonal contrast neutralisation in this context. Each tone in this figure consists of a bundle of features corresponding to the information presented in Table 7-3. Additionally, three notable aspects deserve further discussion.

The tones that share similar F0 realisations, for example, phrase-initial tones 7 (I7) and 8 (I8), are not categorically related to each other, since they present differences in two other dimensions—length and syllable coda. This fact further indicates that the single dimension of F0/pitch is not sufficient to characterise Zhangzhou tones.

Two processes of tonal contrast neutralisation occur phrase initially. Phrase-initial tones 1 and 2 (I1 and I2) are clustered together. Similarly, phrase-initial tones 5 and 8 (I5 and I8) are clustered together, without significant differences in terms of the multidimensional framework. The finding further supports the right-dominant system of Zhangzhou tone sandhi in which the number of tonal contrasts is significantly reduced.

The tones that are neutralised in citation are categorically contrastive in the phrase-initial context. Compared to the phylogram in Figure 5-17, tones 2 and 8 (I2 and I8) are not closely related to each other phrase initially because of the differences in the realisations of F0 and low vowel phonation, signifying they are two separate categories. This finding justifies the notion that the single context of citation is not sufficient to construct the number of tonal contrasts in Zhangzhou.

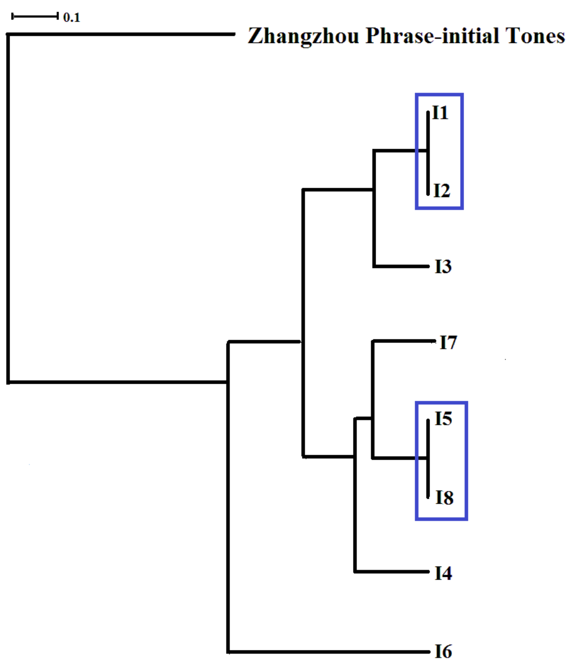


Figure 7-44. Phylogram representation of Zhangzhou phrase-initial tones (I=Phrase-initial).

7.3. Previous Studies

No previous acoustic or statistical studies have explored how the phrase-initial tones are realised and how they interact with the individual phrase-final tones in Zhangzhou. The descriptions of this study address the research gap and advance the understanding of the nature of tonal interactions in Zhangzhou.

7.4. Conclusion

This chapter has provided a detailed exploration into the acoustic realisations of phrase-initial tones across 64 disyllabic combinations, supplemented by statistical testing and articulatory explanation. The results show that the realisations of the phrase-initial tones are not categorically affected by the tones in the phrase-final position but are sensitive to surrounding phonetic environments and present statistically significantly different variants in some tonal combinations, especially in F0 realisations. The tonal realisations in phrase-initial position are also multidimensional, involving systematic interactions among various parameters. The obstruent codas that are not realised in the monosyllabic setting are preserved in this context, significantly shortening the syllable length and giving rise to falsetto voice in the stopped tone 6.

In addition, significant processes of contrast neutralisation occur in the phrase-initial context. The length contrasts are reduced to only two of either extra-short or medium; the number of tonal contrasts is also significantly reduced by two processes of neutralisation. The findings provide a solid and objective foundation for the theoretical explanation of the mapping between the citation and phrase-initial tones in Chapter 9.

Chapter 8: Acoustic Properties of Zhangzhou Phrase-Final Tones

This chapter explores the acoustic properties of Zhangzhou phrase-final tones. It largely examines what phrase-final tones are realised in acoustics with respect to the multidimensional parameters across 64 tonal combinations. It applies the techniques of pairwise *t*-testing and the hierarchical clustering algorithm to determine whether the F0 and duration realisations of individual phrase-final tones may be affected by their preceding tones. It employs the SplitsTree software to assess the relatedness between phrase-final tones from the phonological perspective.

8.1 *Z*-Score Normalised F0 Model

Three major issues are addressed in this section. (1) What are individual phrase-final tones realised across 64 tonal combinations in terms of normalised F0? (2) Are the F0 realisations of phrase-final tones affected by preceding tones? If so, to what extent are they affected, and what conditions the variations? (3) How many normalised F0 levels are contrastive among the phrase-final tones?

The F0 contours presented in this section were obtained by *z*-score normalising the raw F0 values of the related contours from 21 speakers' utterances. Each phrase-final tone also logically would have 28 (=8*7/2) paired F0 differences to be tested at either the 10% or the 100% sampling point, assuming that the normalised values are all independent and identically normally distributed. The Bonferroni corrected alpha of 0.00186 (=0.05/28) was consistently performed to control for the Type I Error for the values at the sampling point under consideration. The hierarchical clustering algorithms was employed to visualise the testing result. The threshold of 1 was consistently utilised to cluster the values and examine how many normalised F0 levels are contrastive from the statistical point of view.

The normalised values and the values of pairwise *t*-testing result for individual phrase-final tones are tabulated in Appendix B (Tables B13-B16). The methodologies of acoustic and statistical processing are described in Chapter 2.

8.1.1. Phrase-final tone 1

Phrase-final tone 1 was indicated as being perceived as having a rising pitch contour with considerable variation (e.g., [34], [23], and [24]) across different combinations in Chapter 6. The auditory impressions are basically justified by the acoustically quantified results. As Figure 8-1 shows, all the normalised F0 contours are rising but differ slightly in both contour onset and offset height. For example, the contour onset value is lower after tone 5 than after tone 1 while the offset value is apparently lower after tones 3, 4, and 6.

The results of the pairwise *t*-test comparison by effect sizes revealed significant differences among the normalised F0 values of phrase-final tone 1 at both the 10% and 100% sampling points, as shown in Figure 8-2. The left panel indicates the normalised F0 values of phrase-final tone 1 across its eight preceding tones at the 10% sampling point (onset), while the right panel indicates the corresponding values at the 100% sampling point (offset). As indicated, the values are clustered into two levels at both sampling points at the threshold of 1, representing statistically significant differences among them. The normalised F0 contour onset is significantly higher after tones 1, 2, and 3 than it is after other tones, while the contour offset is significantly lower after tones 4, 3, and 6.

Therefore, the F0 realisations of phrase-final tone 1 are not quite the same over the phrase-initial tones from the auditory, acoustic, and statistical perspectives. Both onset and offset values are sensitive to surrounding F0 environments. If a preceding tone has the feature of [-falling] (tones 1, 2, and 3), the onset value is statistically significantly higher. If a preceding tone has the feature of either [+mid level] (tones 1 and 2) or [+mid falling] (tones 5, 7, and 8), the offset value tends to be significantly higher.

8.1.2. Phrase-final tone 2

In Chapter 6, the phrase-final tone 2 was described as having a falling contour with a low level plateau varying slightly in the pitch onset ([211] or [311]). The acoustically normalised results basically justify the auditory description. As shown in Figure 8-3, the contours largely present a falling tendency but have a low level plateau during the second half across different combinations. Apparent variation can be seen in the contour onset. For example, the contours after tones 1 and 3 have numerically higher onsets than the contours after tones 7 and 8.

The results of pairwise *t* testing by effect size also support the auditory observation of pitch onset variations. As indicated in Figure 8-4, the terminal nodes at the bottom in the left panel have been clustered into two levels, indicating that the normalised F0 values of the phrase-final tone 2 are statistically significantly different across the eight phrase-initial tones, with the values being higher after tones 1 and 3 than after any other tones at the 10% sampling point. Nevertheless, the effect is marginal depending on the threshold selection. Conversely, the values are clustered together at the 100% sampling point across the eight combinations as shown in the right panel, signifying that the offset realisations are not affected by the preceding tones.

Therefore, phrase-final tone 2 has statistically significant but marginal differences in the realisations of F0 onset. Nevertheless, the conditioning factor for the onset variations tends to be tonally relevant. For example, the F0 onset is significantly higher after tone 1 than after tone 2, but the two phrase-initial tones have very similar realisation in both F0 height and contour. Similarly, the onset is significantly higher after tone 3 than tone 6 while tones 3 and 6 share a very similar F0 offset phrase initially.

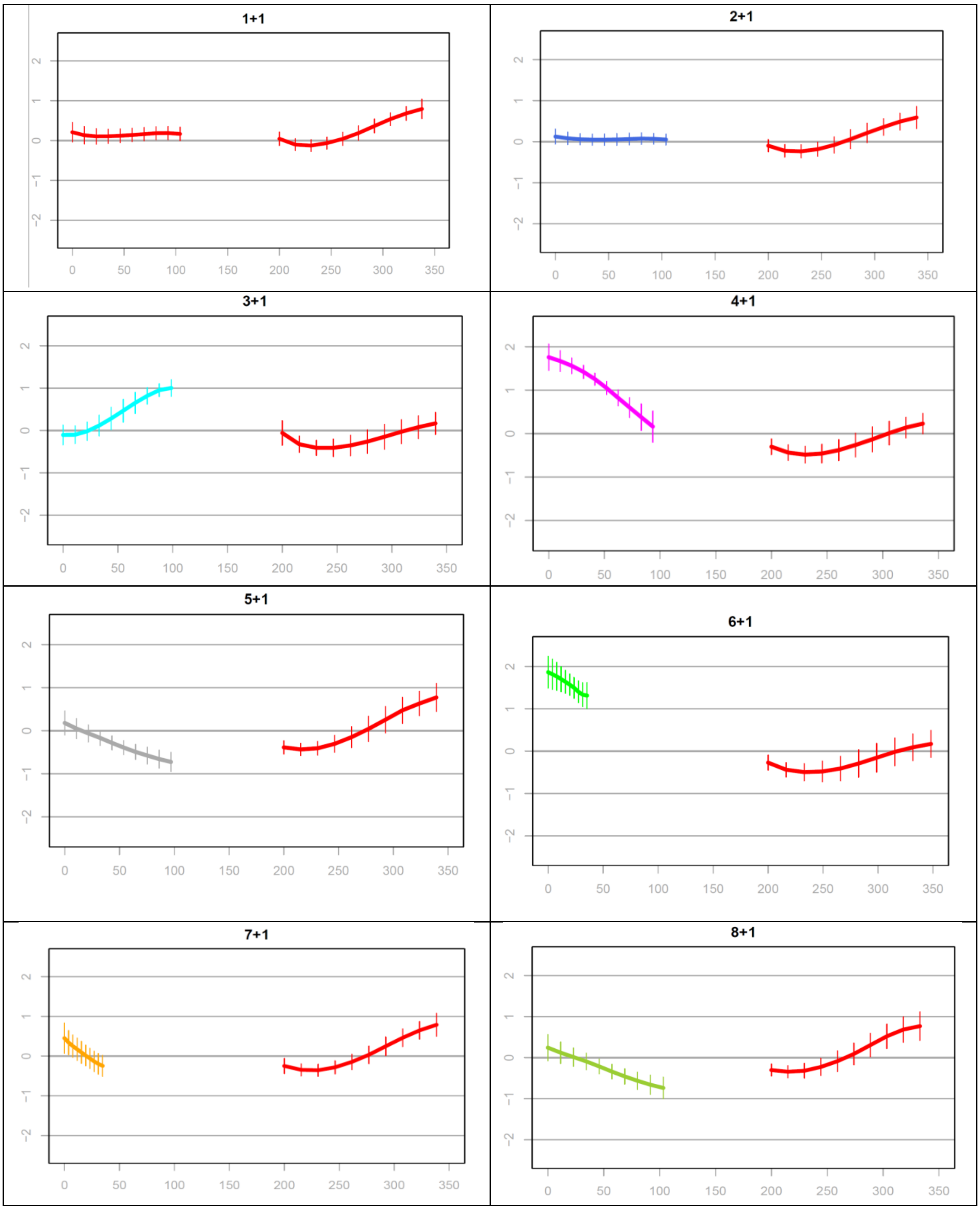


Figure 8-1. ***Z*-score normalised F0 shapes of phrase-final tone 1 across phrase-initial tones from 21 speakers.**

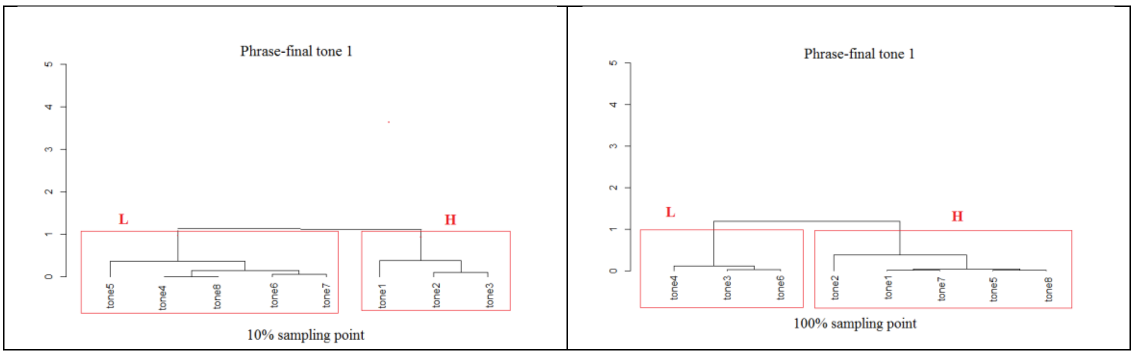


Figure 8-2. Clustering of normalised F0 levels of phrase-final tone 1 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

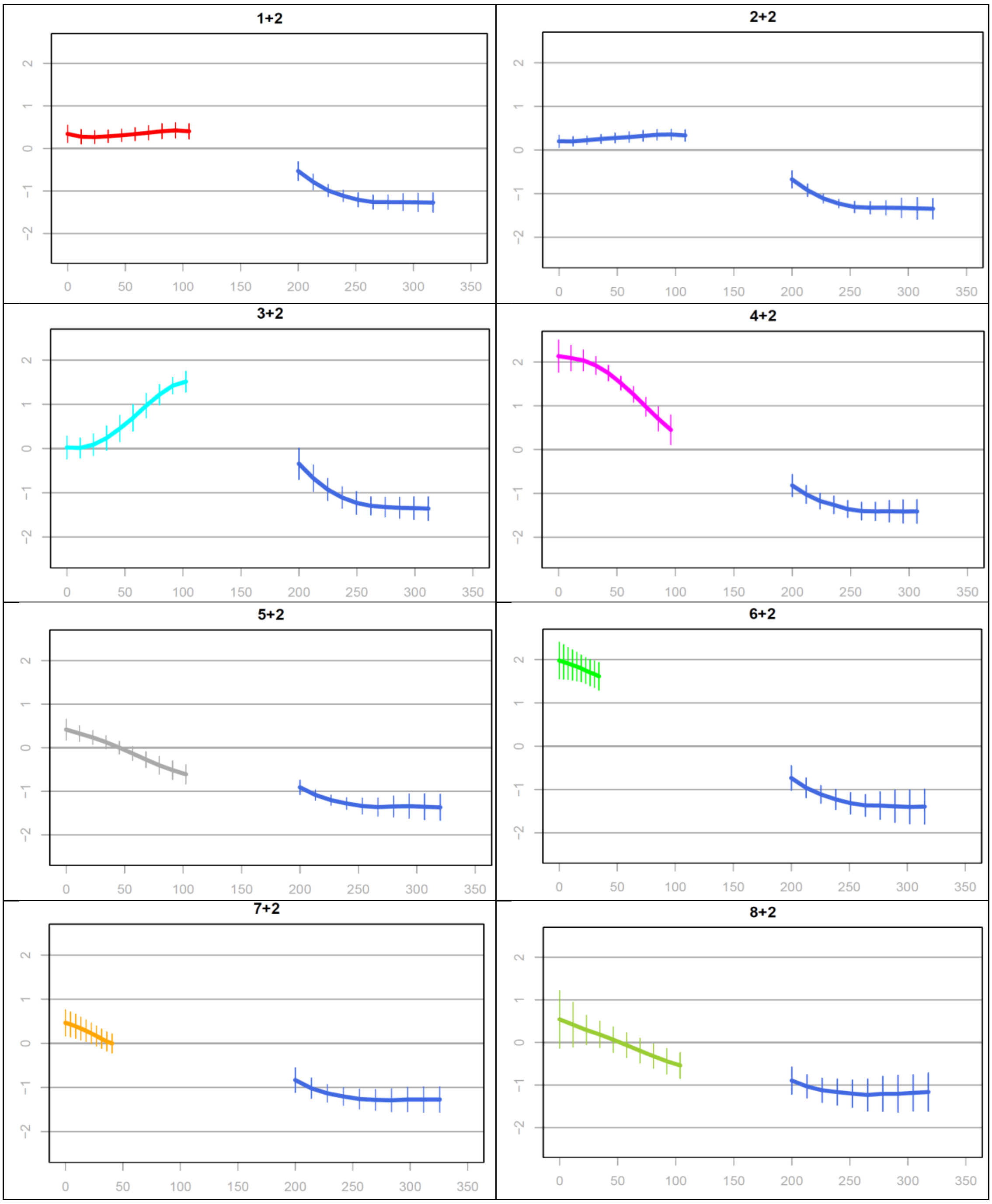


Figure 8-3. *Z*-score normalised F0 shapes of phrase-final tone 2 across phrase-initial tones from 21 speakers.

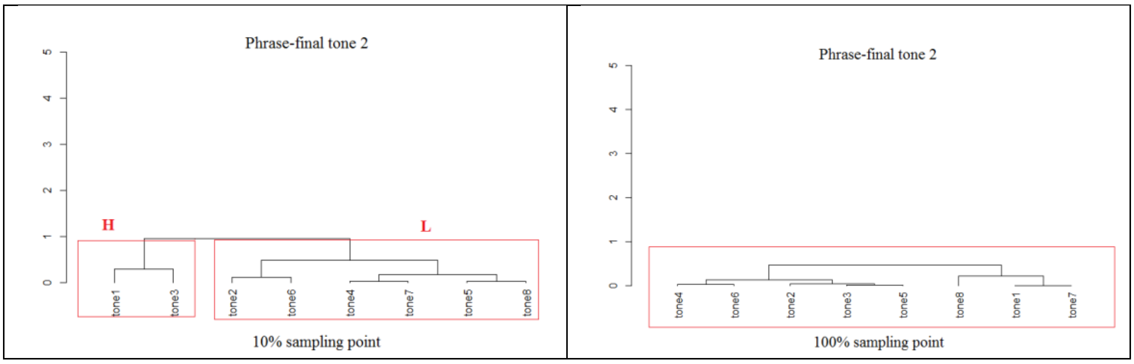


Figure 8-4. Clustering of normalised F0 levels of phrase-final tone 2 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

8.1.3. Phrase-final tone 3

The pitch of phrase-final tone 3 was described in Chapter 6 as a high-falling contour [52] across eight combinations. The acoustically normalised results support the auditory impression. As shown in Figure 8-5, the contours of this tone consistently show a high-falling tendency across different patterns but with slight differences in the F0 onset values. For example, the onset appears lower after tone 4 than after tone 3.

Statistically, the pairwise *t*-test comparisons revealed no significant difference among the normalised F0 values of the phrase-final tone 3 as a function of the phrase-initial tones at both the 10% and 100% sampling points. As shown in Figure 8-6 using the hierarchical clustering algorithms, the eight terminal nodes, representing the normalised F0 values of the phrase-final tone 3 across their preceding tones, are clustered into one class at both the 10% (left) and 100% (right) sampling points. This clustering result indicates that the F0 realisations of phrase-final tone 3 are statistically undifferentiated across different phrase-initial tones.

8.1.4. Phrase-final tone 4

Phrase-final tone 4 was observed as having two pitch variants in Chapter 6. It is realised as a mid-high falling contour [41] in most disyllabic combinations but as a high falling contour [51] after tones 3 and 6, which have high pitch offsets. The results from the acoustic quantification and normalisation justify the auditory observation. As shown in Figure 8-7, the contours of this tone largely descend from slightly above the midpoint, but after tones 3 and 6, the F0 contours apparently have higher onsets.

Statistically, the pairwise t test comparisons by effect size also support the auditory impression. As indicated in Figure 8-8, the terminal nodes at the bottom in the left panel have been clustered into two levels, indicating that the normalised F0 values of the phrase-final tone 4 are statistically significantly different across the eight phrase-initial tones at the 10% sampling point at the threshold of 1, with the values being significantly higher after tones 3 and 6 than after any other tones. Conversely, the values are clustered together as one group at the 100% sampling point across the eight combinations as shown in the right panel, signifying that the offset realisations of the phrase-final tone 4 are not affected by the preceding tones.

Therefore, the F0 onset realisation of phrase-final tone 4 is also phonetically sensitive to the F0 offset of preceding tones. If the phrase-initial tone has the feature of [+high offset] (tones 3 and 6), its onset will be statistically significantly higher.

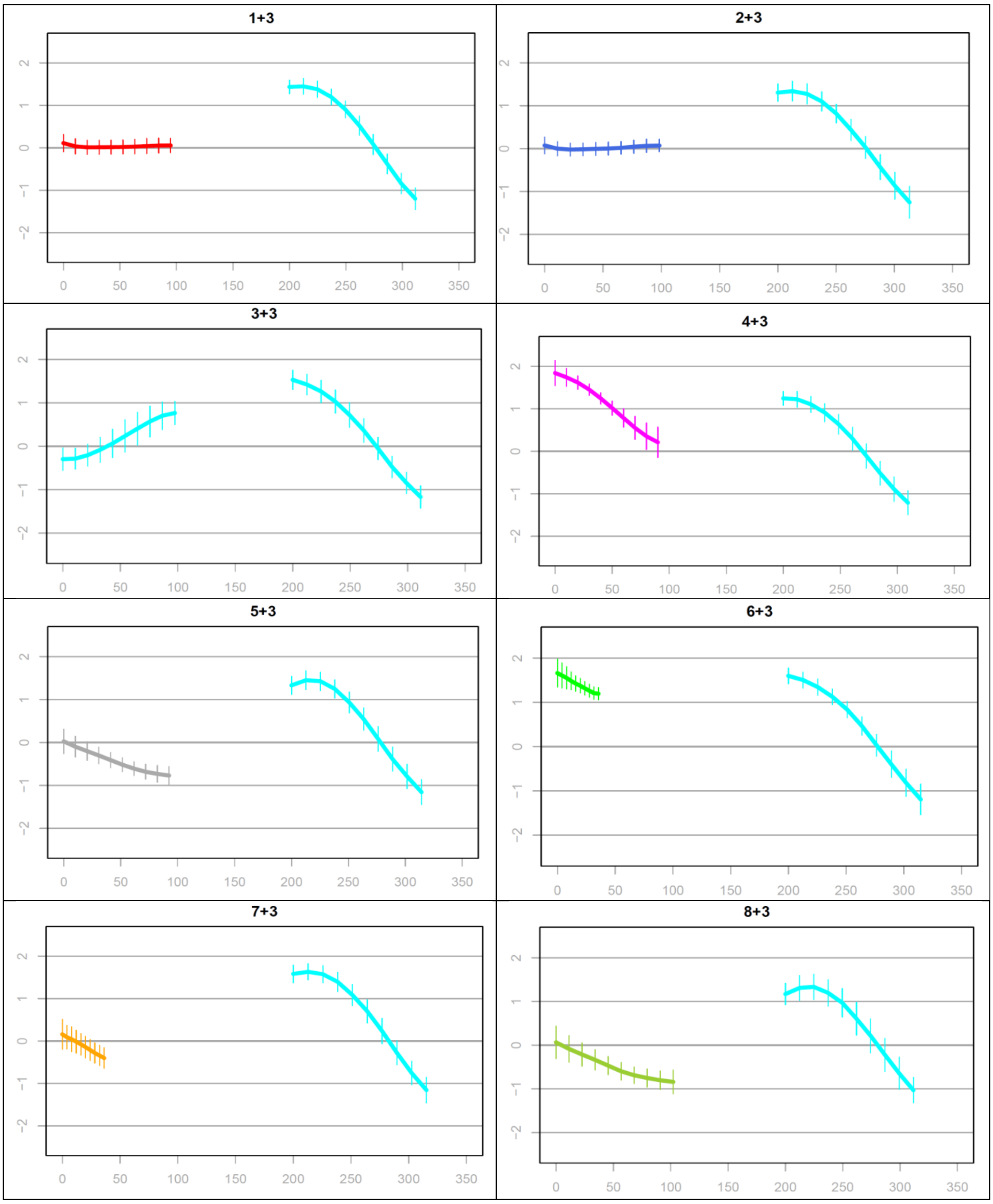


Figure 8-5. *Z*-score normalised F0 shapes of phrase-final tone 3 across phrase-initial tones from 21 speakers.

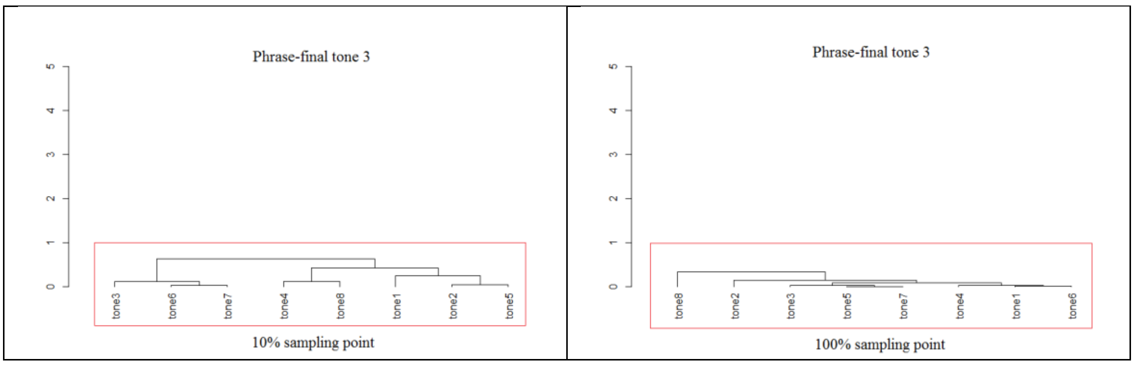


Figure 8-6. Clustering of normalised F0 levels of phrase-final tone 3 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

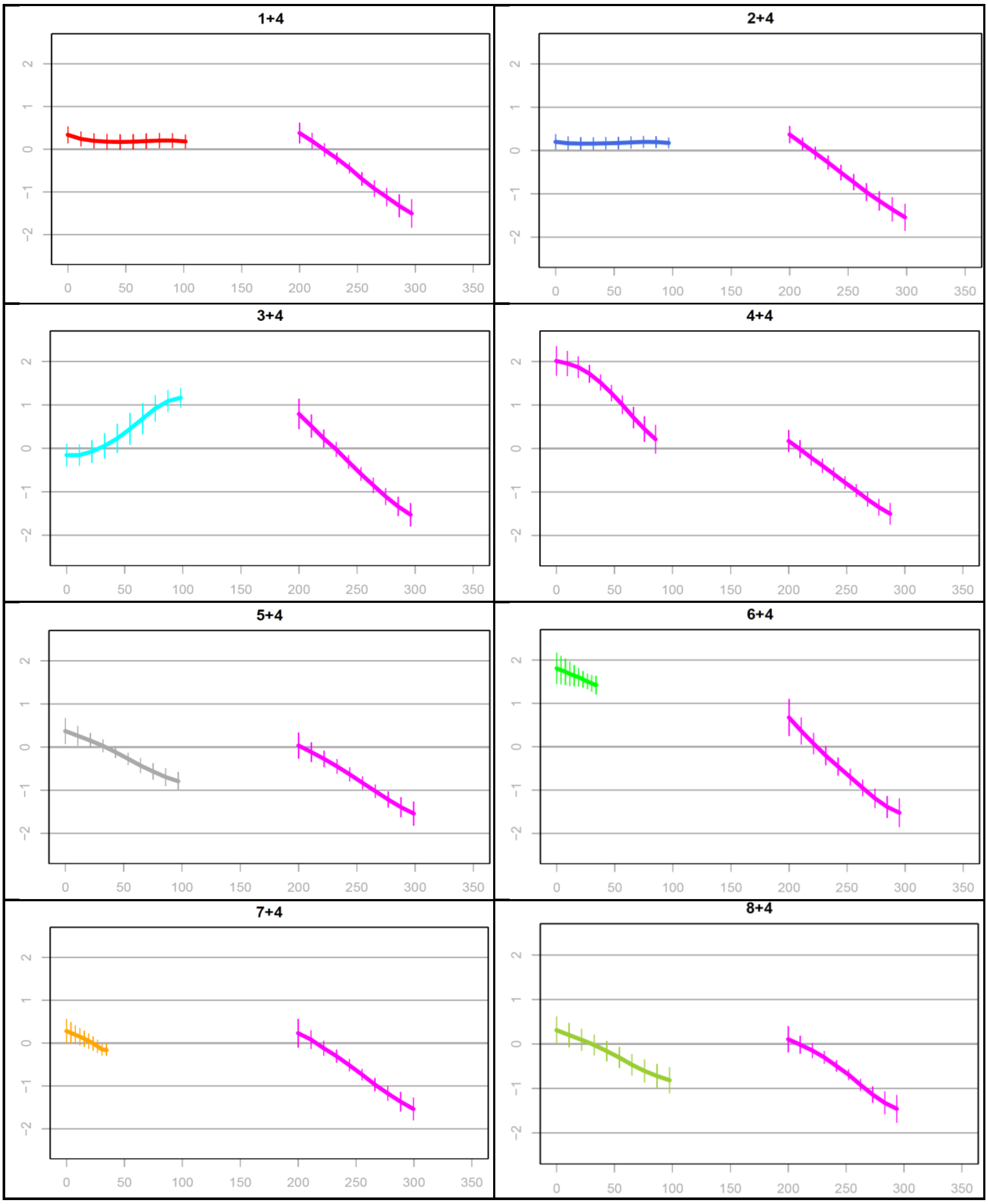


Figure 8-7. *Z*-score normalised F0 shapes of phrase-final tone 4 across phrase-initial tones from 21 speakers.

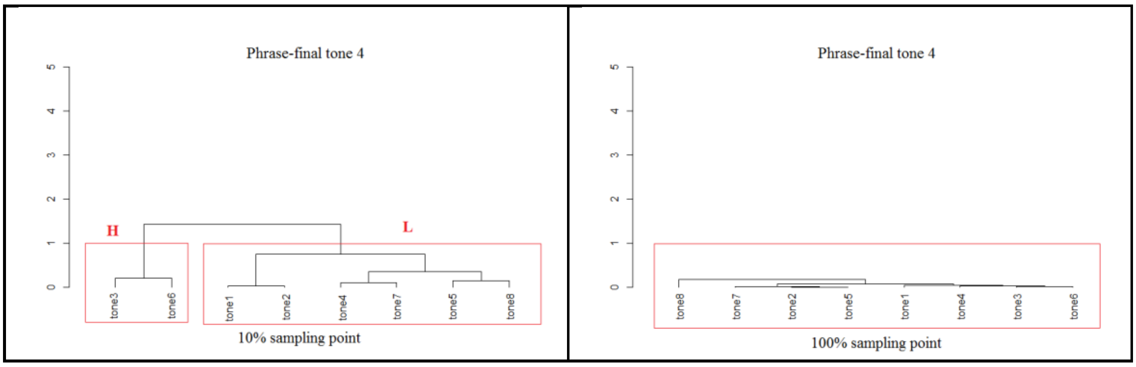


Figure 8-8. Clustering of normalised F0 levels of phrase-final tone 4 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

8.1.5. Phrase-final tone 5

Phrase-final tone 5 was perceived as a mid-level [33] across all tonal combinations in Chapter 6, and identifying whether its F0 ralisations may be influenced by surrounding tones tends to be perceptually difficult. The acoustic quantification and normalisation results show a level F0 contour around the midpoint but with small variation in the contour distribution. For example, the contours have slightly higher levels after tones 1, 2, and 5 than after other tones (Figure 8-9).

The pairwise t test comparisons by effect size also support the auditory impression. As indicated in Figure 8-10, the terminal nodes in the left panel are clustered into two levels at the threshold of 1, indicating that the normalised F0 values of the phrase-final tone 5 are statistically significantly different across the eight phrase-initial tones at the 10% sampling point, with the values being significantly lower after tones 4, 5, and 8 than after any other tones. Conversely, the terminal nodes are clustered together as one group at the 100% sampling point across the eight combinations as shown in the right panel, demonstrating that the offset realisations of the phrase-final tone 5 are not affected by the preceding tones.

Therefore, the F0 realisations of phrase-final tone 5 are not quite the same over the phrase-initial tones from the acoustic, and statistical perspectives. The conditioning factor for the onset variations of phrase-final tone 5 can be generalised as [- low offset] of phrase-initial tones. If the preceding tone (tones 4, 5, and 8) has an offset lower than the midpoint, the phrase-final tone 5 should have a statistically significantly lower onset, and vice versa.

8.1.6. Phrase-final tone 6

Phrase-final tone 6 was described in Chapter 6 as having a falling contour with variation in pitch onset height ([41] and [51]). This auditory impression is basically justified by the acoustically quantified and normalised results. As Figure 8-11 shows, all contours of this tone are falling but vary gradually in the onset levels. For example, the contour after tone 4 apparently has a lower onset than the contour after tone 3.

The statistical testing results of pairwise *t*-test comparisons also support the auditory discription. As Figure 8-12 shows, the terminal nodes in the left panel, representing the normalised F0 values of phrase-final tone 6 across eight phrase-initial tones, have been clustered into two levels at the 10% sampling points. The clustering result indicates significant differences with the onset values significantly lower after tones 4, 5, 7, and 8 than after other tones. Conversely, there is no statistically significant difference among the values at the 100% sampling point.

Therefore, the F0 realisations of phrase-final tone 6 are also acoustically sensitive to the F0 offset of preceding tones. The conditioning factor for the variations can be generalised as [- low offset]. If the preceding tone (tones 1, 2, 3, and 6) has an offset at or above the midpoint, tone 6 should have a statistically significantly higher onset phrase finally, and vice versa.

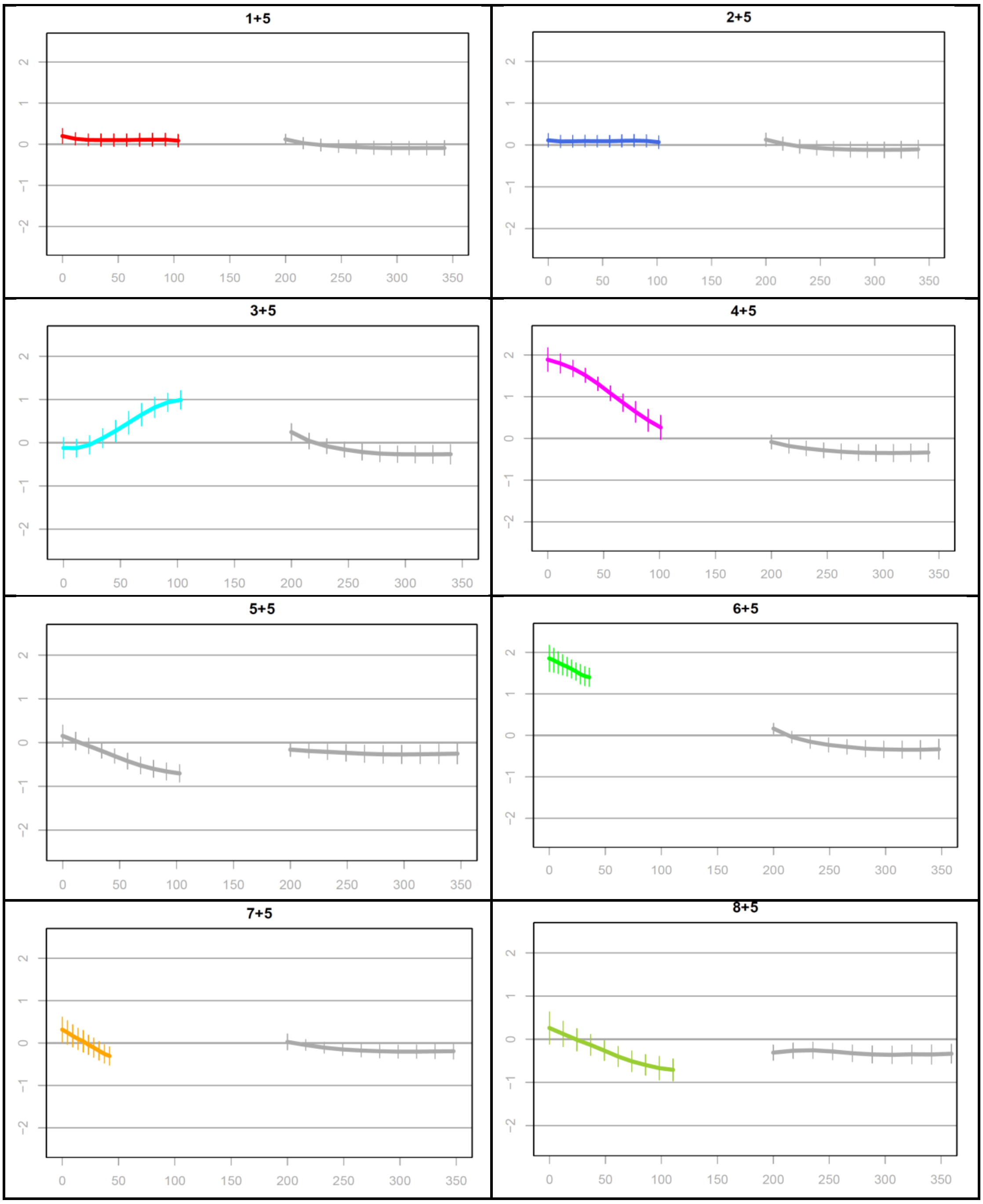


Figure 8-9. *Z*-score normalised F0 shapes of phrase-final tone 5 across phrase-initial tones from 21 speakers.

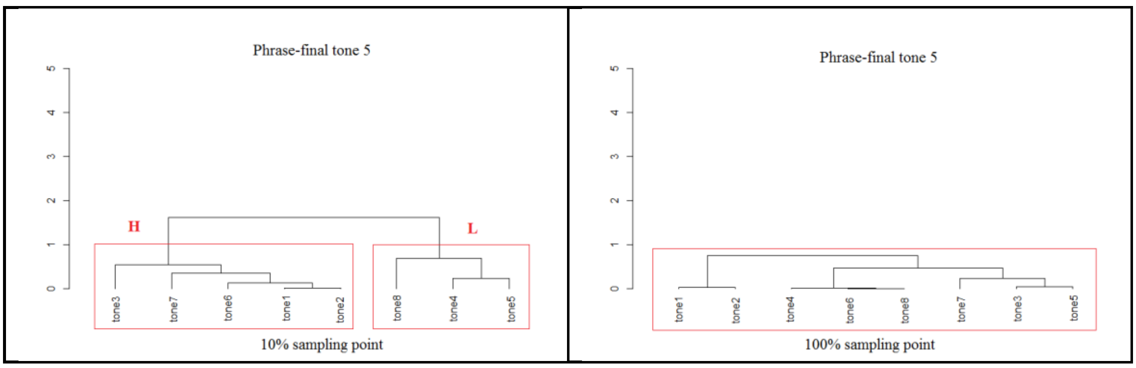


Figure 8-10. Clustering of normalised F0 levels of phrase-final tone 5 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

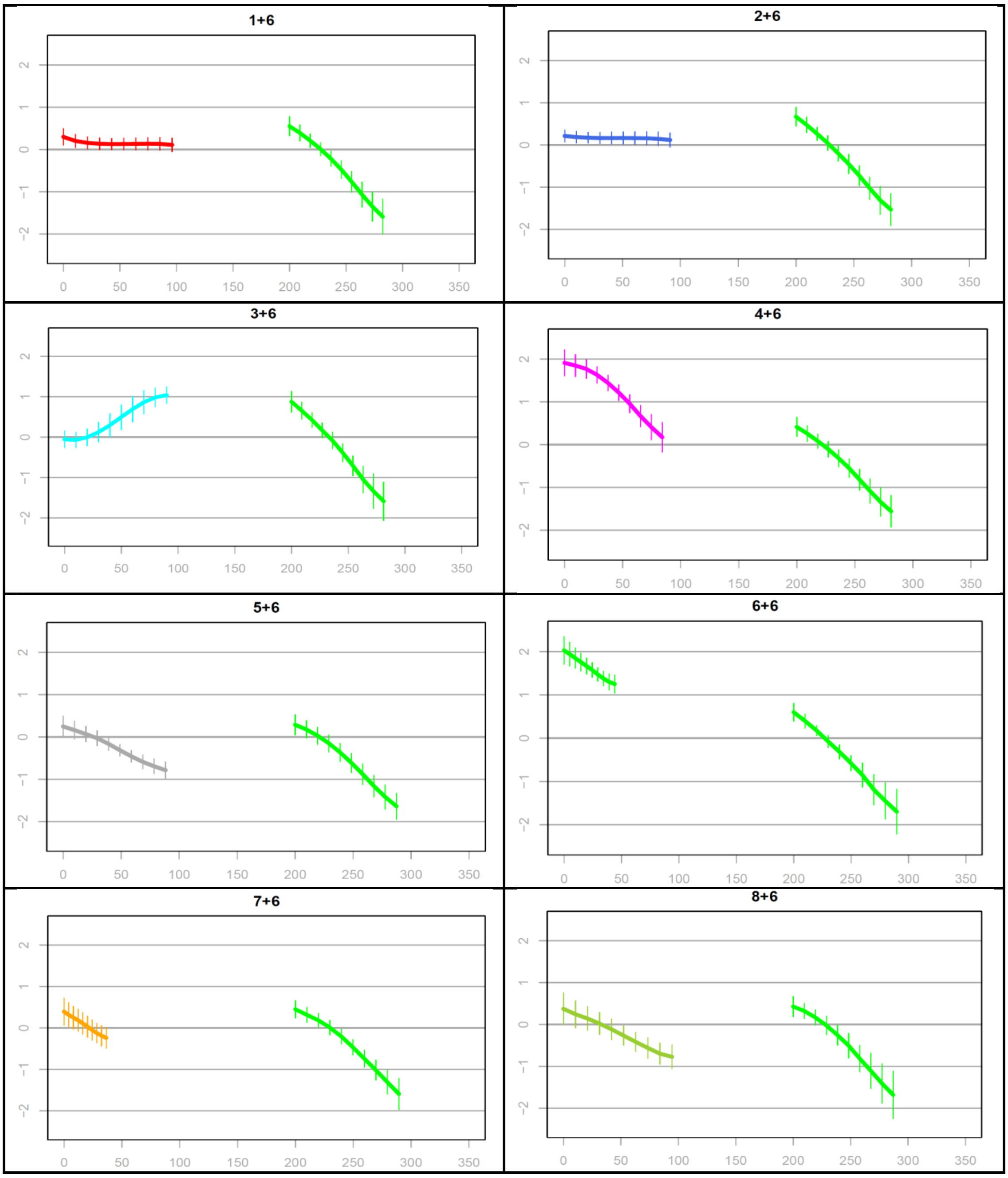


Figure 8-11. ***Z*-score normalised F0 shapes of phrase-final tone 6 across phrase-initial tones from 21 speakers.**

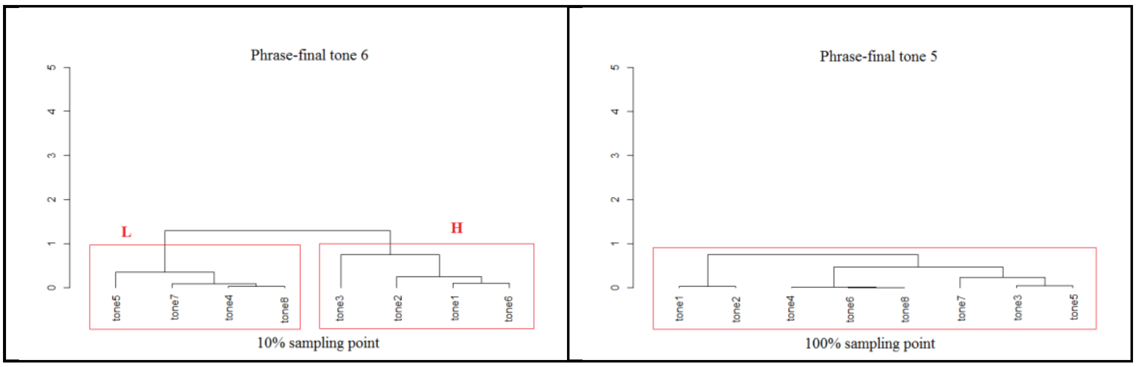


Figure 8-12. Clustering of normalised F0 levels of phrase-final tone 6 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

8.1.7. Phrase-final tone 7

The pitch of phrase-final tone 7 was perceived as having two variants of [311] and [211] as a consequence of the influence of the preceding pitch offset. The acoustically quantified results justify this auditory impression. As shown in Figure 8-13, the contours of this tone all present a falling tendency but have considerable variation in the normalised F0 onset values. For example, the contour has a higher onset after tone 3 than after tone 4.

The statistical testing results also justify the acoustic and auditory observations. As shown in Figure 8-14, the eight phrase-initial tones in the left panel, representing the eight normalised F0 values of phrase-final tone 7 as a function of its preceding tones, are clustered into two levels at the 10% sampling point, with the values after tones 1, 2, 3 and 6 being significantly higher than the values in other group. On the contrary, the normalised F0 values at the 100% sampling point are statistically undifferentiated across eight combinations at the threshold of 1.

Therefore, the F0 realisations of phrase-final tone 7 are not exactly the same as a function of the phrase-initial tones; instead, they are phonetically sensitive to the F0 offset of preceding tones. Similar to the realisations of phrase-final tone 6, the conditioning factor can also be generalised as [-low offset]. If the F0 offset of preceding tone (tones 1, 2, 3, and 6) is at or above the midpoint, the onset of this phrase-final tone 7 should be statistically significantly higher, and vice versa.

8.1.8. Phrase-final tone 8

The pitch of phrase-final tone 8 was indicated in Chapter 6 as a falling contour with a low level plateau in the second half but having variations in the pitch onset ([311] or [211]) because of regressive assimilation to the pitch offset of preceding tones. This auditory observation is acoustically justified. As Figure 8-15 indicates, the contours of this tone largely have a falling tendency in the first half but tend to be level in the second half. In addition, apparent variations occur in the contour onset. For example, the onset is higher after tone 3 than after tone 4.

Statistically, the results of pairwise t-test comparison also support the auditory observation. As Figure 8-16 shows, the eight phrase-initial tones are clustered into two levels, indicating that the normalised F0 values of phrase-final tone 8 are statistically distinguished with the values higher after tones 1, 2, 3, and 6 than after other tones at the 10% sampling point; however, the effect is marginal depending on the threshold selected. Conversely, the phrase-initial tones are clustered together as a group at the 100% sampling points, indicating that the normalised F0 offset values of phrase-final tone 8 are statistically undifferentiated across different tonal combinations.

Therefore, the F0 realisations of phrase-final tone 8 are also phonetically sensitive to the F0 contours of the phrase-initial tones. The conditioning environment for its F0 variations is similar to that for phrase-final tones 6 and 7. If the preceding tone has an offset at or above the midpoint, the onset of this tone is supposed to be statistically significantly higher, and vice versa.

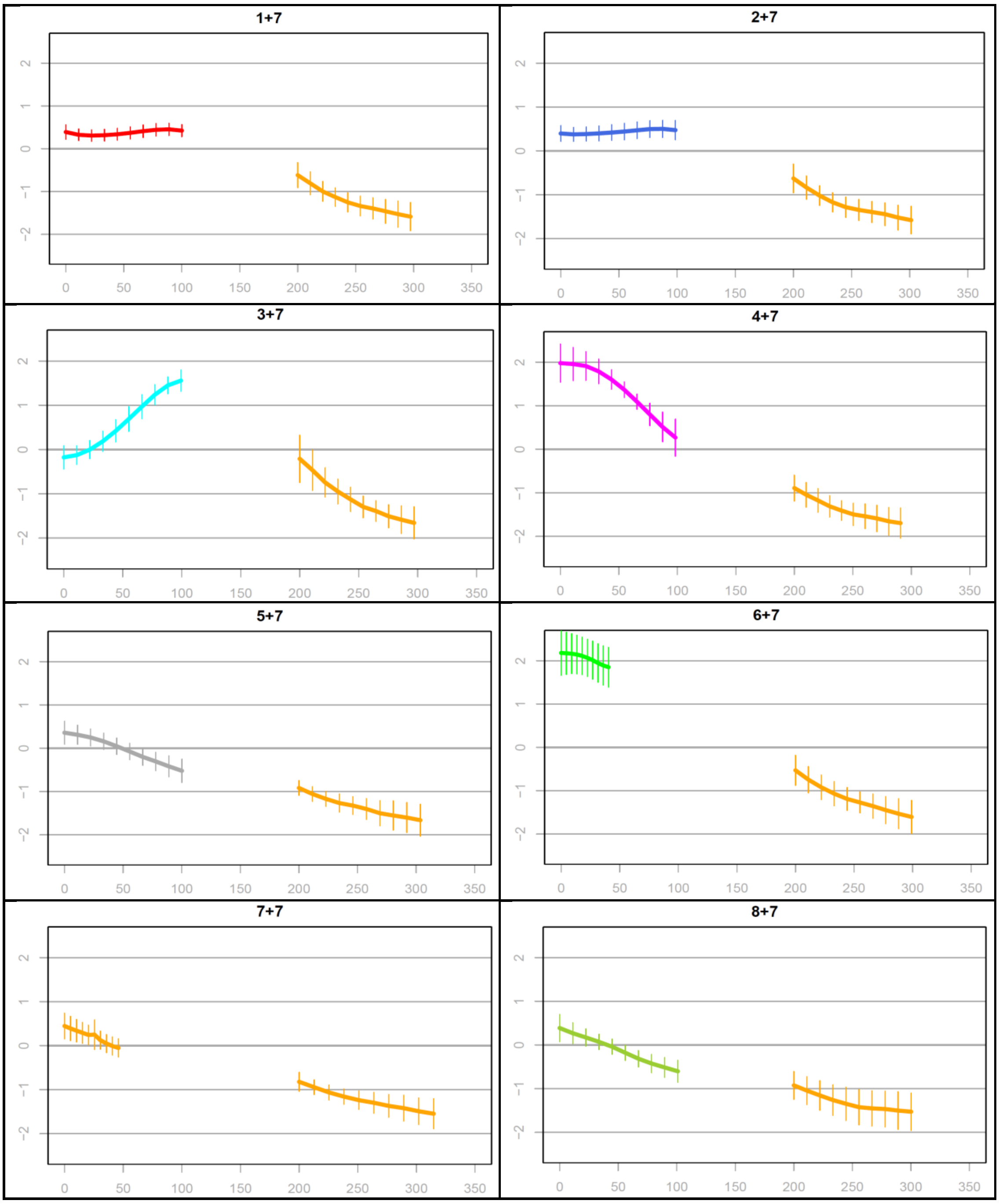


Figure 8-13. *Z*-score normalised F0 shapes of phrase-final tone 7 across phrase-initial tones from 21 speakers.

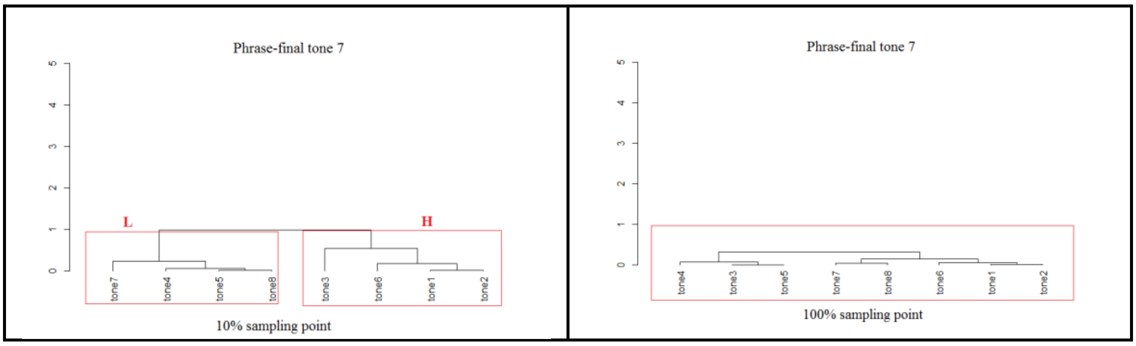


Figure 8-14. Clustering of normalised F0 levels of phrase-final tone 7 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests.

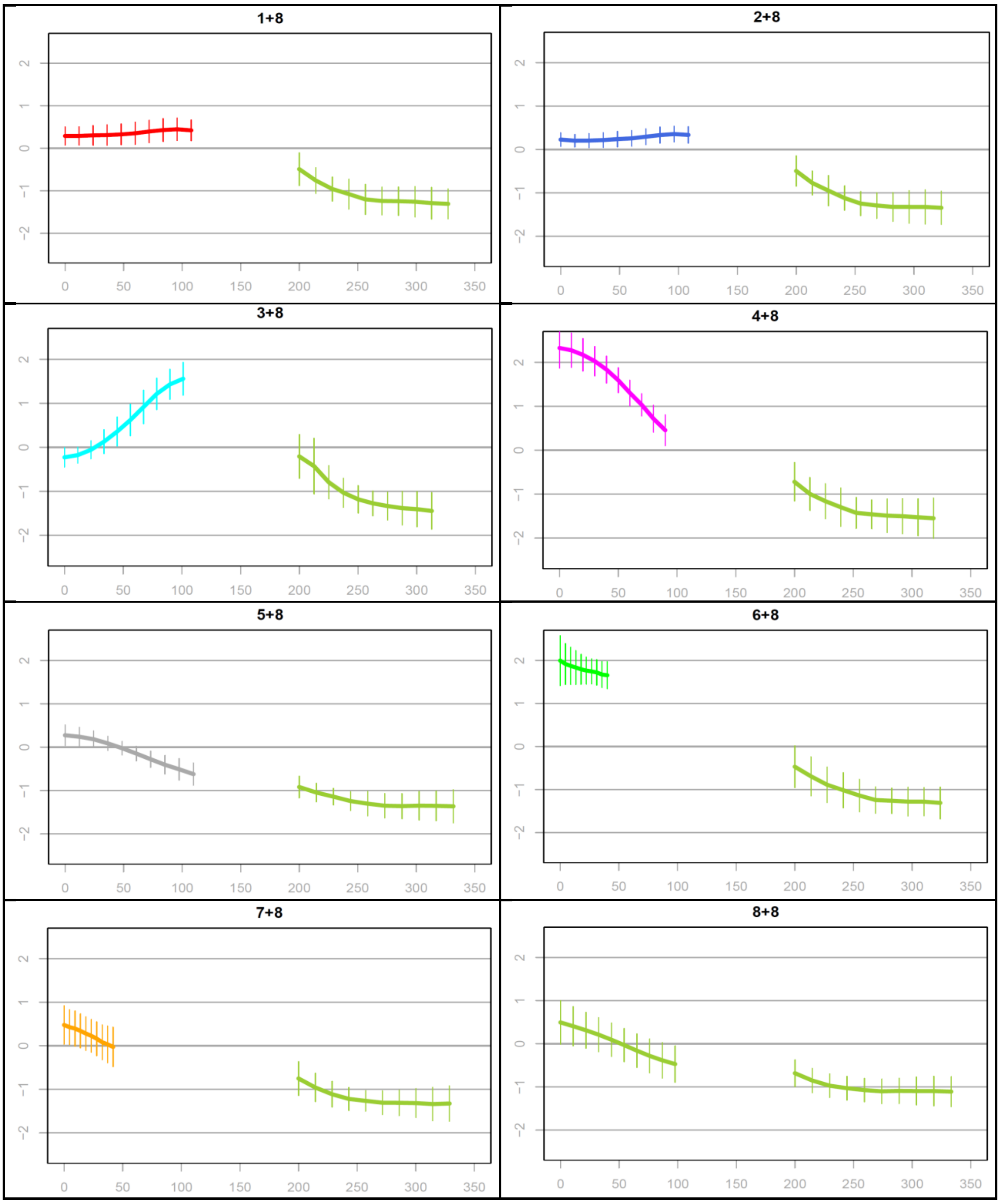


Figure 8-15. *Z*-score normalised F0 shapes of phrase-final tone 8 across phrase-initial tones from 21 speakers.

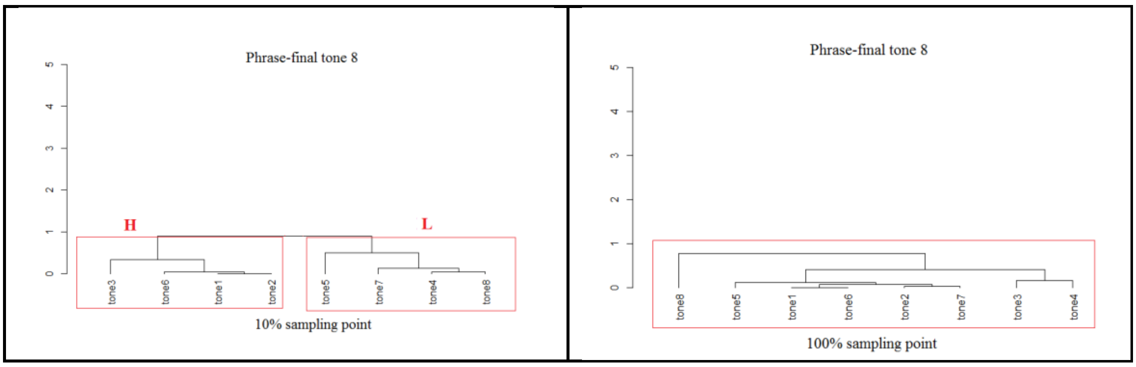


Figure 8-16. Clustering of normalised F0 levels of phrase-final tone 8 at 10% (left) and 100% (right) sampling points based on pairwise *t*-tests from 21 speakers.

8.1.9. Summary

As discussed, the acoustic F0 realisations of individual phrase-final tones tend not to be categorically affected by tones in the phrase-initial position, but their onset values are phonetically sensitive to the offset values of phrase-initial tones while their offset values show less variation. In general, the F0 contours tend to have a higher onset after tones ending in a [-low offset], but have a lower onset after tones having the features [+falling] and [-high offset]. This tendency is statistically justified by the pairwise *t*-test comparison by effect size in most tonal combinations, as summarised in Figure 8-17; all tones except tone 3 have two statistically significantly different variants as a function of their preceding tones. According to the acoustic quantifying and statistical testing results, the linguistically phonetic F0 system of Zhangzhou phrase-final tones is plotted in the Figure 8-18, representing the central tendency of Zhangzhou as an independent system. Nevertheless, several problems occur with respect to the categorisation and representation of different normalised F0 levels.

- The diverse F0 onset levels propose challenges to both statistical testing and the Chao pitch notation system concerning how many F0 levels are categorically distinguished. Seven out of eight tones have variants in the onset realisations, signifying that logically there may exist at least 14 phonetic levels to be tested by conducting exhaustive comparisons on at least 91 paired differences with the Bonferroni corrected alpha of 0.00054945 to achieve significance. Nevertheless, the clustering algorithm turns out to be unsatisfactory to cluster the levels into reasonable categories. The Chao system appears insufficient to characterise different F0 levels.
- The F0 range of phrase-final tones tends to be lowered than that of the phrase-initial tones. The highest level appears lower than the highest level found among the phrase-initial tones, and the lowest level also shows lower than the lowest level at the phrase-initial position. Such a lowering F0 range is considered being caused by the position of utterance-final. In other words, the utterance-final position has a lowering effect on the F0 range realisation.
- The F0 onsets of most phrase-final tones tend to be higher after a non-low F0 offset but lower than after a F0 contour having a falling tendency and ending in a non-high offset. It raises another problem which onset form, the raised F0 or the lowed F0 onset, should be treated as the unmarked form, and which one should be considered as derived from.

Incorporating the auditory observation, acoustic distribution, and position-induced lowering effect, it is appropriate to consider the lowered F0 onsets as the unmarked forms and the raised F0 onsets as the marked forms, whose distributions are conditioned by the F0 offset levels of preceding tones. Thus, the F0 system of the phrase-final tones can be represented using numerical notation, with *5* indicating the highest F0 level and *1* the lowest, as shown in Table 8-1. In addition, as indicated in this table, the single dimension of F0 appears insufficient to distinguish the contours among phrase-final tones 2, 7, and 8, and the contours among phrase-final tones 4 and 6. Taking other phonetic parameters into consideration is thus imperative to understand the nature of phrase-final tonal contrast in Zhangzhou.

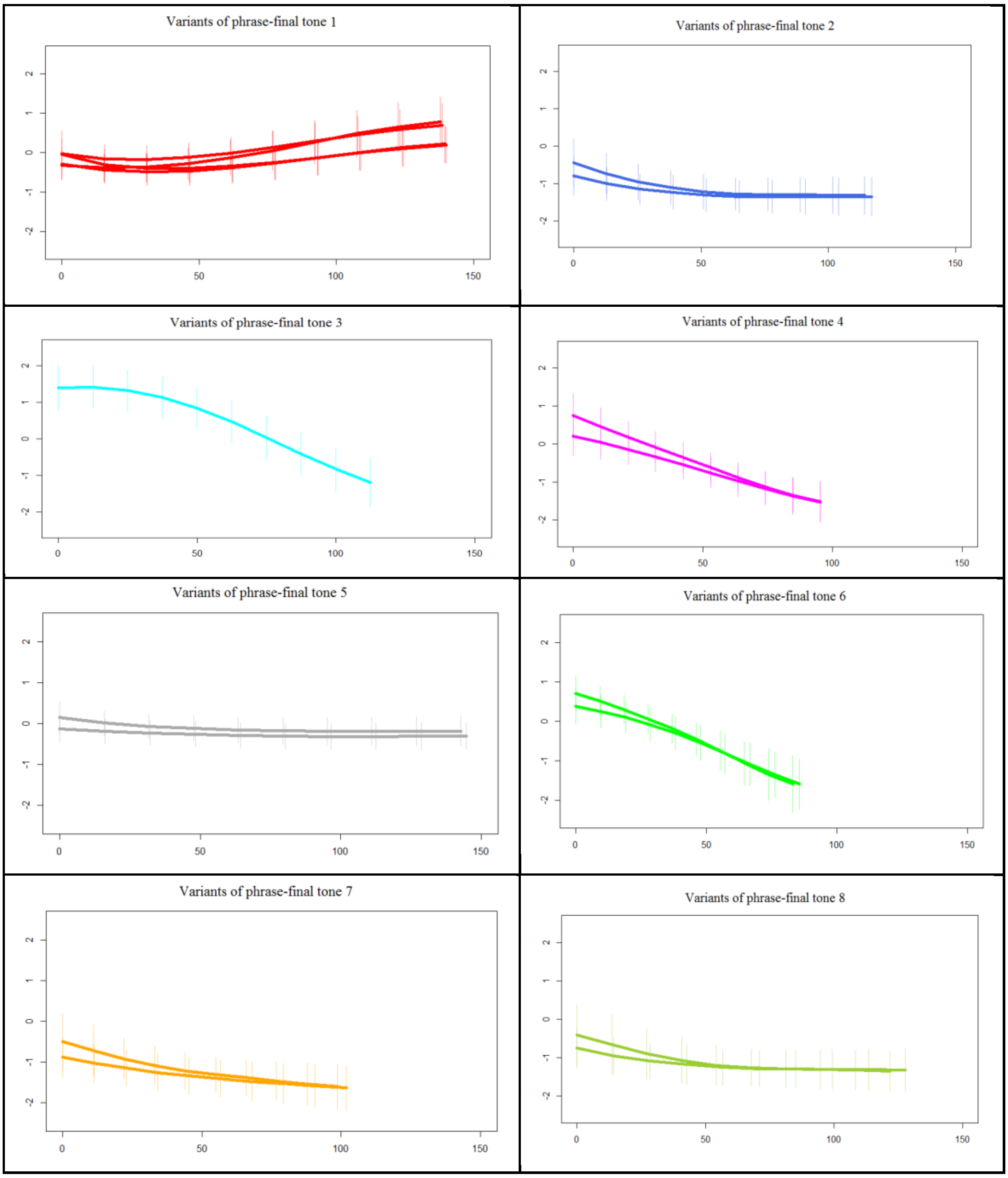


Figure 8-17. Summary of the F0 realisations of individual phrase-final tones in Zhangzhou from 21 speakers.

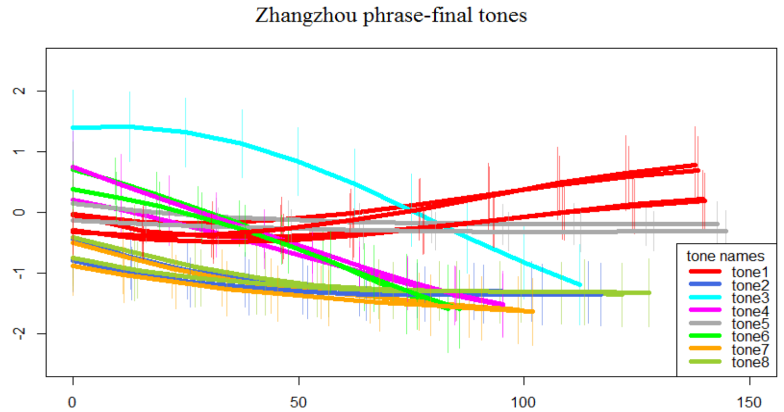


Figure 8-18. *Z*-score normalised F0 system of Zhangzhou phrase-final tones.

Table 8-1. The numerical F0 representation of Zhangzhou phrase-final tones

Tone	F0 contour description	Unmark form	Marked form
1	mid rising	[34]	[35]
2	low falling	[211]	[311]
3	high falling	[52]	[52]
4	mid-high-falling	[41]	[51]
5	mid level	[33]	[43]
6	stopped mid-high falling	[41]	[51]
7	stopped low falling	[211]	[311]
8	low falling	[211]	[311]

8.2. Interactions Among Acoustic Parameters

This section mainly focuses on how various segmental and suprasegmental parameters interact to make the distinctions in acoustics among phrase-final tones. The techniques of pairwise *t* testing and the hierarchical clustering algorithm were applied to determine whether the duration realisations of individual phrase-final tones may be affected by their preceding tones. The SplitsTree software was used to assess the relatedness between phrase-final tones from the phonological perspective.

8.2.1. Duration

This subsection addresses three issues. (1) What are the length realisations of phrase-final tones across 64 disyllabic combinations? (2) Are their length realisations affected by phrase-initial tones? If so, to what extent are they affected, and what conditions the variations? (3) How many length levels are statistically contrastive among the phrase-final tones?

The duration contours presented in this section have been normalised and expressed as a percentage of the average duration of all phrase-final tones from 21 speakers' utterances. Each phrase-final tone also logically would have 28 (=8*7/2) paired length differences to be tested across eight phrase-initial tones, assuming that the normalised values are all independently and identically normally distributed. The Bonferroni corrected alpha of 0.00186 (=0.05/28) was consistently performed to control for the Type I Error. The threshold of 1 was consistently employed to cluster the values and examine how many length levels are contrastive among phrase-final tones from the statistical perspective.

The normalised duration values and the values of pairwise *t*-testing result for individual phrase-final tones are tabulated in Appendix B (Tables B17-B19). The methodologies of acoustic and statistical processing are described in Chapter 2.

8.2.1.1. Phrase-final tone 1

Phrase-final tone 1 was observed as having very similar length realisations across phrase-initial tones in Chapter 6. Acoustically, a very small amount of variation can be seen in the normalised values across all combinations. For example, as plotted in Figure 8-19, the duration is slightly longer after tone 6 than after tone 7. Statistically, the clustering results based on the pairwise *t* tests revealed no significant differences in the normalised duration values of this tone across eight phrase-initial tones. As shown in Figure 8-20, the terminal nodes at the bottom, representing the normalised duration values of phrase-final tone 1 across the eight phrase-initial tones, are clustered together at the threshold of 1, indicating no statistically significant difference among them. In other words, the duration realisations of phrase-final tone 1 can be considered statistically undifferentiated across its preceding tones. Therefore, the auditory observation is not supported by the statistical testing result, and the acoustic differences are not statistically significant.

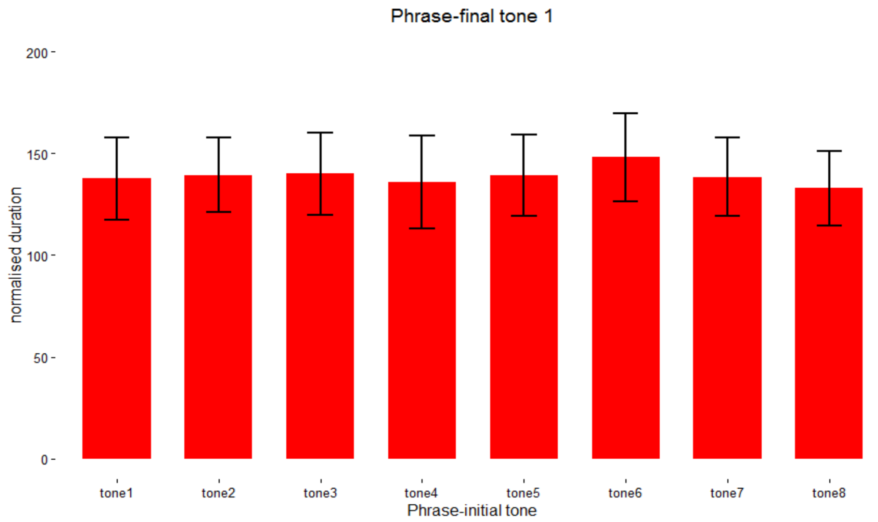


Figure 8-19. Normalised duration of phrase-final tone 1 across phrase-initial tones from 21 speakers.

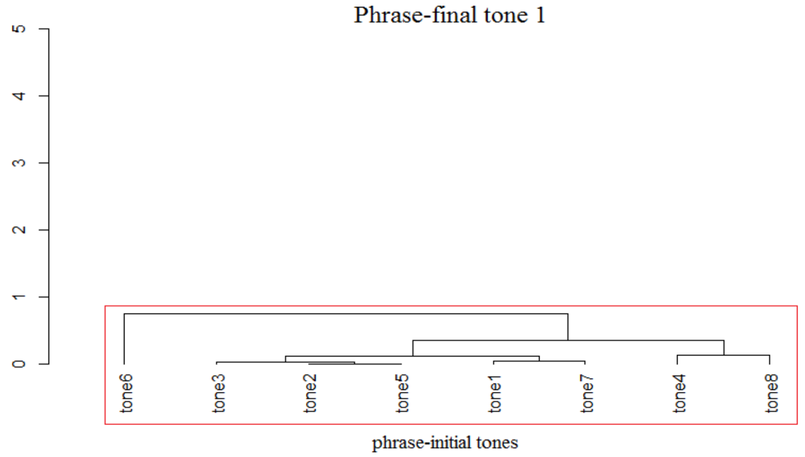


Figure 8-20. Clustering of normalised length levels of phrase-final tone 1 across eight phrase-initial tones based on pairwise *t*-tests.

8.2.1.2. Phrase-final tone 2

Perceptually, phrase-final tone 2 has very similar length realisations across all tonal combinations. Acoustically, the duration values vary somewhat across different combinations. For example, as shown in Figure 8-21, the F0 contour after tone 4 is numerically shorter than the contour after tone 7. Nevertheless, the clustering results based on pairwise *t* tests revealed no statistically significant differences among the normalised duration values of this phrase-final tone 2, since the terminal nodes at the bottom, representing the eight tonal combinations, are all clustered together at the threshold of 1 (Figure 8-22). In other words, the durational realisations of this phrase-final tone 2 are not affected by its preceding tones phonologically.

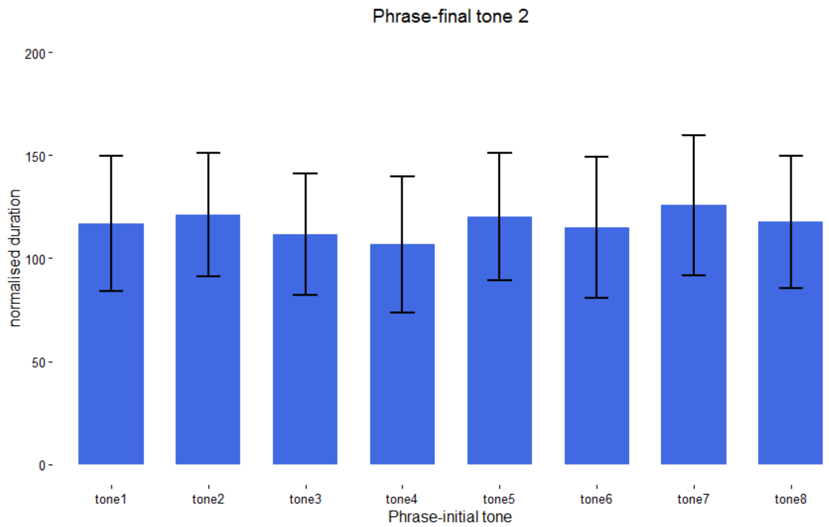


Figure 8-21. Normalised duration of phrase-final tone 2 across phrase-initial tones from 21 speakers.

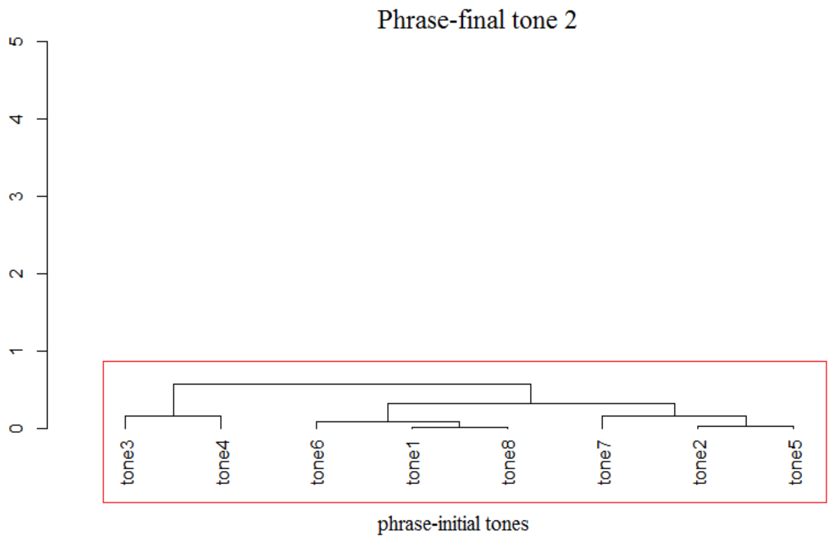


Figure 8-22. Clustering of normalised length levels of phrase-final tone 2 across eight phrase-initial tones based on pairwise *t*-tests.

8.2.1.3. Phrase-final tone 3

Phrase-final tone 3 was indicated in Chapter 6 as having similar length realisations across all disyllabic combinations. The acoustically quantified results justify the auditory impression, as shown in Figure 8-23. The contours tend to have very similar duration without significant variations in the normalised values.

The statistical testing results of pairwise *t*-test comparison also support the auditory assumption. As indicated in Figure 8-24, all the terminal nodes at the bottom, representing the normalised duration values of this phrase-final tone 3 across the eight phrase-initial tones, are clustered together at the threshold of 1. Therefore, the duration realisations of phrase-final tone 3 can be considered statistically undifferentiated across its preceding tones.

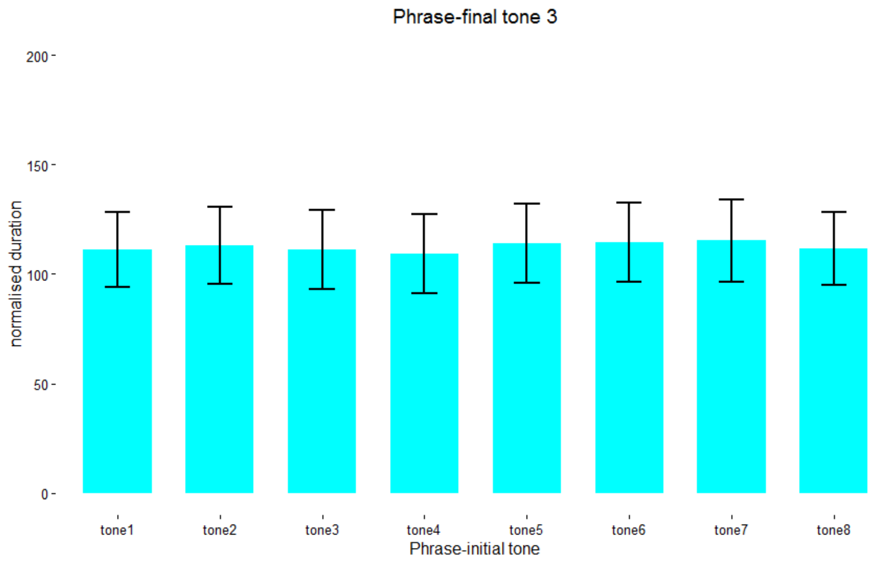


Figure 8-23. Normalised duration of phrase-final tone 3 across phrase-initial tones from 21 speakers

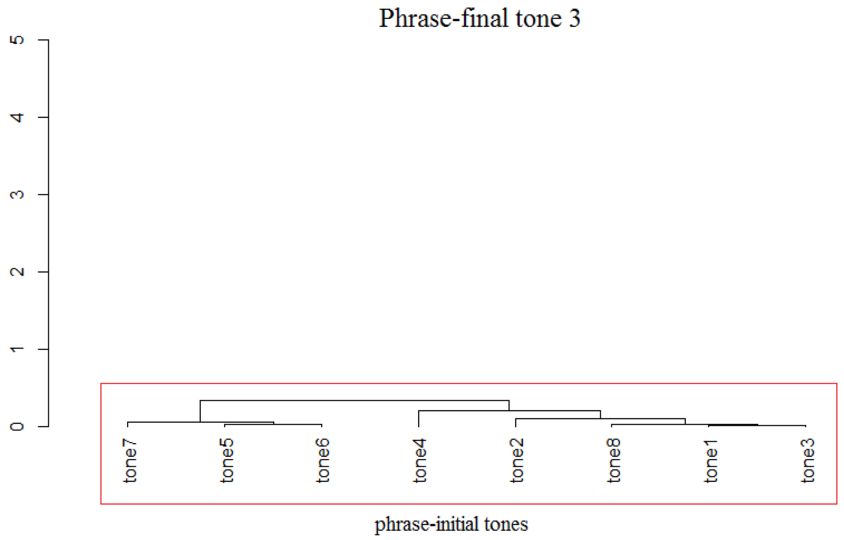


Figure 8-24. Clustering of normalised length levels of phrase-final tone 3 across eight phrase-initial tones based on pairwise *t*-tests.

8.2.1.4. Phrase-final tone 4

Phrase-final tone 4 was described in Chapter 6 as having very similar realisations in terms of tonal duration across eight disyllabic combinations. Nevertheless, the acoustically quantified results indicate a slight degree of variation in the normalised duration values. For example, as shown in Figure 8-25, the normalised duration value is shorter after tone 4 than after tone 5.

Statistically, no significant difference is revealed among the eight disyllabic combinations on the basis of pairwise t tests by effect size. As shown in Figure 8-26, the values of this phrase-final tone 4 are clustered together at the threshold of 1 across their preceding phrase-initial tones. Therefore, the duration realisations of phrase-final tone 4 can be considered statistically undifferentiated across the eight combinations. In other words, the duration realisations of phrase-final tone 4 are not affected by the preceding tones at the phonological level.

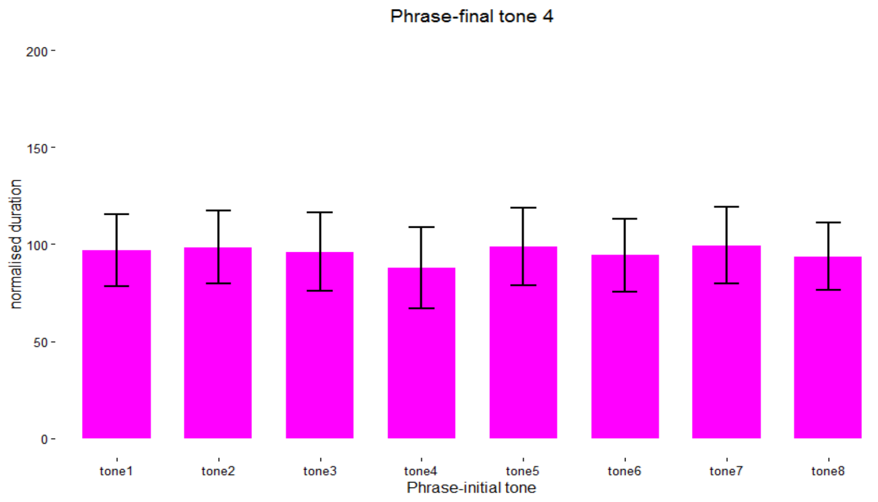


Figure 8-25. Normalised duration of phrase-final tone 4 across phrase-initial tones from 21 speakers.

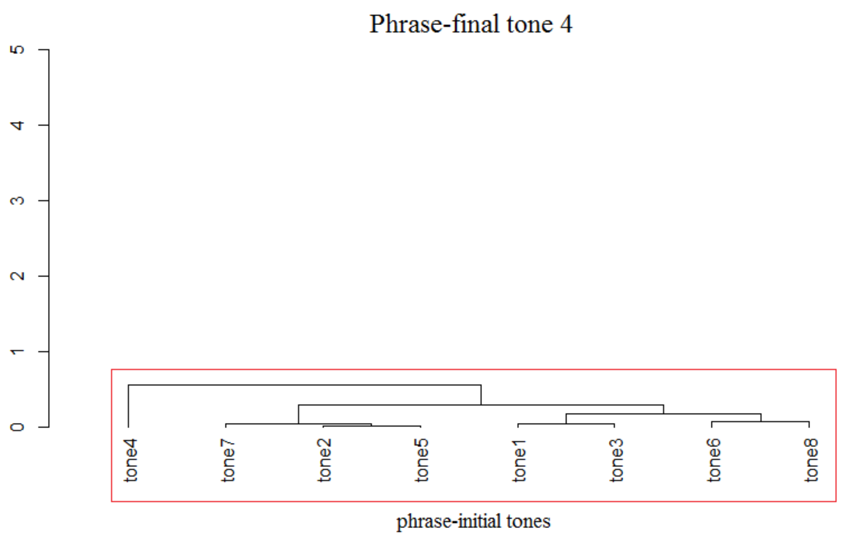


Figure 8-26. Clustering of normalised length levels of phrase-final tone 4 across eight phrase-initial tones based on pairwise *t*-tests.

8.2.1.5. Phrase-final tone 5

Perceptually, phrase-final tone 5 is realised very similarly across eight disyllabic combinations. Acoustically, most F0 contours have very close normalised duration values, but slightly different variations can also be found. For example, as shown in Figure 8-27, the contour is slightly longer after tone 8 than it is after any other tones. Nevertheless, no statistically significant differences are discovered by the pairwise *t* tests among the eight normalised duration values of the phrase-final tone 5. As shown in Figure 8-28, the terminal nodes at the bottom, representing the eight phrase-initial tones, have been clustered to two levels at the threshold of 1, indicating that the differences in the normalised duration of this phrase-final tone 5 are statistically significant, with the value after tone 8 being marginally longer than the values after other tones.

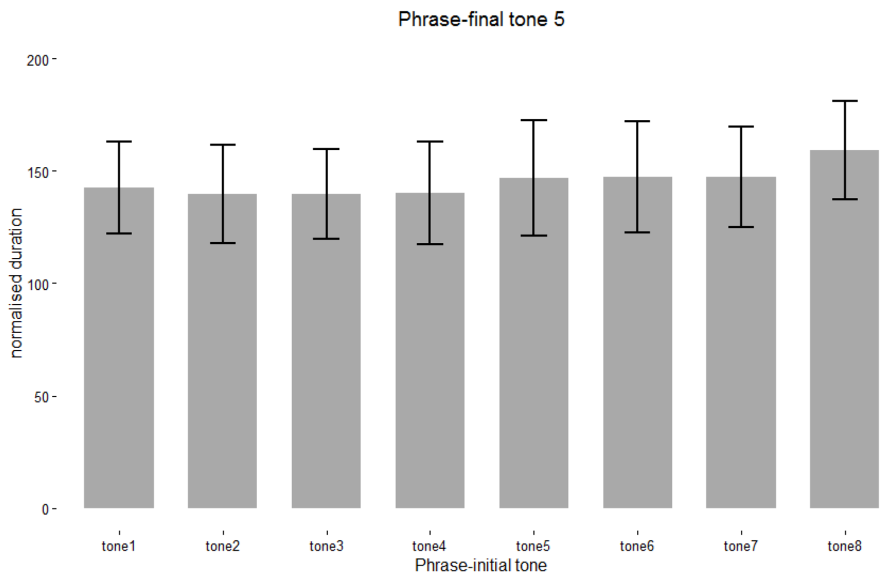


Figure 8-27. Normalised duration of phrase-final tone 5 across phrase-initial tones from 21 speakers.

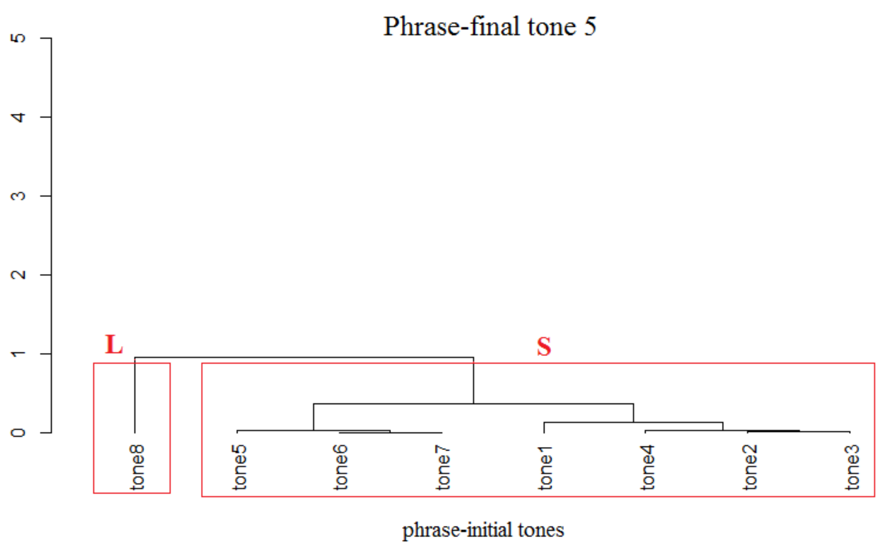


Figure 8-28. Clustering of normalised length levels of phrase-final tone 5 across eight phrase-initial tones based on pairwise *t*-tests.

8.2.1.6. Phrase-final tone 6

Phrase-final tone 6 was observed in Chapter 6 as having very similar length realisations across phrase-initial tones. The acoustically quantified results basically justify the auditory observation, but a very small amount of variation can be seen in the normalised duration values as well. As plotted in Figure 8-29, the duration after tones 1, 2, 3, and 4 are very similar, but all are trivially shorter than they are after tones 5, 6, 7, and 8, which show very similar duration values.

Nevertheless, the differences in the normalised duration values of this tone across its phrase-initial tones are not statistically significant according to the pairwise *t*-test comparisons. As shown in Figure 8-30, the eight terminal nodes, representing the duration values of this tone across eight tonal combinations, are clustered into one group at the threshold of 1. In other words, the duration realisations of phrase-final tone 6 are not phonologically affected by its preceding tones.

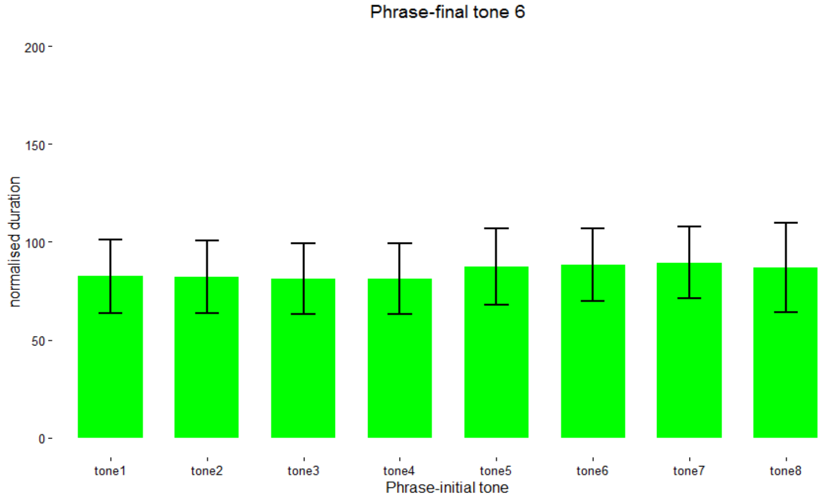


Figure 8-29. Normalised duration of phrase-final tone 6 across phrase-initial tones from 21 speakers.

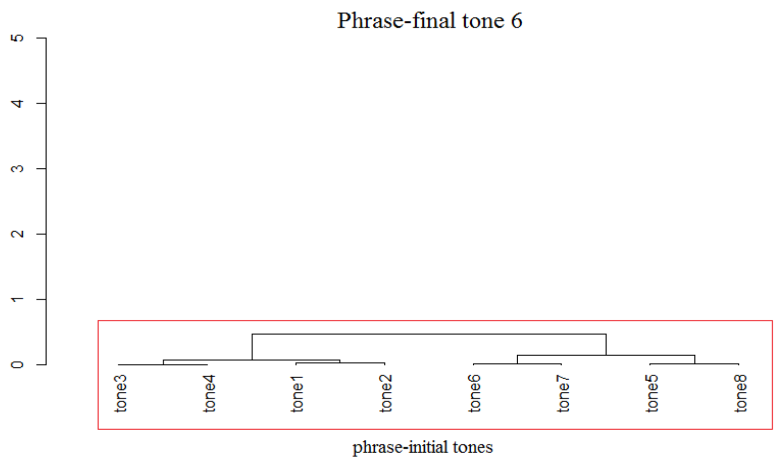


Figure 8-30. Clustering of normalised length levels of phrase-final tone 6 across eight phrase-initial tones based on pairwise *t*-tests.

8.2.1.7. Phrase-final tone 7

Perceptually, phrase-final tone 7 has very similar length realisations across phrase-initial tones. Acoustically, the normalised duration values vary somewhat among the eight combinations, as shown in Figure 8-31. For example, the duration after tone 4 is apparently shorter than it is after tone 7. Statistically, the differences among the eight normalised duration values are marginally significant according to the pairwise *t*-test comparisons, as plotted in Figure 8-32.

Statistically, the pairwise *t*-test comparison results showed significant differences among the eight normalised duration values. As Figure 8-32 indicates using the hierarchical clustering algorithm, the eight terminal nodes at the bottom, which represent the normalised duration values of this phrase-final tone 7 across eight phrase-initial tones, have been clustered into two levels at the threshold of 1. This clustering indicates that the two levels are statistically significantly different, with the values after tone 7 being significantly but also marginally longer than the values before other tones.

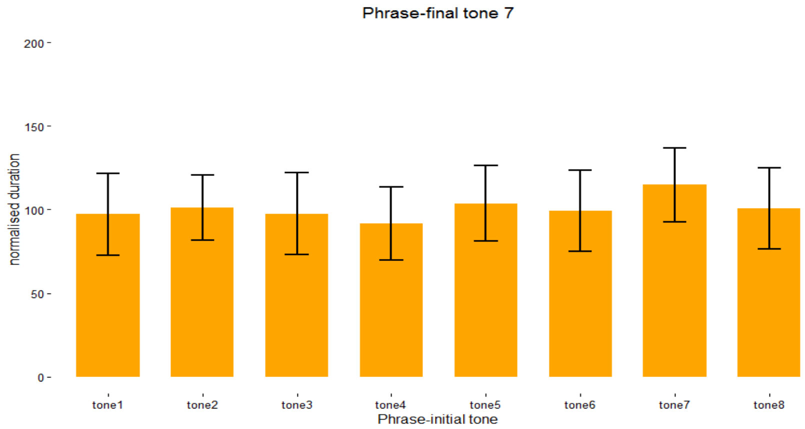


Figure 8-31. Normalised duration of phrase-final tone 7 across phrase-initial tones from 21 speakers.

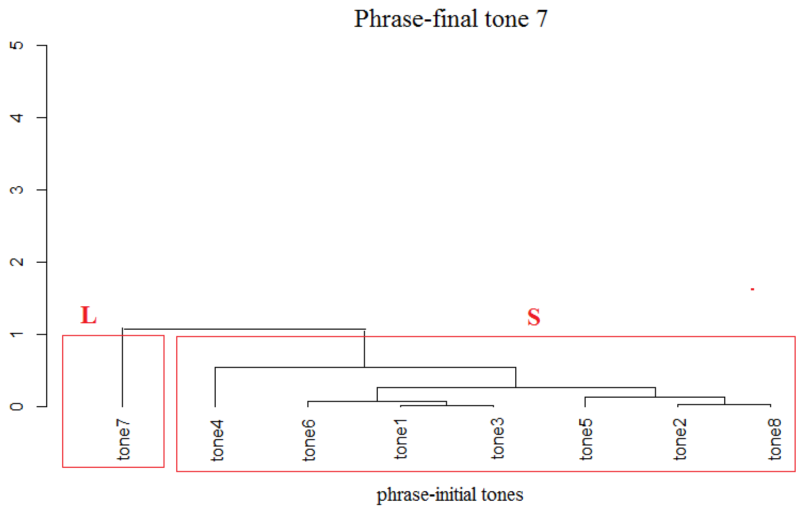


Figure 8-32. Clustering of normalised length levels of phrase-final tone7 across eight phrase-initial tones based on pairwise *t*-tests.

8.2.1.8. Phrase-final tone 8

Perceptually, determining whether the length realisation of phrase-final tone 8 may be influenced by tones in the phrase-initial position is difficult. Acoustically, a considerable number of variations are shown among the eight normalised duration values. For example, as plotted in Figure 8-33, the duration after tone 4 is apparently shorter than it is after tone 5. Statistically, according to the results of pairwise t test, all eight normalised duration values with respect to the eight phrase-initial tones are clustered into one class at the threshold of 1, as shown in Figure 8-34. This clustering indicates the phrase-initial tones do not exert statistically significant influence on the length realisation of phrase-final tone 8.

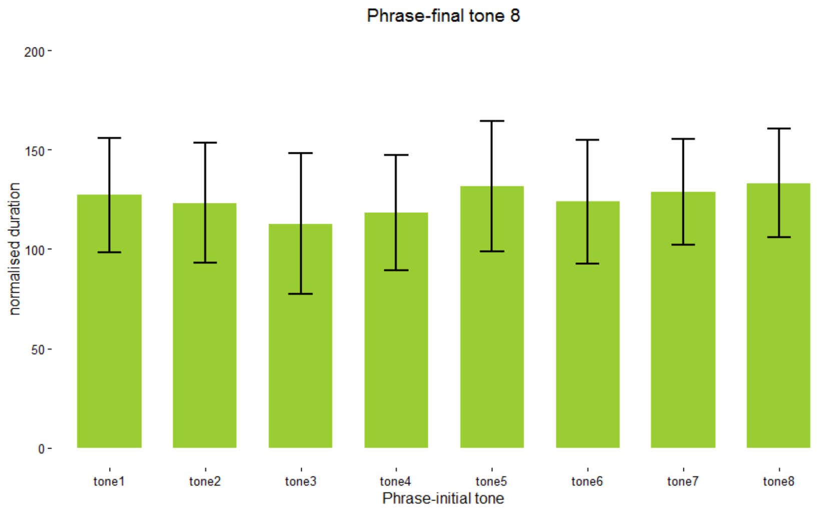


Figure 8-33. Normalised duration of phrase-final tone 8 across phrase-initial tones from 21 speakers.

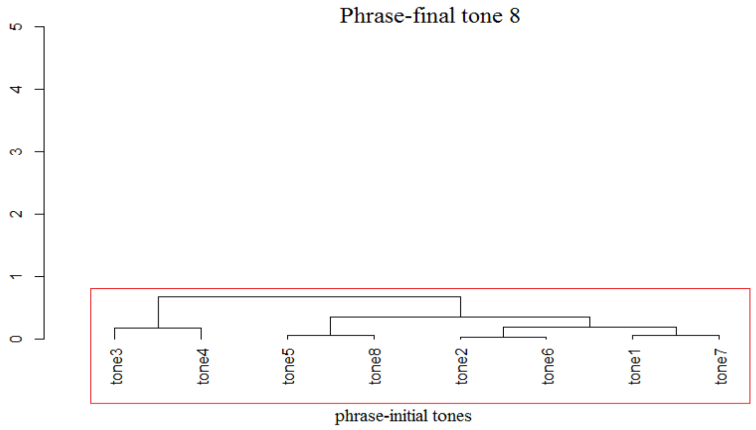


Figure 8-34. Clustering of normalised length levels of phrase-final tone 8 across eight phrase-initial tones based on pairwise *t*-tests.

8.2.1.9. Summary

As described, the length realisations of phrase-final tones are not categorically affected by tones in the phrase-initial position, but a certain degree of phonetic variation in the normalised values can be observed. In general, the duration tends to be slightly shorter after tone 4, which has an extra-high falling contour, and slightly longer before tones having mid-falling contours (e.g., tones 5, 7, and 8). Nevertheless, the acoustic differences are largely not statistically significant based on the results of the pairwise t-test comparison by effect size across 64 tonal combinations. Exceptions are phrase-final tones 5 and 7, showing two variants, as summarised in Figure 8-35, although even these are only marginally statistically significant.

The bars are expressed as a percentage of the average duration value of all phrase-final tones across 64 combinations and across 21 speakers. As indicated, tonal length also varies across different phrase-final tones; for example, phrase-final tone 6 tends to have the shortest duration. Nevertheless, whether the length differences of the 10 phonetic variants are statistically significant, how many of them are categorically contrastive, and how many of them are redundant remain crucial issues to be further tested and investigated using the pairwise t tests and the clustering algorithm, assuming that these variants are all independent and identically normally distributed.

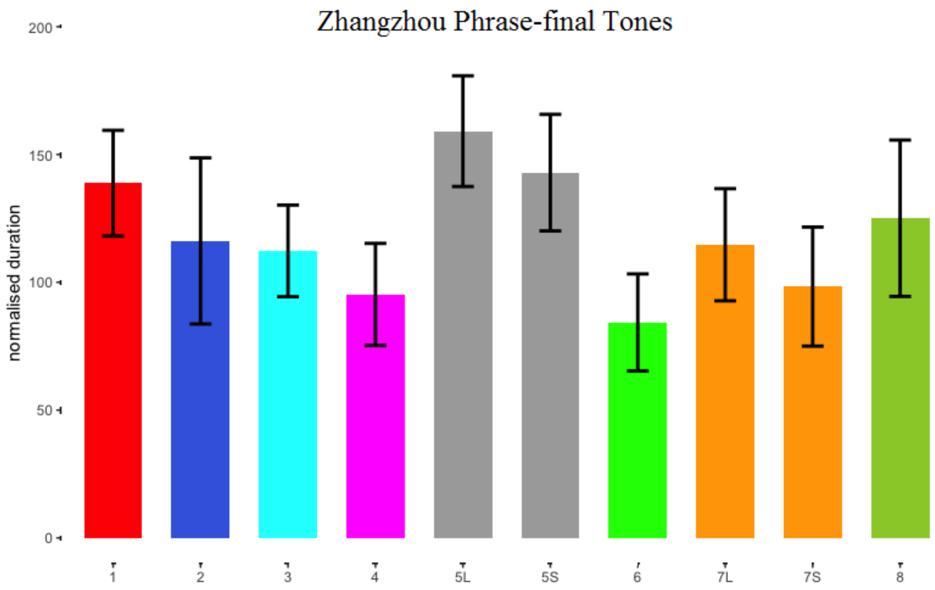


Figure 8-35. Summary of normalised F0 duration of individual Zhangzhou phrase-final tones from 21 speakers.

An exhaustive pairwise *t* test was conducted on 45 (=10*9/2) paired differences in duration among the 10 variants with the Bonferroni corrected alpha of 0.00111 to control for the Type 1 Error and to achieve significance. The testing result was visualised using the clustering algorithm as shown in Figure 8-36. At the threshold of 1, the 10 phonetic variants were clustered into three classes, with *1* representing the extra-long duration, *2* the long duration, and *3* the medium duration.

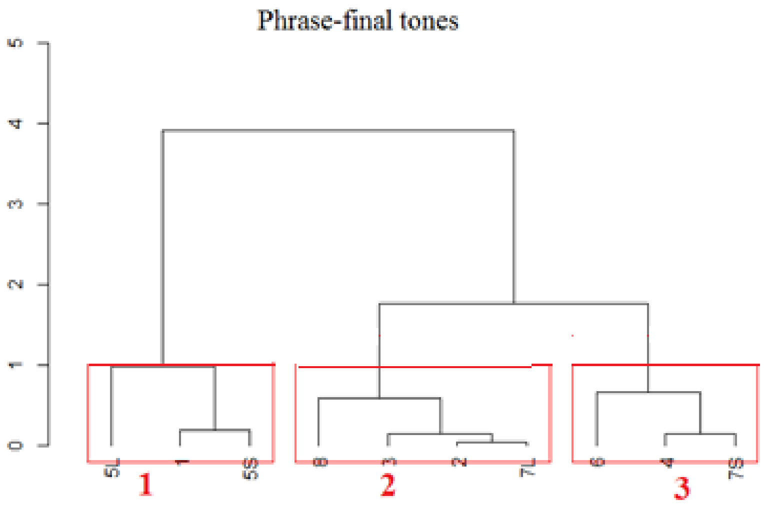


Figure 8-36. Clustering of normalised tonal length levels of all phrase-final tones based on pairwise *t*-tests.

Therefore, on the basis of acoustic quantification and statistical testing results, a linguistic-tonetic length representation of Zhangzhou phrase-final tones is shown in Table 8-2, including extra long [V::] , long [V:], and medium [V]. In addition, tone 7 has a medium length across most phrase-final tones but has a marked long variant after tone 8.

Table 8-2. Length system of Zhangzhou phrase-final tones based on acoustics and statistics

Tone	F0 contour description	Length	Notation
1	mid rising [34]/[35]	extra long	[V::]
5	mid level [33]/[43]		
2	low falling [211]/[311]	long	[V:]
3	high falling [52]		
8	low falling [211]/[311]		
4	mid-high-falling [41]/[51]	medium	[V]
6	stopped mid-high falling [41]/[51]		
7	stopped low falling [211]/[311]	(marked long)	

8.2.2. Vowel quality

Similar to the vowel quality realisation in the monosyllabic context, the high vowels were observed in Chapter 6 as undergoing diphthongisation in the phrase-final stopped tones 6 and 7 while other vowels are realised as monophthongs across all phrase-final tones. From the acoustic point of view, the changing articulatory trajectory for diphthongs indicates the corresponding acoustic correlates F1 and F2 present dynamic changing curves. Figures 8-37, 8-38, 8-39, and 8-40 show the acoustic realisations of Zhangzhou vowel quality with respect to phrase-final stopped and unstopped tones. As only high and low vowels can occur in stopped tonal environments, no examples of mid vowels are provided.

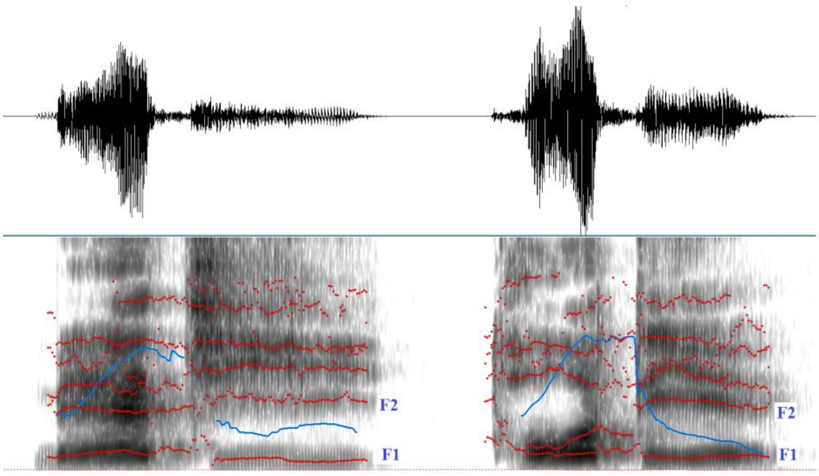


Figure 8-37. Acoustic realisations of high vowel /i/ in phrase-final unstopped tone 8 and stopped tone 7 (WYF, male).

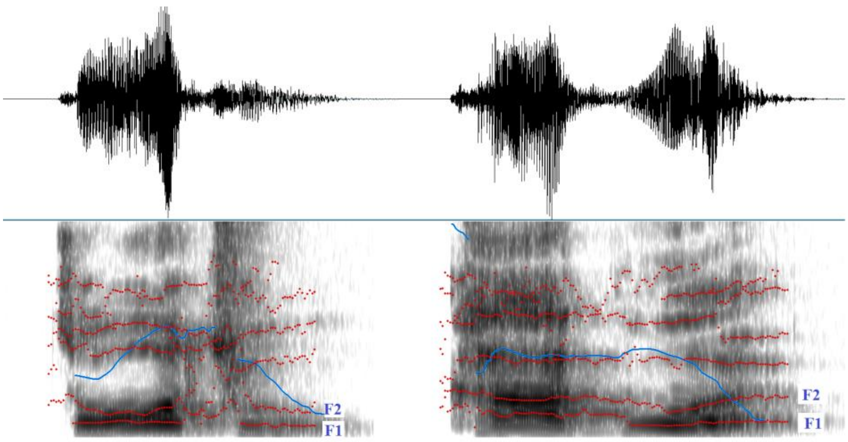


Figure 8-38. Acoustic realisations of high vowel /u/ in phrase-final unstopped tone 4 and stopped tone 6 (WYF, male).

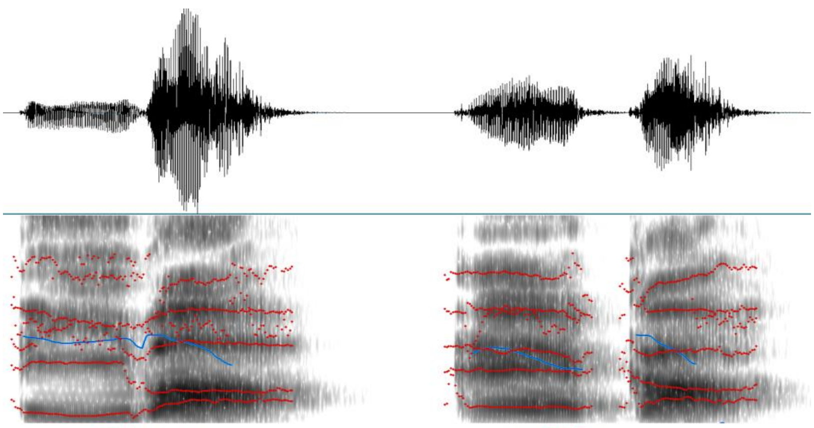


Figure 8-39. Acoustic realisations of low vowel /ɐ/ in phrase-final unstopped tone 4 and stopped tone 6 (WYF, male).

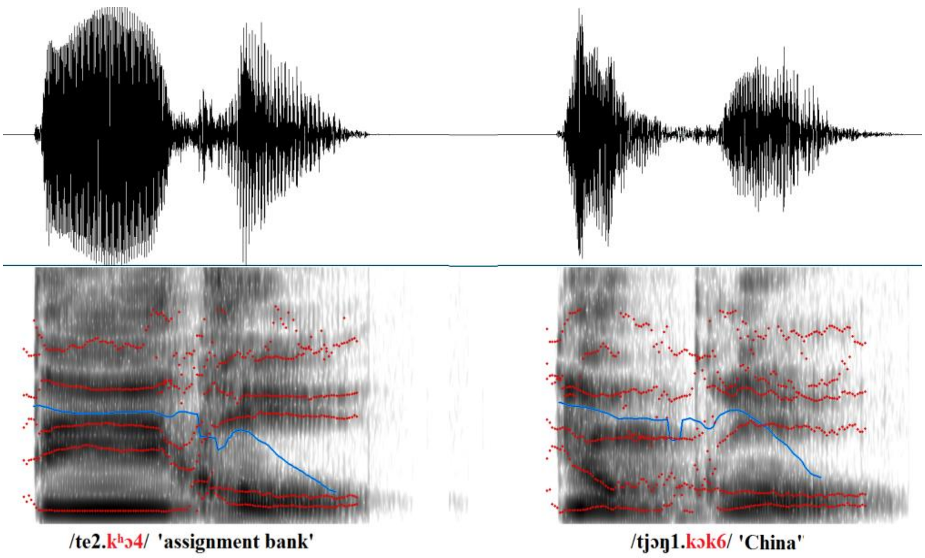


Figure 8-40. Acoustic realisations of low vowel /ɔ/ in phrase-final unstopped tone 4 and stopped tone 6 (WYF, male).

In Figure 8-37, the last morphemes of the two disyllabic tokens have the same high vowel /i/ at the underlying level and similar F0 realisations at the low frequency range. The formant patterns of the /tsi8/ morpheme in the token /ɓɛ3.tsi8/ 'horse tongue' are relatively steady and horizontal, without significant changes in the last parts of the formant curves. The relatively steady manifestations of F1 and F2 curves indicate a monophthong realisation of this high vowel in the phrase-final unstopped tone 8.

Conversely, the formant patterns of the /tit7/ morpheme in token /tʰɔ3.tit7/ 'straightforwardly' change during the second halves of the curves. The F1 has a slight rising tendency while the F2 shows a falling trend. The dynamic manifestations of a rising F1 and a falling F2 signify a gradually downward and backward movement of the tongue during the articulation of this sound. In other words, the tongue moves to a lower and central position of the oral cavity, supporting a diphthongisation of the high vowel /i/ in the phrase-final stopped tones.

In Figure 8-38, the right-most of the two disyllabic tokens have the same high vowel /u/ at the underlying level and very similar falling F0 contours at the surface. The F1 curve of the /tsʰu4/ morpheme of the disyllabic token /tsɔ3.tsʰu4/ 'ancestral house' is relatively horizontal without significant changes across time while its F2 curve is largely steady, except for the first 10% presenting a falling trend that is reasonably perturbed by the aspirated onset [tsʰ]. Therefore, the formants basically indicate a monophthong realisation of this high vowel /u/ in this unstopped tonal environment.

In contrast, the F1 and F2 curves of the /kut6/ morpheme in the second token both present a rising tendency after the second halves of the curves. The rising F1 indicates a downward movement of the tongue while the rising F2 indicates a fronting movement of the tongue. In other words, the production essentially involves two articulatory trajectories, indicating a diphthongisation of this high vowel /u/ in this stopped tone.

In Figure 8-39, the last morphemes of the two disyllabic tokens have the same underlying low vowel /ɐ/ and very similar F0 realisations. The F1 and F2 curves are both relatively stable across time in the morphemes /ɓɐ4/ in the unstopped tone 4 and /kɐk6/ in the stopped tone 6. Therefore, the acoustic manifestations essentially justify the auditory observations of monophthong realisation of the low vowel /ɐ/ across phrase-final tones in Zhangzhou. Similarly, in Figure 8-40, the first and second formants of the last morphemes /kʰɔ4/ and /kɔk6/ do not show significant change after the first 10%, indicating a monophthong realisation of the low vowel /ɔ/ in both stopped and unstopped tones.

In summary, the formant patterns suggest two high vowels undergo diphthongisation in the phrase-final stopped tones, while other vowels are realised consistently as monophthongs across the tones in the phrase-final position.

8.2.3. Voice quality

Three types of phonation—modal, breathy, and creaky—were noted in Chapter 6, but their distributions are simultaneously conditioned by vowel quality, pitch and tone. The breathy voice dominantly occurs on high vowels but only in unstopped tones. The creaky voice largely occurs on low vowels with falling pitch contours but is also often seen on the falling-pitch mid vowel [e] among male speakers. The modal voice, on the other hand, is largely found on mid vowels across different tones and on low vowels with non-falling pitch contours.

As discussed in Chapter 5, the correlation between voice quality, vowel quality, pitch, and tone has a robust articulatory basis, involving a complicated but systematic covariation of articulators in the laryngeal and supraglottal positions, as well as the volume and velocity of aerodynamic forces (Ladefoged, 1971, 2003; Laver, 1980, 1994; Gordon & Ladefoged, 2002, Lotto et al., 1997). From the acoustic point of view, the changing articulatory configurations for different phonations inevitably cause changes in the acoustic manifestations.

Figures 8-41, 8-42, and 8-43 show how the acoustic realisations of voice quality for phrase-final morphemes change with respect to vowel quality, pitch and tone in Zhangzhou. The illustrated vowels include high, mid, and low. Because all vowels undergo laryngealisation and high vowel undergoes diphthongisation in stopped tones, the illustrated tones are all unstopped covering four major F0 shapes—rising [35], level [33], low-falling [211], and high-falling [52]. The waveforms and spectrums are consistently extracted from the last 10% of the related vowel articulations.

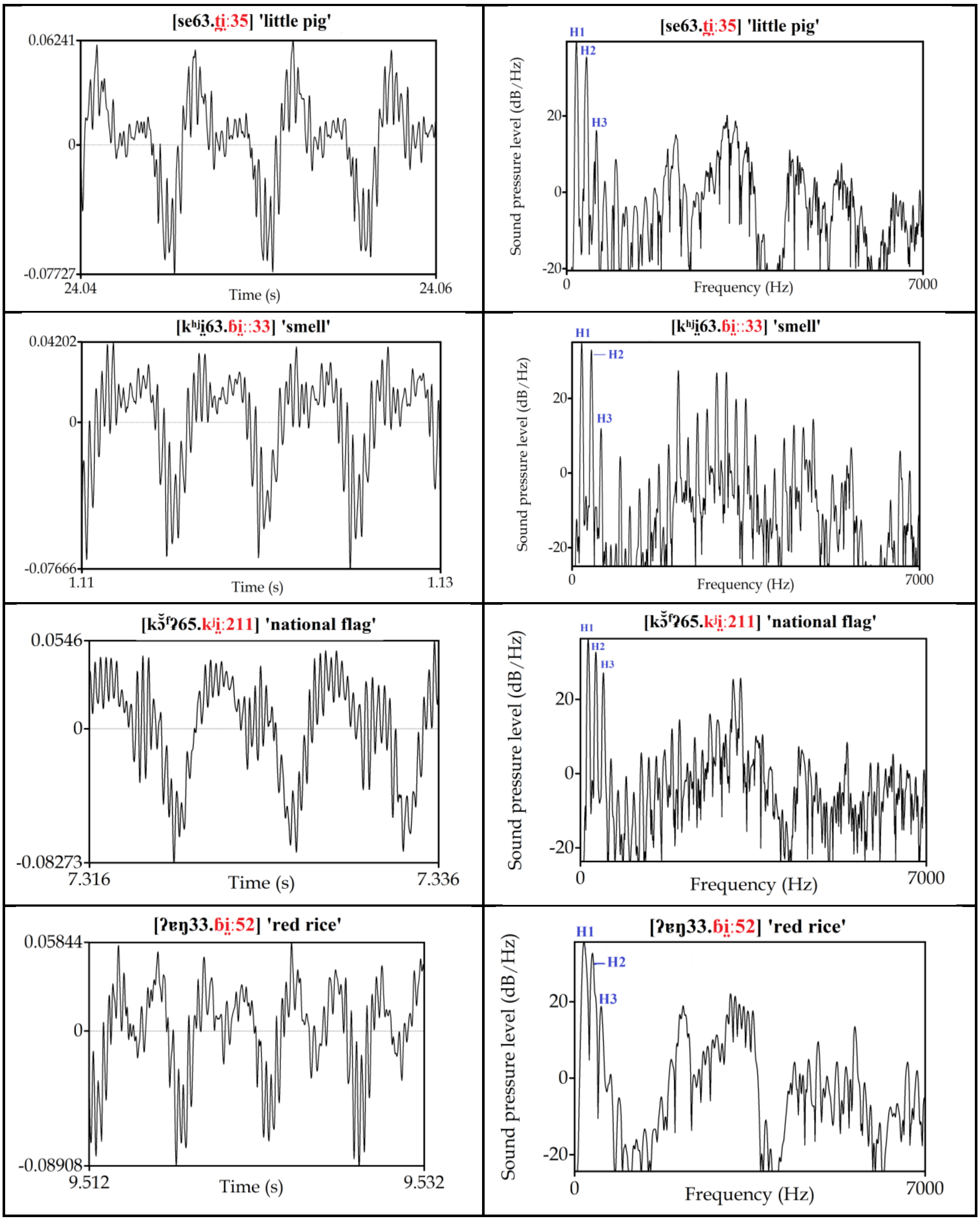


Figure 8-41. Phonation realisations of phrase-final high vowel /i/ in four different F0 contexts (WYF, male).

Figure 8-41 shows the voice quality realisations of phrase-final high vowel /i/ across four F0 contours. In the waveforms, a dense amount of random and rapid fluctuation is always superimposed on the glottal pulses across time regardless of whether the related F0 contour is rising [35], level [33], low falling with a level plateau [211], or falling [52]. Correspondingly, the first harmonic consistently has a higher amplitude level than any other harmonics across the frequency ranges; that is, the spectral tilt of H1-H2 is steeply positive. The two acoustic manifestations are characteristic of breathy voice; therefore, this high vowel is produced with a breathy voice regardless of the shapes of the F0 contours.

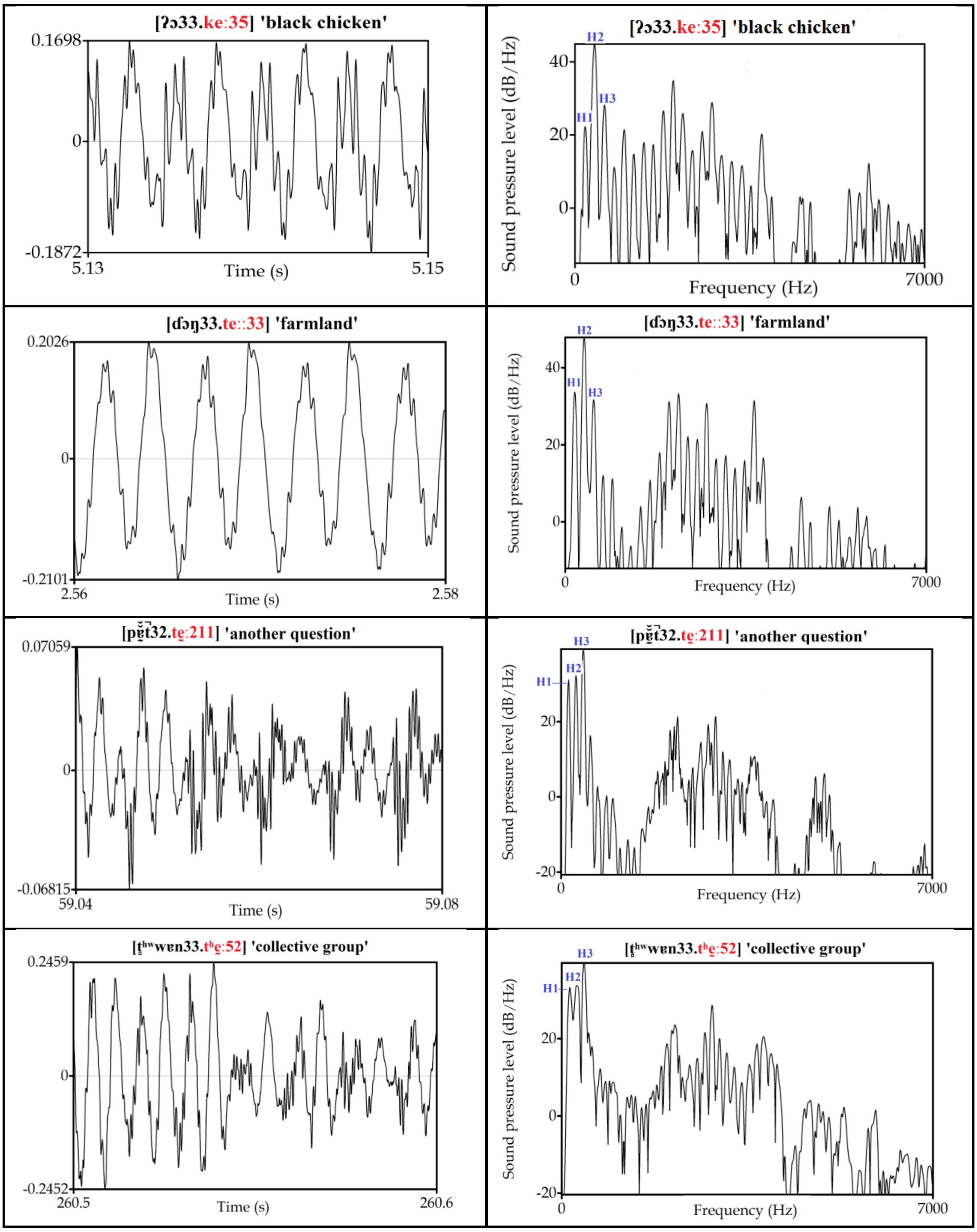


Figure 8-42. Phonation realisations of the phrase-final mid vowel /e/ in four different F0 contexts (WYF, male).

Figure 8-42 shows the phonation realisations of mid vowel /e/. In the rising [35] and mid-level [33] contexts, the waveforms are not superimposed with a large amount of random and persistent fluctuation. The time intervals between glottal pulses are relatively regular and evenly spaced. In addition, the amplitude of the second harmonic is steeply higher than the amplitude of its surrounding harmonics. That is, it presents a negative spectral tilt of H1-H2. Such acoustic manifestations of regular glottal pulses and negative H1-H2 are characteristic of a typical modal voice.

Nevertheless, in the falling tonal environments [211] and [52], the mid vowel presents different acoustic realisations. The time intervals of glottal pulses become irregular and less frequent. The amplitude values of the first harmonic and the second harmonic are very similar, but both are lower than the value of the third harmonic. The irregular glottal pulses and nearly zero spectral tilt of H1-H2 are among major acoustic characteristics of creaky voice because the vocal folds at this point are tightly adducted and form a compressed mass, causing them to vibrate irregularly and less frequently. Therefore, the mid vowel is essentially articulated with creaky phonation in these two falling environments by this male speaker.

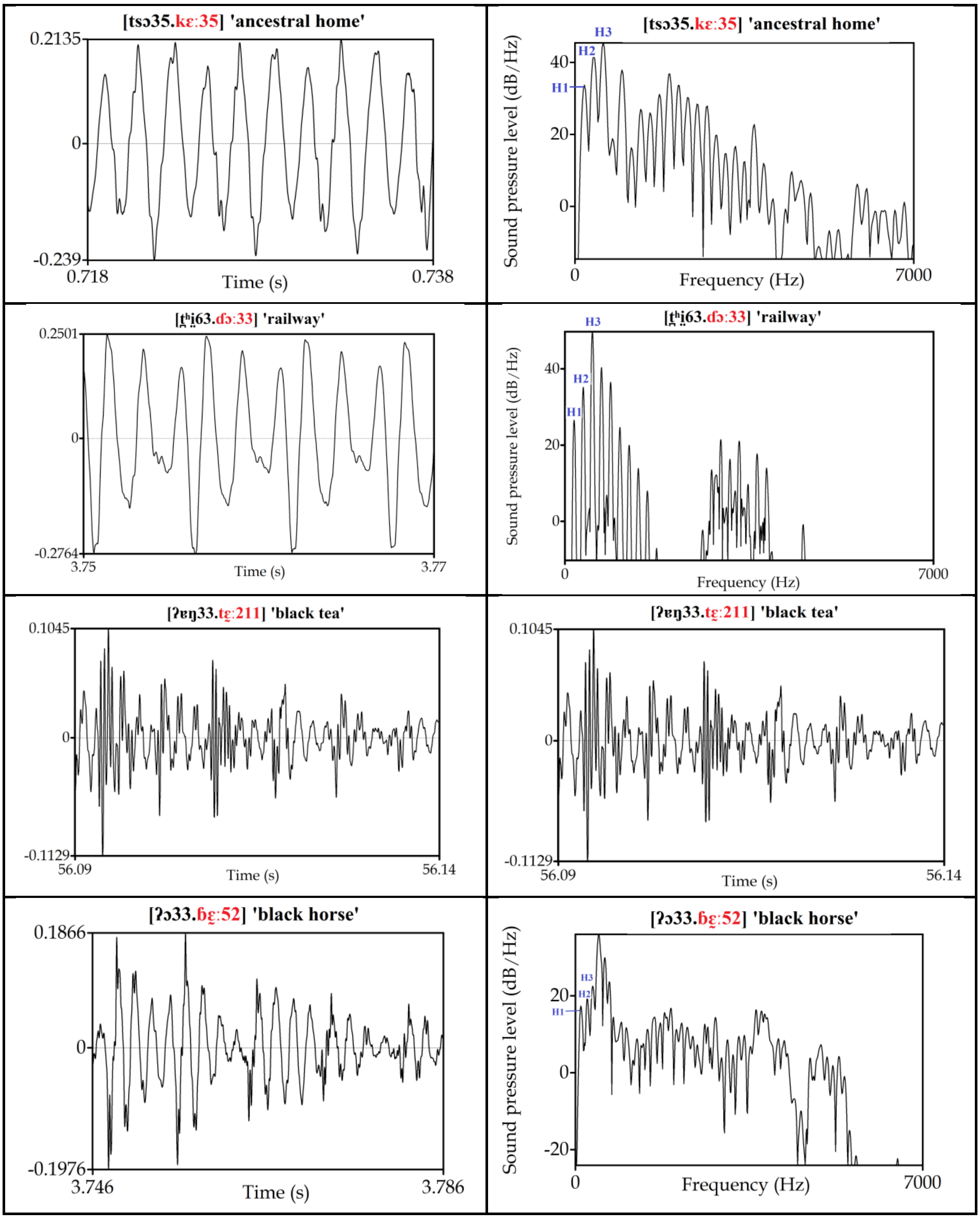


Figure 8-43. Phonation realisations of the phrase-final low vowel /ɛ/ or /ɔ/ in four different F0 contexts (WYF, male)

Figure 8-43 shows the voice quality realisations of low vowels. Similar to the acoustic realisations of mid vowel /e/, the low vowels also present acoustic manifestations of a typical modal voice in the rising [35] and mid level [33] contours: regular glottal pulses and negative spectral tilt of H1-H2.

Nonetheless, in the falling contours [211] and [52], the glottal pulses are irregularly distributed while the spectral tilt of H1-H2 is neither steeply positive, as for the breathy voice, nor steeply negative, as for the modal phonation, but rather waving slightly around zero. For example, in the [211] contour, the spectral tilt of H1-H2 is slightly higher than zero while, in the [52] context, the tilt becomes slightly lower than zero. Such acoustic manifestations are generally characteristic of a creaky voice. Therefore, the low vowels change their phonation from modal voice to creaky voice in the falling-pitched tones. In summary, the correlations between voice quality, pitch, tone and vowel quality are acoustically supported.

8.2.4. Syllable coda

In contrast to the preservation of obstruent coda in the phrase-initial syllables, obstruent codas were indicated as perceptually not realised in the phrase-final position in Chapter 6, a finding similar to that for their realisations in the monosyllabic setting. From the articulatory point of view, the nonrealisation of obstruent codas indicates no oral constriction is created after production of preceding vowels. Correspondingly, no formant transition is expected to show on spectrograms. Figure 8-44 shows the acoustic realisations of different obstruent codas in phrase-final syllables that have same vowels and tones underlyingly.

As indicated in Figure 8-44, the formant patterns of F1 and F2 are consistently stable without significant changes across time to all phrase-final syllables investigated. The steady manifestation of formant patterns acoustically indicates (a) the related vowel is a monophthong and (b) no obstruent coda is produced following the vowel. The low vowels have been shown to be realised as monophthongs in the stopped tones (Section 8.2.2), therefore, the acoustic manifestation supports the auditory observation of non-realisation of obstruent codas phrase finally.

On the contrary, if obstruent codas are realised phrase finally in Zhangzhou, the formant transitions are expected to occur on the spectrograms. For example, in the /tɐt7/ morpheme, the F2 must have had a rising tendency toward the frequency range around 1600 Hz as a result of influence from the alveolar coda production; however, no rising trend can be seen on the F2 curve, justifying the obstruent coda is not realised in this utterance.

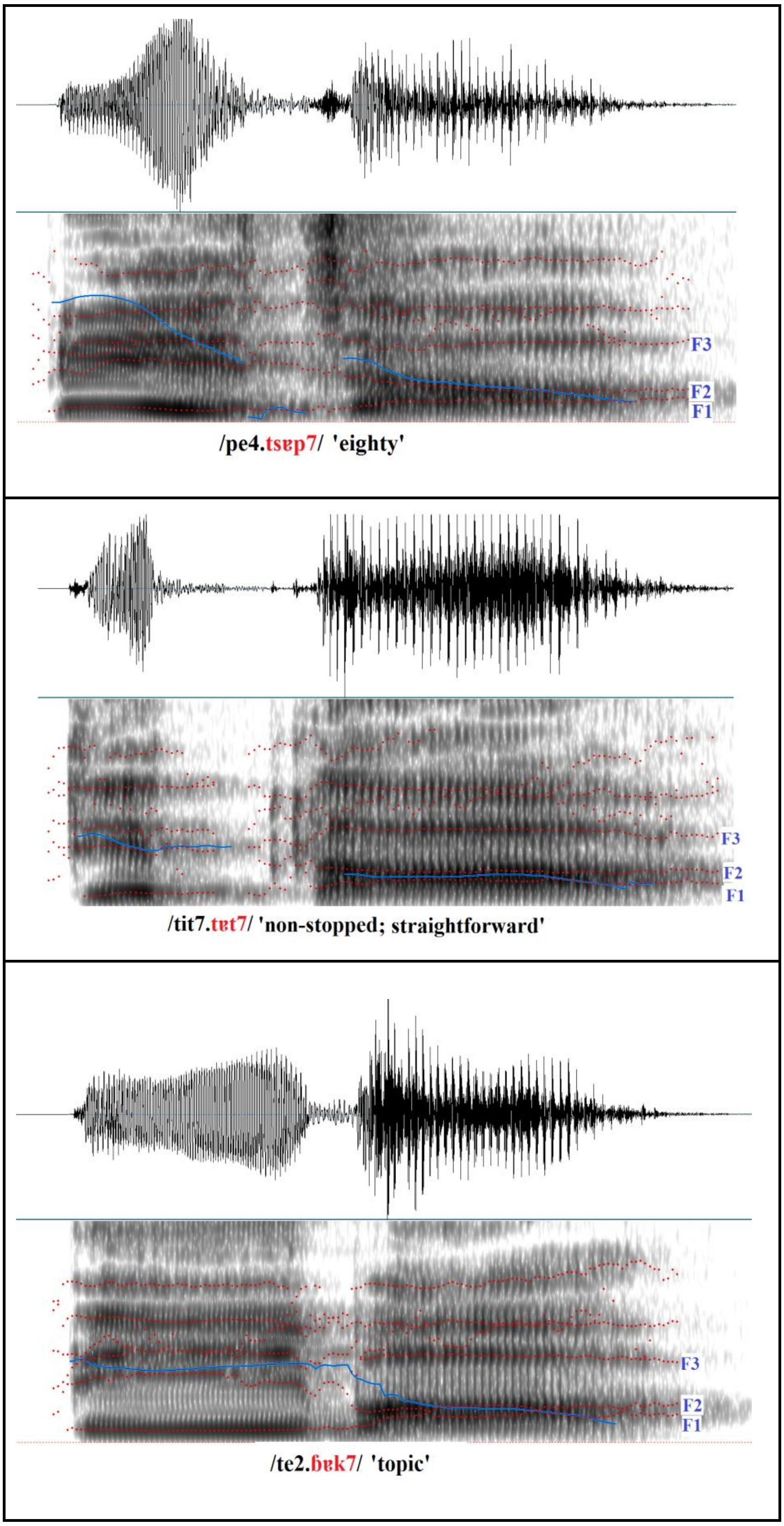


Figure 8-44. Acoustic realisation of the obstruent codas in phrase-final position.

The non-realisation of obstruent codas in the phrase-final position on the other hand causes several phonetic effects on the whole syllables, including syllable lengthening, high vowel diphthongisation, vowel laryngealisation, and F0 contour depression.

- The syllable lengthening effect can be seen in the disyllabic token /tit7.tɐt7/ ‘non-stopped; straightforward’ in Figure 8-44. The morphemes /tit7/ and /tɐt7/ both have tone 7 and the alveolar obstruent coda /t/ at the underlying level; however, the phrase-final morpheme /tɐt7/ apparently has a much longer duration than the phrase-initial morpheme /tit7/.
- The high vowel diphthongisation effect can be seen in Figures 8-37 and 8-38 (Section 8.2.2), where both F1 and F2 curves show dynamic changes in the second halves of phrase-final morphemes in the stopped tones.
- The vowel laryngealisation effect can be seen in Figure 8-45. The low vowel [ɐ] consistently presents the creaky manifestation of similar H1 and H2 amplitude values across three phrase-final morphemes that have underlyingly identical vowels and tones but differ in obstruent coda types.
- The F0 contour depression was addressed in Section 8.1.7, with phrase-final tone 7 consistently having a low-falling contour of either [211] or [311] across different phrase-initial tones.

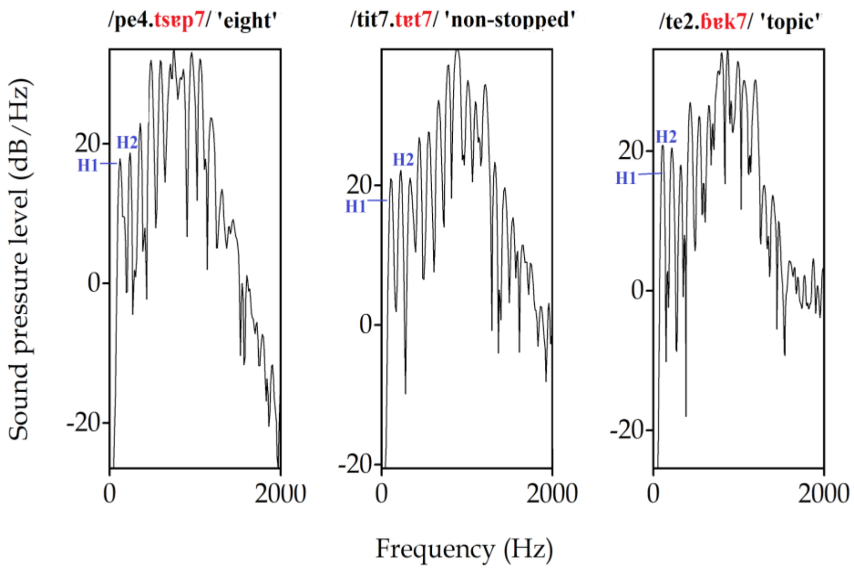


Figure 8-45. Acoustic manifestation of low vowel laryngealisation in the phrase-final stopped tone (WYF, male).

8.2.5 Summary

According to the description in this section, a bundle of segmental and suprasegmental parameters function together to code the tonal realisations and distinctions in the phrase-final context. Table 8-3 summaries the multidimensional realisations of Zhangzhou phrase-final tones, including six dimensions. In the F0 column, the unmarked form is shown on the left while the unmarked form on the right. For example, the phrase-final tone 6 indicates a mid-high falling contour, medium length, high vowel diphthongisation, creaky vowel, and not-realised obstruent coda; while phrase-final tone 1 suggests a mid-rising contour, extra long length, no diphthongisation, breathy high vowel, modal-voiced low vowel and sonorant coda. Nevertheless, questions remain as to how these phrase-final tones are related to each other from the phonological point of view, as well as how the patterns hidden within this framework can be revealed and interpreted.

Similarly, to address these issues, the SplitsTree software was employed to generate a phylogram in order to visualise the internal relations among these phrase-final tones. The multidimensional properties for each tone were firstly transformed into a multiple sequence alignment in order to be computed in the SplitsTree software. Figure 8-46 shows the generated result of the phonological mapping between Zhangzhou phrase-final tones. The root of the tree represents the set of phrase-final tones being investigated. The horizontal lines are branches representing the amount of similarity in terms of multidimensional realisations across eight phrase-initial tones. The vertical lines indicate the relatedness between tones. The closer the lines, the higher the probability that the tones belong to the same categories. The more divergent the lines, the higher the probability that the tones are not related categorically in this phrase-final context. As suggested by this figure, two aspects deserve further discussion.

The tones that share similar F0 realisation can be categorically different on the basis of the multidimensional framework. For examples, phrase-final tones 4 and 6 have similar mid-high falling contour but are classified as two different tones, since they differ in the parameters of phonation and syllable coda type. Similarly, phrase-final tone 7 (F7) is not clustered with phrase-final tones 2 and 8 (F2 and F8), although they share similar F0 contours. These facts further indicate that the single dimension of F0/pitch is not sufficient to characterise Zhangzhou tones.

The process of tonal contrast neutralisation can also occur in the phrase-final environment. As indicated in Figure 8-46, phrase-final tones 2 and 8 are clustered together without significant differences in the realisations of multiple parameters.

In summary, the multidimensional realisations of Zhangzhou phrase-final tones are also not only auditorily supported but also firmly grounded in both acoustic and articulatory reality.

Table 8-3. Multidimensional realisations of Zhangzhou phrase-final tones

Tone	F0	Length	Diphthong	High vowel	Low vowel	Syllable coda
1	[34]/[35]	extra long	-	breathy	modal	sonorant
2	[211]/[311]	long	-	breathy	creaky	sonorant
3	[52]	long	-	breathy	creaky	sonorant
4	[41]/[51]	medium	-	breathy	creaky	sonorant
5	[33]/[43]	extra long	-	breathy	modal	sonorant
6	[41]/[51]	medium	+	creaky	creaky	not realised obstruent
7	[211]/[311]	medium/long	+	creaky	creaky	not realised obstruent
8	[211]/[311]	long	-	breathy	creaky	sonorant

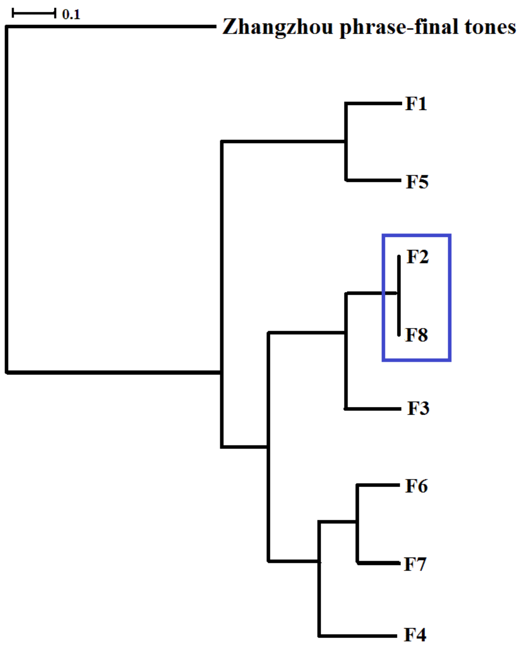


Figure 8-46. Phylogram representation of Zhangzhou phrase-final tones (F = Phrase-final).

8.3. Previous Studies

No previous studies have explored how the phrase-final tones are realised acoustically and how they interact with the individual phrase-initial tones in Zhangzhou. The descriptions in this study address the research gap and advance the understanding of the nature of tonal interactions in Zhangzhou.

8.4. Conclusion

This chapter has provided a detailed exploration of the acoustic properties of phrase-final tones, supplemented by statistical testing and articulatory explanation. The results show the realisations of the phrase-final tones are not categorically affected by tones in the phrase-initial position but are higly sensitive to their surrounding phonetic environments. For example, the F0 onsets of most phrase-final tones have statistically significantly different variants conditioned by the F0 offset of their preceding tones.

Tonal realisations in the phrase-final environment are also multidimensional, involving systematic interactions among various parameters. The tones that share similar F0 contour can be categorically different on the basis of the multidimensional framework. In addition, the neutralisation of tonal contrast also occurs in this context. The findings provide a solid and objective foundation for the theoretical explanation of the mapping of the relatedness between the citation and disyllabic tones in Chapter 9.

Part D: Mappings Between Citation and Disyllabic Tones

Chapter 9: Theoretical Considerations

Previous chapters have provided detailed descriptions of the auditory and acoustic properties of Zhangzhou citation and disyllabic tones in terms of multidimensional parameters. This chapter primarily considers how the disyllabic tones are related to the citation tones and explains how native speakers structure the relations among the tonal realisations in different linguistic contexts in their mental grammar.

9.1. Mapping Between Citation and Phrase-Initial Tones

Two questions are examined in this section. (1) To what extent do the tonal realisations in the phrase-initial position differ from those in citation with respect to the multidimensional parameters? (2) How are the phrase-initial tones related into the citation tones from the phonological point of view?

9.1.1. Pitch/F0

Table 9-1 shows the pitch/F0 comparison of Zhangzhou citation and phrase-initial tones. Several notable aspects can be generalised.

- All phrase-initial tones have completely different pitch/F0 realisations from their corresponding citation forms. For example, the pitch/F0 of tone 1 is a mid-rising contour [35] in citation but changes to a mid-level [33] in the phrase-initial position.
- Tones with different pitch/F0 realisations in citation can have very similar pitch/F0 contours in the phrase-initial context. For example, the pitches/F0 of tones 1 and 2 are different in citation but are neutralised to [33] in the phrase-initial position. Similarly, tones 5 and 8 neutralise their pitch/F0 contrasts to [32] phrase initially.
- Tones with similar pitch/F0 realisations in citation can have different pitch/F0 contours in the phrase-initial position. For example, the pitch/F0 contours of tones 2 and 8 are indistinguishable in citation but are different phrase initially.
- The pitch/F0 inventory of phrase-initial tones tends to overlap with that of the citation tones to a certain degree but presents new contours. For example, both systems have rising [35] and level [33], but new pitch/F0 contours are produced phrase initially, including extra-high falling [63], extra-high and short falling [65], and mid-falling [32].

9.1.2. Duration

Table 9-2 shows the comparison of the length realisations of Zhangzhou citation and phrase-initial tones. Three notable aspects should be further mentioned.

- The number of length contrast is considerably reduced phrase initially. Four different length levels are identified in citation while only two are observed in the phrase-initial context.
- A new length realisation of an extra-short property is generated phrase initially for stopped tones 6 and 7.

- Length realisations classify phrase-initial tones into either stopped tones or unstopped tones. All unstopped tones have statistically indistinguishable medium lengths while stopped tones have similar extra-short lengths.

Table 9-1. Pitch/F0 comparison among Zhangzhou citation and phrase-initial tones

Tone	Citation	Phrase-initial
1	[35]	[33]
2	[22]	[33]
3	[51]	[24]/[25]
4	[41]	[63]
5	[33]	[32]
6	[41]	[65]
7	[221]	[32]
8	[22]	[32]

Table 9-2. Length comparison among Zhangzhou citation and phrase-initial tones

Tone	Citation	Phrase-initial
1	extra long	medium
2	extra long	medium
3	medium	medium
4	medium	medium
5	extra long	medium
6	short	extra short
7	long	extra short
8	extra long	medium

9.1.3. Vowel quality

The high vowels /i/ and /u/ have different realisations across different linguistic contexts. In citation, they are diphthongised to [iɛ̯] and [uɤ̯] in the stopped tones 6 and 7, but they are realised as monophthongs in the phrase-initial context, as shown in Table 9-3. The diphthongisation of high vowels is thought to be induced by the non-realisation of obstruent codas in the monosyllabic setting while their monophthong realisations are considered to be associated with the preservation of obstruent codas in the phrase-initialsyllables.

Table 9-3. Vowel quality realisation comparison of Zhangzhou citation and phrase-initial tones (C = citation; PI = phrase-initial)

	/i/		/u/		/e/		/ɵ/		/ɛ/		/ɐ/		/ɔ/	
Tone	C	PI	C	PI	C	PI	C	PI	C	PI	C	PI	C	PI
1	[i]	[i]	[u]	[u]	[e]	[e]	[ɵ]	[ɵ]	[ɛ]	[ɛ]	[ɐ]	[ɐ]	[ɔ]	[ɔ]
2	[i]	[i]	[u]	[u]	[e]	[e]	[ɵ]	[ɵ]	[ɛ]	[ɛ]	[ɐ]	[ɐ]	[ɔ]	[ɔ]
3	[i]	[i]	[u]	[u]	[e]	[e]	[ɵ]	[ɵ]	[ɛ]	[ɛ]	[ɐ]	[ɐ]	[ɔ]	[ɔ]
4	[i]	[i]	[u]	[u]	[e]	[e]	[ɵ]	[ɵ]	[ɛ]	[ɛ]	[ɐ]	[ɐ]	[ɔ]	[ɔ]
5	[i]	[i]	[u]	[u]	[e]	[e]	[ɵ]	[ɵ]	[ɛ]	[ɛ]	[ɐ]	[ɐ]	[ɔ]	[ɔ]
6	[iɛ̯]	[i]	[uɤ̯]	[i]	*	*	*	*	*	*	[ɐ]	[ɐ]	[ɔ]	[ɔ]
7	[iɛ̯]	[i]	[uɤ̯]	[i]	*	*	*	*	*	*	[ɐ]	[ɐ]	[ɔ]	[ɔ]
8	[i]	[i]	[u]	[u]	[e]	[e]	[ɵ]	[ɵ]	[ɛ]	[ɛ]	[ɐ]	[ɐ]	*	*

9.1.4. Voice quality

Two notable aspects can be observed with respect to the voice quality realisations among Zhangzhou citation and phrase-initial tones, as shown in Table 9-4 (P-initial=phrase-initial).

- The voice quality realisations of the same vowels in the same tones can change across citation and phrase-initial environments in Zhangzhou. For example, the low vowels in tones 3 and 8 alternate between modal and creaky while high vowels in tone 7 alternate between creaky and breathy in these two contexts.
- A new phonation type occurs in the specific phrase-initial tone. For example, the falsetto voice is observed in the vowels, regardless of high or low, in the phrase-initial stopped tone 6, which has an extra-high pitch/F0 and an extra-short length.

Table 9-4. Voice quality realisation comparison among Zhangzhou citation and phrase-initial tones

	High vowel		Mid vowel		Low vowel	
Tone	Citation	P-initial	Citation	P-initial	Citation	P-initial
1	breathy	breathy	modal	modal	modal	modal
2	breathy	breathy	modal	modal	modal	modal
3	breathy	breathy	modal/creaky	modal	creaky	modal
4	breathy	breathy	modal/creaky	modal/creaky	creaky	creaky
5	breathy	breathy	modal	modal/creaky	modal	creaky
6	creaky	falsetto	*	*	creaky	falsetto
7	creaky	breathy	*	*	creaky	creaky
8	breathy	breathy	modal	modal/creaky	modal	creaky

9.1.5. Syllable coda

The realisations of obstruent codas are different across phrase-initial and citation contexts. They tend not to be realised in citation but are realised in the phrase-initial environment, as shown in Table 9-5. Regardless of whether the obstruent codas are realised or not, sets of phonetic effects can be induced on the whole syllables, as shown in Table 9-6. For example, syllables are significantly shortened in the phrase-initial position, where the obstruent codas are preserved, but are lengthened in the citation, where they tend not to be realised.

Table 9-5. Obstruent coda realisation comparison among Zhangzhou citation and phrase-initial tones

	Sonorant coda		Obstruent coda	
Tone	Citation	Phrase-initial	Citation	Phrase-initial
1	+	+	-	-
2	+	+	-	-
3	+	+	-	-
4	+	+	-	-
5	+	+	-	-
6	-	-	not realised	realised
7	-	-	not realised	realised
8	+	+	-	-

Table 9-6. Comparison of Phonetic effects induced by the obstruent coda realisations among Zhangzhou citation and phrase-initial tones

Parameters	Citation	Phrase-initial
Obstruent coda	not realised	realised
Duration	lengthened	shortened
Pitch/F0	depressed	raised
Vowel quality	high vowel diphthongisation	monothongisation
Voice quality	creaky	falsetto

9.1.6. Summary

As discussed above, phrase-initial and citation tones are both multidimensional. Each tone consists of a bundle of different phonetic features. Nevertheless, the same tones can be realised differently in the phrase-initial position with respect to the parameters. Exploring how the phrase-initial tones are mapped onto the citations and discovering to what extent they are associated with each other is crucial to understanding the nature of tone sandhi in Zhangzhou. To address this issue, the SplitsTree software was applied to visualise the phonological relatedness among the citation and phrase-initial tones on the basis of the multidimensional framework discussed.

The multidimensional properties of 16 tones including 8 citation and 8 phrase-initial tones were firstly transformed into a multiple sequence alignment to be computed in the SplitsTree for generating a phylogram, as shown in the Figure 9-1. The terminal nodes on the right indicate the 16 tones being investigated. The horizontal lines suggest the amount of phonetic similarity shared by these tones. The shorter the lines, the stronger amount of similarity that tones share in phonetics, and vice versa. The lines in the vertical dimension on the other than represents the relatedness between tones. The closer the lines, the higher probability the tones are related phonologically.

As indicated in Figure 9-1, the phrase-initial tones are largely not related to their corresponding citation tones because most of the pairs are not clustered together in the phylogram. For example, citation tone 1 (C1) is not clustered with phrase-initial tone 1 (I1), and citation tone 2 (C2) and phrase-initial tone 2 (I2) are on different layers. Phrase-initial tone 4 (I4) appears more likely to related with citation tone 4 (C4), indicating the considerable amount of similarities they share in the phonetics, but their contrasts are not fully neutralised because they differ in pitch/F0. For most tonal pairs, their associations tend not to be significant. On the whole, the mapping result indicates the realisations of phrase-initial tones and their corresponding citation tones are largely not associated at either the phonetic or the phonological level; thus, they should be considered two independent systems.

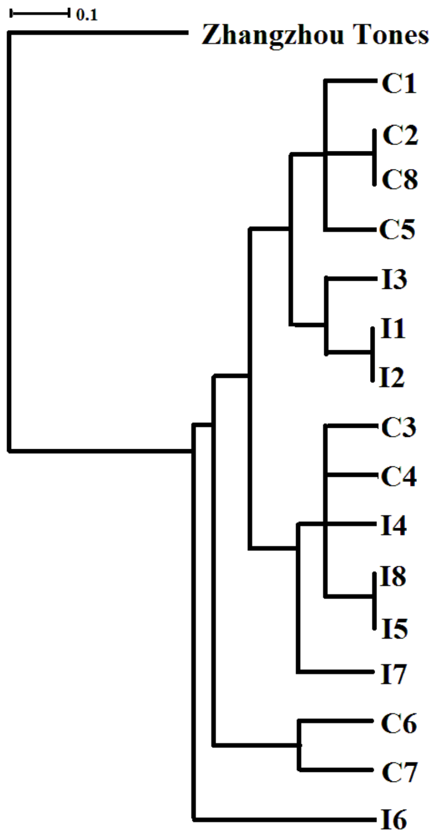


Figure 9-1. Mapping between Zhangzhou citation and phrase-initial tones (C = citation tone; I = phrase-initial tone).

9.2. Mapping Between Citation and Phrase-Final Tones

This section addresses two issues. (1) To what extent do the tonal realisations in phrase-final position differ from those in citation with respect to the multidimensional parameters? (2) How are the phrase-final tones mapped onto the citation tones from the phonological point of view?

9.2.1. Pitch/F0

Table 9-7 shows the comparison of the pitch/F0 realisations of Zhangzhou citation and phrase-final tones. In general, the pitch/F0 realisations of phrase-final tones do not always resemble their corresponding citation forms, although categorically most of them are very similar. The divergence can be ascribed to the following two factors.

- Most phrase-final tones have two pitch/F0 variants because of their phonetic sensitivity to the pitch/F0 offset of preceding tones. For example, the marked forms generally occur when the preceding tones end in a higher pitch/F0 offset as discussed in Chapters 6 and 8.
- The unmarked forms of phrase-final tones generally resemble their corresponding citation forms. Nevertheless, the phrase-final tones 2, 7, and 8 tend to have lower pitch/F0 levels, which can reasonably be considered a consequence of the declination effect of utterance-final position.

Table 9-7. Pitch/F0 comparison among Zhangzhou citation and phrase-final tones

Tone	Citation	Phrase-final	
		Unmarked	Marked
1	[35]	[35]	[34]
2	[22]	[211]	[311]
3	[51]	[52]	*
4	[41]	[41]	[51]
5	[33]	[33]	[43]
6	[41]	[41]	[51]
7	[221]	[211]	[311]
8	[22]	[211]	[311]

9.2.2. Duration

The length realisations of phrase-final tones are not quite the same as those of their corresponding citation forms (Table 9-8), largely reflected in the following two aspects.

- The number of length contrasts is reduced to three (extra-long, long, and medium) in the phrase-final position from four (extra-long, long, medium, and short) in citation.
- The length realisations of most phrase-final tones are not exactly the same as their corresponding citation forms. For example, tones 2 and 8 have slightly shorter durations than their citation forms while tone 6 becomes slightly longer in phrase-final position. Nevertheless, tones 1 and 5 maintain their extra-long properties while tone 4 keeps a medium realisation across citation and phrase-final contexts.

Table 9-8. Length comparison among Zhangzhou citation and phrase-final tones

Tone	Citation	Phrase-final
1	extra long	extra long
2	extra long	long
3	medium	long
4	medium	medium
5	extra long	extra long
6	short	medium
7	long	medium (long)
8	extra long	long

9.2.3. Vowel quality

The vowels have similar realisations across citation and phrase-final contexts. The high vowels alternate between monophthongs in the unstopped tones and diphthongs in the stopped tones 6 and 7 while other vowels are consistently realised as monophthongs across different tones and different environments, as shown in Table 9-9. In this study, the high vowel diphthongisation is posited as induced by the non-realisation of obstruent codas in the utterance-final position.

Table 9-9. Vowel quality realisation comparison among Zhangzhou citation and phrase-final tones (C=citation; PF=phrase-final)

	/i/		/u/		/e/		/ɵ/		/ɛ/		/ɐ/		/ɔ/	
Tone	C	PF	C	PF	C	PF	C	PF	C	PF	C	PF	C	PF
1	[i]	[i]	[u]	[u]	[e]	[e]	[ɵ]	[ɵ]	[ɛ]	[ɛ]	[ɐ]	[ɐ]	[ɔ]	[ɔ]
2	[i]	[i]	[u]	[u]	[e]	[e]	[ɵ]	[ɵ]	[ɛ]	[ɛ]	[ɐ]	[ɐ]	[ɔ]	[ɔ]
3	[i]	[i]	[u]	[u]	[e]	[e]	[ɵ]	[ɵ]	[ɛ]	[ɛ]	[ɐ]	[ɐ]	[ɔ]	[ɔ]
4	[i]	[i]	[u]	[u]	[e]	[e]	[ɵ]	[ɵ]	[ɛ]	[ɛ]	[ɐ]	[ɐ]	[ɔ]	[ɔ]
5	[i]	[i]	[u]	[u]	[e]	[e]	[ɵ]	[ɵ]	[ɛ]	[ɛ]	[ɐ]	[ɐ]	[ɔ]	[ɔ]
6	[iɛ̯]	[iɛ̯]	[uɤ̯]	[uɤ̯]	*	*	*	*	*	*	[ɐ]	[ɐ]	[ɔ]	[ɔ]
7	[iɛ̯]	[iɛ̯]	[uɤ̯]	[uɤ̯]	*	*	*	*	*	*	[ɐ]	[ɐ]	[ɔ]	[ɔ]
8	[i]	[i]	[u]	[u]	[e]	[e]	[ɵ]	[ɵ]	[ɛ]	[ɛ]	[ɐ]	[ɐ]	*	*

9.2.4. Voice quality

Voice quality changes with respect to vowel quality, pitch and and tone in both monosyllabic and phrase-final envrionments. Three types of phonation—creaky, breathy, and modal—are observed in the two contexts, as indicated in Table 9-10, but three notable aspects deserve further attention.

- The voice quality realisations of high vowels are consistently similar across citation and phrase-final environments. They alternate between breathy voice in the unstopped tones to creaky voice in the stopped tones 6 and 7.

- The voice quality realisations of mid vowels are basically similar across these two contexts. They tend to be produced with modal phonation but occasionally are produced with creaky voice in the falling pitch/F0 contours, especially among male speakers.
- The low vowels have slight differences in voice quality realisations in the phrase-final context. For example, they are largely produced with a creaky voice in the phrase-final tones 2 and 8 but with a modal phonation in their corresponding citation forms, which can be ascribed to the pitch/F0 declination effect phrase finally.

Table 9-10. Voice quality comparison among Zhangzhou citation and phrase-final tones

	High vowel		Mid vowel		Low vowel	
Tone	Citation	P-final	Citation	P-final	Citation	P-final
1	breathy	breathy	modal	modal	modal	modal
2	breathy	breathy	modal	modal	modal	creaky
3	breathy	breathy	modal/creaky	modal/creaky	creaky	creaky
4	breathy	breathy	modal/creaky	modal/creaky	creaky	creaky
5	breathy	breathy	modal	modal	modal	modal
6	creaky	creaky	*	*	creaky	creaky
7	creaky	creaky	*	*	creaky	creaky
8	breathy	breathy	modal	modal/creaky	modal	creaky

9.2.5. Syllable coda

The obstruent codas tend not to be realised in both monosyllabic and phrase-final settings, as indicated in Table 9-11. Correspondingly, the non-realisation of obstruent codas gives rise to several similar phonetic effects on the whole syllables in these two contexts, as shown in Table 9-12, including syllable lengthening, pitch/F0 contour depression, high vowel diphthongisation, and vowel laryngealisation.

Table 9-11. Obstruent coda realisation comparison with respect to Zhangzhou citation and phrase-final tones

Tone	Sonorant coda		Obstruent coda	
	Citation	Phrase-final	Citation	Phrase-final
1	+	+	-	-
2	+	+	-	-
3	+	+	-	-
4	+	+	-	-
5	+	+	-	-
6	-	-	not realised	not realised
7	-	-	not realised	not realised
8	+	+	-	-

Table 9-12. Comparison of phonetic effects induced by obstruent coda realisations among Zhangzhou citation and phrase-final tones

Parameter	Citation	Phrase-final
Obstruent coda	not realised	not realised
Duration	lengthened	lengthened
Pitch/F0 contour	depressed	depressed
Vowel quality	high vowel diphthongisation	high vowel diphthongisation
Voice quality	creaky	creaky

9.2.6. Summary

As discussed, each tone is made up of a complex of phonetic parameters in both monosyllabic and phrase-final contexts. In general, the tonal realisations are categorically similar across the two environments but are slightly different in phonetic details, especially in length and pitch/F0 realisations. The phonetic variations are largely predictable because of the declination effects of utterance-final position or phonetic sensitivity to preceding tones (Table 9-12). Exploring how the phrase-final tones are mapped onto the citation tones and discovering to what extent they are associated with each other is crucial to understanding the nature of Zhangzhou tones. The SplitsTree software was also applied to visualise the phonological relatedness among the citation and phrase-final tones on the basis of the multidimensional framework described.

The multidimensional properties of 16 tones including 8 citation and 8 phrase-final tones were all firstly transformed into a multiple sequence alignment to be computed in the SplitsTree software to generate a phylogram, as shown in the Figure 9-2. The terminal nodes on the right indicate the 16 tones being investigated. The horizontal lines indicate the amount of phonetic similarity shared by these tones. The vertical lines indicate the relatedness between tones. The longer the lines, the higher the probability that the tones are not related phonologically.

The mapping result between Zhangzhou citation tones and phrase-final tones is shown in Figure 9-2, using the SplitsTree software. Several aspects need to be further noted.

- Most of phrase-final tones are clustered with their corresponding citations, indicating they are categorically related. For example, citation tone 1 (C1) and phrase-final tone 1 (F1) are clustered together, so are C3 and F3, C4 and F4, and so on.
- The related pairs can have slight differences in their phonetic realisations. For example, C1 and F1 essentially have the same realisations, but C3 and F3 have differences in length realisations, reflecting the gradual continuous relatedness in the tonal mappings.
- Phrase-final tones 2 (F2) and 8 (F8) appear unrelated to their corresponding citation forms (C2 and C8, respectively), primarily because of the double effects of phonetic sensitivity to preceding pitch/F0 contours and declination of utterance-final position giving rise to variations in the realisations of pitch/F0, length, and voice quality. Nevertheless, these variations are articulatorily understandable and predictable. Thus, they should be considered as allophonic variants in their categories.

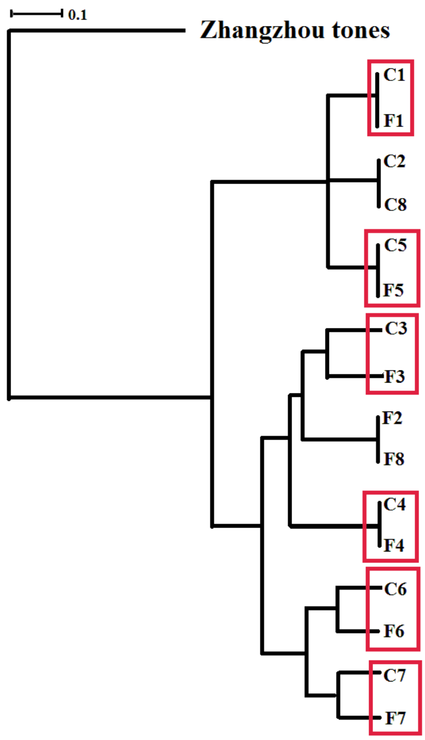


Figure 9-2. Mapping between Zhangzhou citation and phrase-final tones (C = citation tones; F= phrase-final tones).

9.3. Morphological Nature of Zhangzhou Tones

Combining the results of the mappings of the relatedness between the phrase-initial and citation tones (Figure 9-1) and between the phrase-final and citation tones (Figure 9-2), it is crucial to explore and explain how different forms of tonal realisation are structured abstractly.

The different phonetic forms of Zhangzhou tones across linguistic contexts are considered to be morphologically structured in speaker's mental grammar for the following three reasons.

- Because the realisations of phrase-initial tones are largely unrelated to those of their corresponding citation tones at either the phonological or the phonetic level (see Section 9.1.6), it is appropriate to consider the phrase-initial tones and citation tones as two independent systems.
- Because the realisations of most phrase-final tones appear to be categorically related to those of their corresponding citation tones with a certain degree of phonetically predictable variations (see Section 9.2.6), it is appropriate to consider the phrase-final tones and citation tones as belonging to the same system.
- Because tonal contrast neutralisation occurs across citation (see Section 5.1.7), phrase-initial (see Section 7.2.5), and phrase-final environments (see Section 8.2.5), the direction of tonal alternation is essentially indeterminate. Determining which forms are underlying and which forms are derived is difficult.

As a whole, it is appropriate to consider that the nature of Zhangzhou tones is morphological. Each lexical tone in Zhangzhou functions as a single morpheme, as indicated in Figure 9-3. When morphemes combine in a given domain, they may influence each other's sound structures, giving rise to one or more phonetically distant allomorphs for the same morphemes (Haspelmath & Sims, 2002; Lieber, 2009; Aronoff & Fudeman, 2011; Booij, 2012). In this study, each tonal morpheme is observed having two alternating allomorphs (tonemes) that are independently stored in the mental grammar of native speakers and phonetically distinct on the surface.

The distributions of alternating tonemes for a given lexical tonal morpheme in Zhangzhou are complementary and primarily constrained by the boundaries of syntactic phrases. One toneme is generalised for variants occurring in non-phrase-final position, called the *sandhi position*, while another toneme is generalised for variants occurring in non-sandhi positions, including citation and phrase-final contexts, called *non-sandhi positions*. In the non-phrase-final (sandhi) position, morphophonological alternation occurs and results in dramatic changes in the phonetic shapes of related tones, as illustrated in Figure 9-4.

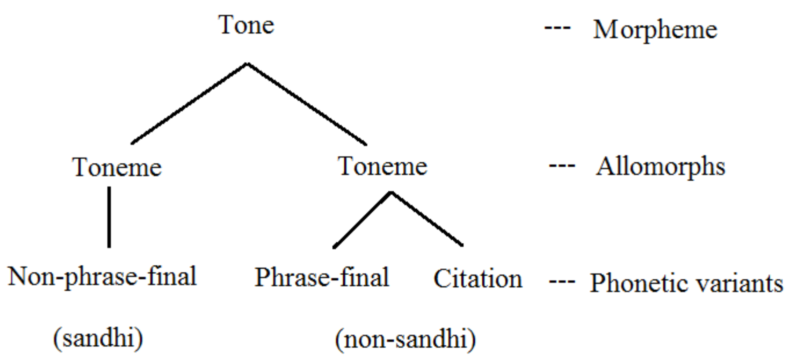


Figure 9-3. Morphological nature of Zhangzhou tones.

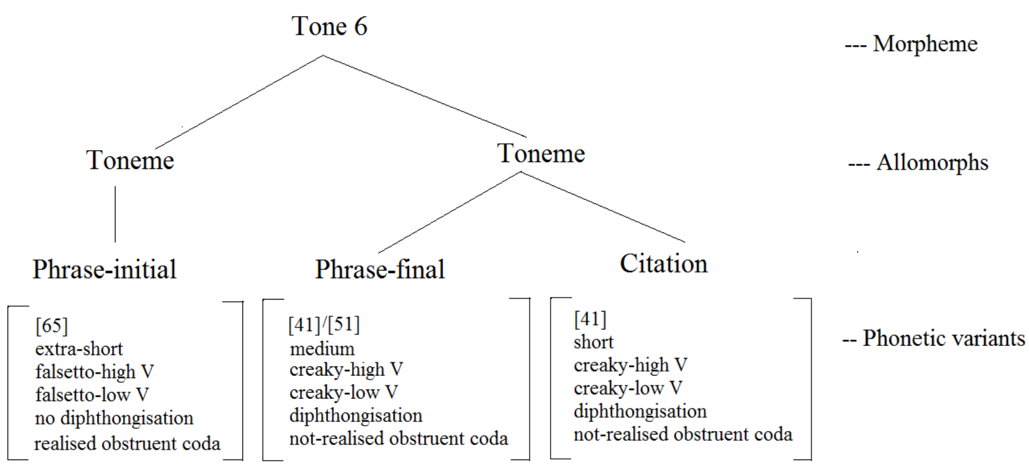


Figure 9-4. Morphological nature of tone 6 in Zhangzhou.

This morphological nature of Zhangzhou tones is analogous to the nature of grammatical morpheme in English. For example, when the plural morpheme is attached to other morphemes to perform its grammatical function, it is found having different phonetic forms conditioned by the manner of articulation of the final segment of its preceding morphemes (Haspelmath & Sims, 2002). As shown in Figure 9-5, it is pronounced as [s] after a voiceless non-sibilant sound (as in cats [kæts]), as [z] after a voiced non-sibilant sound (as in dogs [dɒgz]), and as [əz] after a sibilant of either voiceless or voiced (as in faces [feisəz]). These alternants for the plural morpheme have the same meaning but occur in different environments and in complementary distribution (Haspelmath & Sims, 2002, p. 27).

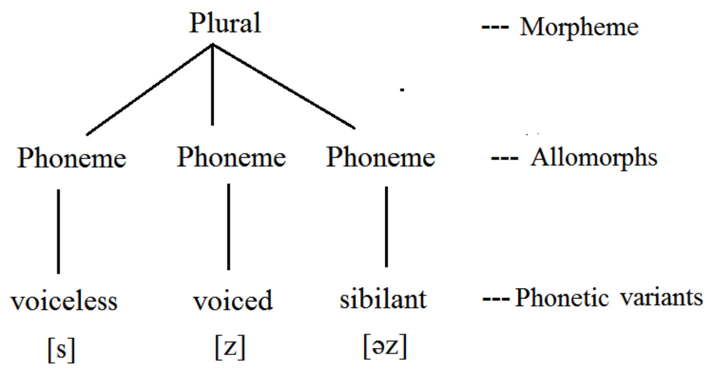


Figure 9-5. Morphological nature of plural morpheme in English.

As indicated, tonal morphemes in Zhangzhou essentially function very similar to the grammatical morphemes in English. They are attached to other morphemes to fulfil their linguistic function of either lexical, grammatical or both. They are found having alternating allomorphs of either tonemes or phonemes that are phonologically independent and phonetically different. The distributions of their alternating allomorphs are complementary and predictable, but are conditioned by different linguistic factors. The allomorph selection for Zhangzhou tones in this study is syntactically conditioned by the boundary of syntactic phrases, while the allomorph selection for English grammatical morphemes can be phonologically constrained, for example, by the manner of articulation of preceding final segments in the plural case.

As discussed, the morphological treatment of Zhangzhou tones is appropriate and plausible, reflecting complex interfaces among phonetics, phonology, morphology, and syntax. The relation between phrase-initial, phrase-final, and citation tones is morphophonemic. The distribution of sandhi tones is morphosyntactically determined. The relation between phrase-final and citation tones is allophonic. The forms of phrase-initial and citation tones are phonetically distant. The forms of phrase-final and citation tones are phonetically similar but present predictable variations. The morphological nature of Zhangzhou tones supports one assertion made by Hyman & Leben (2000, p. 588) that:

> [T]onal morphology is particularly interesting because it exhibits essentially the same range of morphological properties as in all of segmental morphology.

9.4. Previous Studies

No theoretical work has explored and explained the nature of Zhangzhou tones, and the relatedness between citation and disyllabic tones in the literature (see Section 6.4).

Nevertheless, numerous theoretical analyses have been conducted on other Southern Min varieties of Xiamen and Taiwan (see Section 1.6). As reviewed, proposing a unified and adequate account within generative phonology (see Section 1.6.3.1) and optimality theory (see Section 1.6.3.2) for the nature of Southern Min tone sandhi has been considered theoretically challenging and phonetically unnatural.

The unsatisfactory phonological interpretations have caused researchers to question the psychological reality of Southern Min tone sandhi rule and to seek for explanation from other perspectives. For example, a considerable number of production experiments have been conducted, the results of which largely falsify the full productivity of tone sandhi rule in Taiwanese, and favour a lexicon-based interpretation of allomorph selection for the relation between Southern Min sandhi and citation tones (see Section 1.6.3.3).

The result of this chapter (see Section 9.1) also does not support the phonological nature of tone sandhi in Southern Min, because the phrase-initial tones were not related with corresponding citation tones at either the phonological or the phonetic level in Zhangzhou. Thus, there seems impossible to formulate phonological rules or constraints to explain a derivational relation between citation and sandhi tones. This study asserts the nature of Zhangzhou tones is morphological (see Section 9.3). Sandhi (phrase-initial) and non-sandhi (citation and phrase-final) tones function as two alternating allomorphs that are phonologically independent and phonetically distant. The relation between sandhi and citation tones is morphophonemic.

The treatment of allomorph status of Zhangzhou tone sandhi is similar to the lexicon-based interpretation of allomorph selection in Taiwanese tone sandhi (Hsieh, 1976; Wang, 1992; Tsay & Myers, 1996; Chen et al., 2010); however, the two streams have different phonetic bases mainly reflected in the usage of methodologies and research corpus, and the claim with respect to the existence of phonological rules, as compared below.

First, the proposal of Zhangzhou tone sandhi was based on auditory and acoustic phonetics to map the relatedness of tonal realisations in phrase-initial, phrase-final, and citation contexts. The proposals of Taiwanese tone sandhi were based on production experiment to test the productivity of the tone sandhi rule, except the one from Tsay & Myers (1996) that was made by summarising results of previous studies from multiple sources with respect to different arguments.

Second, the research corpus of Zhangzhou tone sandhi involved 8 citation tones and 64 disyllabic combinations from 21 native speakers living in urban areas of Zhangzhou, and examined multidimensional parameters for tonal realisations. The corpus of Taiwanese tone sandhi varied between scholars. Hsieh (1976) involved one bi-dialectal female speaker of coastal and inland Taiwanese and focused on two tones. Wang (1992) investigated seven citation tones from 22 native speakers of Hsin-chu Taiwanese. Tsay & Myers (1996) did not conduct experiment. Chen et al. (2010) focussed on five tones of sonorant-ending syllables. These studies dominantly focussed on monodimension of tonal pitch/F0.

Third, the proposal of Zhangzhou tone sandhi asserted that the relation between sandhi and citation tones is morphophonemic, and there appears impossible to formulate phonological rules to explain a derivational relation between them. The claims of Taiwanese tone sandhi varied among scholars. Hsieh (1976) claimed that the Taiwanese sandhi tone rules are unreal. Wang (1992) asserted that some of Taiwanese tone sandhi rules are real, but the rules are organisational rather than derivational. Tsay & Myers (1996) claimed no rule for Taiwanese tone sandhi, but this claimed was made by summarising previous work rather than basing on observations of empirical experiment of their own. Chen et al. (2010) asserted that the nature of Taiwanese tone sandhi is not grammar-governed, although not all the tonal production errors are allomorph errors.

Thus, the morphological interpretation of this study is proposed on the basis of a solid empirical foundation, which further deepens the understanding of the nature of tone sandhi as a phenomenon in Southern Min, while filling the research gap in relation to the Zhangzhou variety.

9.5. Conclusion

This chapter has investigated the nature of the mappings between Zhangzhou citation and phrase-initial and phrase-final tones. It has shown the phrase-initial tones are largely not related to their corresponding citation tones at both the phonological and the phonetic levels while the phrase-final tones are largely categorically related with citation tones but present a certain degree of variation because of predictable sensitivity to the surrounding phonetic environment and to the effect induced by the utterance-final postion. Considering the morphological nature of Zhangzhou tones is thus appropriate.

In summary, the tonal realisations in Zhangzhou are multidimensional, involving a variety of segmental and suprasegmental parameters. The tonal interactions in Zhangzhou reflect interfaces among different linguistic levels, including phonetics, phonology, syntax, and morphology. Tonal neutralisation occurs across different linguistic contexts, including citation, phrase-initial and phrase-final settings. Therefore, investigating the nature of Zhangzhou tones must go beyond the single parameter of pitch/F0 and beyond the single linguistic context of citation.

Part E: From Now On

Chapter 10: Pitch and Beyond

The aim of this thesis was to explore and explain the nature of Zhangzhou tones with respect to five major research questions. (1) How many tones are there in Zhangzhou? (2) What are tones realised in Zhangzhou? (3) How do segmental and suprasegmental parameters interact to shape the tonal realisations and distinctions in Zhangzhou? (4) How do tones interact with each other in constructions beyond monosyllables? (5) How are the disyllabic tones, especially sandhi tones, related to the citation tones?

Several original findings have been observed and discussed based on analyses of the data from 21 speakers by drawing various intersecting approaches, including field linguistics; auditory, acoustic, and articulatory phonetics; and statistics. This final chapter primarily reviews the research findings of this thesis and discusses directions for future studies of Zhangzhou tones.

10.1. Linguistically Significant Findings

This section primarily summarises the major findings in each part of this thesis.

10.1.1. Part A: Preliminaries

Part A introduced the preliminaries before formal descriptions and discussions of Zhangzhou tones. Chapter 1 discussed the research background for this study in terms of tonology, tonetics, tone sandhi, previous studies of Zhangzhou, and previous studies of Southern Min tone sandhi. It primarily reviewed four issues. (1) Tonal realisations are complicated. It is necessary to investigate phonetic parameters besides pitch to understand the nature of tones. (2) Tonal representations are controversial among different models. Consensus has not yet been reached on some essential issues. For example, what is the nature of primitive features of tones? What is the tone-bearing unit? Are tonal features binary or privative? How are different features arranged? (3) Tonal sandhi is typologically diverse. Languages may differ in the motivation of tonal alternation, the dominancy of the sandhi pattern, and the specification of the sandhi domain. (4) Southern Min tone sandhi is theoretically attractive but challenging. Many studies have investigated the nature of SM tone sandhi, resulting in varying opinions concerning the directionality of SM tone sandhi, definition of the sandhi domain, and theoretical interpretation of the motivation of tone sandhi in Southern Min. This chapter also introduced the research goals, research opportunities and structure of this thesis.

Chapter 2 largely addressed two major issues. (1) How was the fieldwork conducted in Zhangzhou, including research area selection, participant selection, corpus design, and the data elicitation procedure? (2) How were the post-field data processed auditorily, acoustically, and statistically? Special attention was given to describing how the acoustic data were measured and normalised, as well as how the discontiguous F0 trajectories were resolved to extract satisfactory and reasonable values. The well-controlled data elicitation and processing largely guaranteed the quality and quantity of phonetic analyses in this study.

Chapter 3 discussed Zhangzhou syllables and segments. While Zhangzhou has a relatively small inventory of phonemes constituting different syllable components, their phonetic realisations are substantially diverse and dynamic. For example, a variety of processes motivated the diverse realisations of Zhangzhou onsets, including dentalisation, palatalisation, lenition, labialisation, laminalisation, nasalisation, and glottalisation. Both vowel quality and voice quality in Zhangzhou can change in certain tonal environments. The obstruent codas tend not to be realised in utterance-final settings but are identifiable in non-utterance-final positions. The glottal stop is considered a laryngeal feature accompanying creaky rhymes rather than a phonemic coda.

10.1.2. Part B: Citation tones

Part B largely discussed the phonetic properties of Zhangzhou citation tones from the perspective of auditory and acoustic phonetics (Chapters 4 and 5, respectively). Three major findings were observed.

First, tonal realisations are multidimensional. The single parameter of pitch/F0 is not sufficient to characterise the contrasts among Zhangzhou citation tones. Instead, a variety of phonetic parameters interact to create tonal realisations and distinctions. For example, tones 4 and 6 share a common mid-high falling pitch/F0, but they differ in other parameters. Tone 4 results in medium-length, breathy high vowels, creaky low vowels, and sonorant codas while tone 6 results in short-length, creaky high and low vowels, diphthongisation, and unrealised obstruent codas. The correlation between segmental and suprasegmental parameters for tonal productions has articulatory motivations.

Second, a considerable number of speaker-dependent variations exist in both pitch/F0 and length realisations of Zhangzhou citation tones. Thus, a normalisation process was necessary to derive a linguistic-phonetic representation of Zhangzhou tones as an independent variety, making this variety comparable to other languages/varieties with respect to language-internal properties.

Third, tonal contrasts can be neutralised in the monosyllabic setting. For instance, the realisations of citation tones 2 and 8 were found to be similar at either the phonological or the phonetic level. Thus, identifying the number of tonal contrasts based simply on tonal realisations in the citation environment is not sufficient. Instead, examining tonal realisations across different linguistic contexts beyond monosyllables is imperative for understanding the nature of tone.

10.1.3. Part C: Disyllabic tones

Part C concerned the properties of Zhangzhou disyllabic tones. Chapter 6 discussed the general properties of Zhangzhou disyllabic tones from the perspective of auditory phonetics while Chapters 7 and 8 explored the acoustic properties of Zhangzhou phrase-initial and phrase-final tones, respectively, supplemented by statistical testing and articulatory explanations. Several results are generalised as follows.

First, Zhangzhou has a right-dominant tone sandhi system. The number of tonal contrasts is largely preserved at the right-dominant position while the number is considerably reduced in the non-right-dominant positions. For example, the length contrasts are reduced dramatically to only two: medium or extra short. Tonal realisations are significantly changed. For example, the pitch/F0 contour of each tone is alternated to the form compeletely different from its citation form.

Second, Zhangzhou tone sandhi is syntactically relevant. The tone sandhi domain in Zhangzhou is not phonologically determined, but rather is aligned with a syntactic phrase XP. Within a given XP, the realisations of the tones in non-phrase-final positions undergo alternations phonologically and phonetically. Nevertheless, the alterations are sensitive only to the phrase boundaries and are not affected by the internal structures of syntactic phrases because the realisations appear similar regardless of whether the phrases are adjectival, verbal, nominal, or adverbial.

Third, Zhangzhou tone sandhi is phonologically inert. The realisations of Zhangzhou tones in disyllabic phrases are not categorically affected by their surrounding tones. For example, they have consistently similar realisations regardless of whether surrounding tones have rising, falling, or level pitch/F0 contours. The phonologically inert property of Zhangzhou tone sandhi is statistically supported by the pairwise *t* tests across 64 tonal combinations from 21 speakers.

Fourth, Zhangzhou tone sandhi is phonetically sensitive. While the realisations of Zhangzhou tones in multisyllabic phrases are not categorically affected by their surrounding tones, the realisations, especially the pitch/F0 realisations of phrase-final tones, are essentially sensitive to the surrounding environments and presents diverse outputs. For example, phrase-initial tones tend to have a higher pitch/F0 range while phrase-final tones tend to have a lower pitch/F0 range. Most of phrase-final tones have two variants which are statistically-significantly different.

Fifth, the realisations of Zhangzhou disyllabic tones are also multidimensional, involving complicated but systematic interactions among a variety of phonetic parameters. For example, the phrase-initial tone 7 indicates a mid-falling pitch/F0 contour, extra-short length, breathy high vowels, creaky low vowels, and realised obstruent codas while the phrase-final tone 7 indicates a low falling pitch/F0 contour, medium length, creaky vowels, and nonrealised obstruent codas. Nevertheless, the phrase-initial tones have several realisations different from both citation and phrase-final tones. For example, the obstruent codas are identifiable in this context, giving rise to extra-short length for the stopped tones and extra-high pitch/F0 and falsetto voice for the vowels in the phrase-initial stopped tone 6.

10.1.4. Part D: Mapping between citation and disyllabic tones

Part D primarily concerned how the disyllabic tones are related to the citation tones in Zhangzhou and how native speakers structure different realisations across different contexts in their mental grammars. Three major findings resulted from Chapter 9.

First, the phrase-initial tones are largely not related with their corresponding citation tones at both the phonological and the phonetic levels with respect to the multidimensional parameters. Therefore, considering them two independent systems is appropriate.

Second, the phrase-final tones are largely related to their corresponding citation tones at the phonological level but phonetically are not exactly the same because of predictable sensitivity to the surrounding phonetic environments. Thus, considering them to be in an allophonic relationship is appropriate.

Third, the nature of Zhangzhou tones is morphological. The structures of tonal realisations across different linguistic contexts are morphologically conditioned. Each tone can be regarded as a single morpheme having two alternating allomorphs (tonemes) that are both abstractly stored in the mental grammar of native speakers but are phonetically distant on the surface. One allomorph is for the variants occurring in the sandhi (non-phrase-final) position while another allomorph is for the variants occurring in the non-sandhi positions, including citation and phrase-final contexts.

10.1.5. Summaries

In summary, Zhangzhou has eight tones rather than seven, as proposed in previous studies. This finding resulted from examining the realisations of diverse parameters across three different contexts—isolation, phrase-initial, and phrase-final—, rather than classifying tones in terms of the preservation of Middle Chinese tonal categories in spontaneous speech, as largely was done in earlier studies.

Tonal contrasts in Zhangzhou can be neutralised across different linguistic contexts. Identifying the number of tonal contrasts based simply on tonal realisations in the citation environment is not sufficient. Instead, examining tonal realisations across different linguistic contexts beyond monosyllables is imperative for understanding the nature of tone.

Tonal realisations in Zhangzhou are multidimensional, involving a variety of segmental and suprasegmental parameters that interact with each other in a complicated but systematic way to code Zhangzhou tonal distinctions. Therefore, the single dimension of pitch/F0 is not sufficient to characterise Zhangzhou tonal contrasts.

Tonal interactions in Zhangzhou are complicated, involving phonetics, phonology, syntax, and morphology. Investigating the nature of tone sandhi in a given tonal language should not merely focus on how tones are realised in multisyllabic settings but should examine how the tone sandhi domain is linguistically specified; whether the tonal realisations may be affected by their surrounding tones and, if so, to what extent they are affected; how the sandhi tones are mapped onto the citation tones; and how native speakers structure different tonal realisations across different linguistic contexts. Thus, researchers must go beyond pitch to properly understand tone as a phenomenon in Southern Min.

10.2. Recommendation for Future Research

This study was a systematic investigation into the nature of Zhangzhou tones across multidimensional parameters, different phonetic branches, and different linguistic contexts, using field data from 21 speakers. The findings also highlight a number of new areas for exploration in future studies. The directions largely include: (1) studying Zhangzhou tones in connected speech, (2) exploring special tone sandhi patterns, (3) using articulatory and perceptual experiments to investigate tonal realisations in Zhangzhou, and (4) exploring the nature of Zhangzhou tones from diachronic and socio-phonetic perspectives of view, as discussed in this section.

10.2.1. Tones in connected speech

The data presented in this thesis were all well controlled to balance the perturbative effects from tautosyllabic segments for the purpose of acoustic experiment. In addition, both monosyllabic and disyllabic tokens were elicited by having them produced in isolation by native speakers. However, it will be of linguistically significantly interesting to investigate how tones are realised and how individual tones interact in constructions beyond disyllables. Additionally, tonal realisations can be considerably affected when produced in connected speech, in which a sentence generally "carries many other communicative functions that either use pitch as their primary or secondary carrier or change pitch inadvertently" (Xu, 2006, p. 743). For example, a question sentence generally has an overall rising intonation. Xu (2006) summarises three substantial discourse-induced effects on tone.

- The ideal form of tone is routinely compromised due to articulatory sluggishness.
- The underlying tonal target itself is sometimes changed.
- The surface tonal form is further distorted by other linguistic functions that also change pitch either deliberately or inadvertently.

Therefore, comprehending the discourse-induced modifying effects on tonal realisations is important for a full understanding of the nature of tone. It will be of linguistic significance for future research to investigate how tones are realised and how they interact in connected speech in

Zhangzhou, how a sequence of tones is articulated in succession, and how tonal realisations are different from those revealed in this study.

10.2.2. Special tone sandhi

The tone sandhi in Zhangzhou was found to be characteristically sensitive to the boundaries of syntactic phrases. In non-phrase-final positions, tonal alternations occur, which is considered regular tone sandhi in this present study. In addition, the tonal realisations in disyllabic constructions categorically appear not affected by internal structures of syntactic phrases and not affected by the categories of their surrounding tones. Therefore, the mapping between sandhi tones and citation tones is morphologically conditioned. Each tone has an allomorph (toneme) for variants occurring in non-phrase-final positions and another allomorph for variants occurring in utterance-final positions.

Nevertheless, two other tone sandhi patterns were also discovered in the field, occurring in very specific morphosyntactic contexts but showing dramatically different realisations from those identified in the regular tone sandhi. First, tones in Zhangzhou have different realisations in nominal compounds with a diminutive suffix /ʔɐ51/. For example, in this morphological environment, the pitch of the eight tones is either a high level [55] or a rising contour [35], depending on their pitch contour shape in citation. The falling pitch contours in citation are all neutralised to a high level [55] while the non-falling pitch contours are neutralised to a rising contour [35] in this specific context, as indicated in Table 10-1.

Table 10-1. Examples of tonal realisations before the diminutive morpheme /ʔɐ51/ in Zhangzhou

Tone	Citation	Regular sandhi	Before /ʔɐ51/	Example
1	[35]	[33]	[35]	ke35.ʔɐ51 ‘little chicken’
2	[22]	[33]	[35]	ɦi35.ʔɐ51 ‘little fish’
3	[51]	[25]	[55]	tsjɐw55.ʔɐ51 ‘little bird’
4	[41]	[63]	[55]	tʰɔ55.ʔɐ51 ‘little rabbit’
5	[33]	[32]	[35]	kʰi35.ʔɐ51 ‘little persimmon’
6	[41]	[65]	[55]	tit55.ʔɐ51 ‘small bamboo’
7	[221]	[32]	[35]	ɗɔk35.ʔɐ51 ‘little deer’
8	[22]	[32]	[35]	ʔjɵ35.ʔɐ51 ‘little tablet’

Second, tones in Zhangzhou have different realisations in the first adjectives of triplicated adjective constructions to express the highest degree of a particular quality. For example, in this specific morphosyntactic context, the pitches of the eight tones are all alternated to a mid-rising contour [35] with an extra-long duration, as shown in Table 10-2.

Table 10-2. Examples of tonal realisations of the adjectives in the first position of triplicated adjective constructions in Zhangzhou

Tone	Citation	Regular sandhi	First adjective	Adjective triplications
1	[35]	[33]	[35]	tsʰɛ35.tsʰɛ33.tsʰɛ35 ‘superb green’
2	[22]	[33]	[35]	ʔɐŋ35.ʔɐŋ33.ʔɐŋ211 ‘superb red’
3	[51]	[25]	[35]	suj35.suj35.suj52 ‘superb beautiful’
4	[41]	[63]	[35]	tsʰe35.tsʰe63.tsʰe41 ‘super crisp’
5	[33]	[32]	[35]	tsiŋ35.tsiŋ32.tsiŋ33 ‘superb quiet’
6	[41]	[65]	[35]	sip35.sip65.sip41 ‘superb humid’
7	[221]	[32]	[35]	tit35.tit32.tit211 ‘superb straight’
8	[22]	[32]	[35]	pɛ35.pɛ32.pɛ211 ‘superb white’

The special sandhi patterns were also broadly documented in the literature, but the descriptions were inconsistent among scholars (Ma, 1994; ZCCEC, 1999; Yang, 2006, 2008) and were different from my observations in the field. For example, in the adjective triplication context (Table 10-3), the pitch descriptions for the first adjective vary widely.

Table 10-3. Pitch description comparison of the first adjective in triplicated adjective constructions in Zhangzhou

Tone	Ma 1994	ZCCEC 1999	Yang 2008	This study
1	44	44	34	[35]
2	44	44	34	[35]
3	44	44	34	[35]
4	44	44	53	[35]
5	44	44	34	[35]
6	44	44	5	[35]
7	44	44	34	[35]
8	*	*	*	[35]

In summary, investigating how the individual tones are realised in these specific morphosyntactic environments with respect to the multidimensional parameters identified in this study will be interesting. In addition, exploring and explaining how these different realisations are related to citation tones and how native speakers structure such diverse realisations in their mental grammars will also be linguistically significantly interesting.

10.2.3. Other experimental methodologies

This study incorporated auditory description, acoustic quantification, articulatory explanation, and statistical testing to explore how tones are realised and how segmental and suprasegmental parameters interact to make tonal distinctions in Zhangzhou. Nevertheless, the research was limited to conducting acoustic quantification and statistical testing only on the parameters of F0 and duration, rather than on all parameters. For example, different types of voice quality, as part of tonal realisations in Zhangzhou, were examined primarily in terms of their acoustic manifestations on the waveforms and the spectral tilt H1-H2, supplemented by articulatory interpretation. Nevertheless, a variety of other parameters, including jitter, shimmer, harmonics-to-noise ratio (HNR), cepstral peak prominence (CPP), H1-A1 and H1-A2, can also be adopted for voice quality assessment (e.g., Andruski & Ratliff, 2000; Esposito, 2010, 2012; Abramson et al., 2015).

Additionally, electroglottograph (EGG), indexing the degree of contact between the vocal folds, has been considered a useful tool in measuring and describing vocal-fold phonatory vibration (Andruski & Ratliff, 2000; Esposito, 2010, 2012; DiCanio, 2009; Abramson et al., 2015). The perceptual validation test is also often adopted to examine how people perceive a contrast (e.g., Abramson, 1975; Esposito, 2006, 2010; Brunelle, 2009; Yu & Lam, 2014; Garellek et al., 2013; Kuo, 2013; Kuang, 2011, 2013). With multiple dimensions of contrasts, the perceptual test can also be used to verify how well listeners weight different parameters for tonal contrasts.

In summary, exploring the role of voice quality in tonal contrasts in Zhangzhou using other acoustic parameters would be of linguistic interest. The results concerning the correlation among vowel quality, voice quality, and pitch/F0 could be improved if articulatory and perceptual experiments were incorporated in future research. Additionally, conducting psycholinguistic studies on the lexical and tonal representation, using wug tests and priming techniques, would also help shed more lights on the nature of tone sandhi.

10.2.4. Other descriptive perspectives

This study focussed on synchronic speech to explore the nature of Zhangzhou tones. The derived results are not only of linguistically phonetic significance to represent Zhangzhou as a whole variety, making it comparable to other languages/varieties with respect to tonal properties. The findings are also of diachronic implication for future studies to trace the development of present-day Zhangzhou speech from the Archaic and Middle Chinese, and to map the shifts and mutations of sound change over centuries. From example, the emergence of tone 8 significantly reflects the split of MC Yangru tone that is motivated by the disappearance of historical glottal stop in the synchronic data.

A host of related questions may also arise for a fuller understanding of the sound system of Zhangzhou, both synchronically and diachronically. For example, how does the change of tonal system spread through a community? Does the change spread from morpheme to morpheme or does it affect all morphemes at the same time? Do speakers have simultaneous access to two grammatical system of tones? How does age or other social variation between speakers lead to the spread of change? Similarly, the widespread existence of nasalised vowels, syllabic nasals, and the three-way nasal coda contrasts in the segmental system also calls for diachronic attention to its historical origins and propagations.

Additionally, while this study concentrated to derive a linguistic phonetic representation of the variety of Zhangzhou as a whole, the acoustic signals essentially contain not only the invariable linguistic contents, but also the variable indexical contents reflecting individual speaker’s sociocultural background, pragmatic intent, vocal tract anatomy and physiology. It thus will be linguistically interesting to conduct a systematic investigation of the fundamental relationship between the variation of the tonal production and social background of speakers from articulatory, acoustic, and dialectological perspectives. For example, how do transgender speakers use individual parameters, for example F0, duration and voice quality, to index their gender identities? How are the phonetic/phonological forms of speech sound related to the social, regional, and interactional-communicative factors?

References

Abramson, A. S. (1975). Tones of Central Thai: Some perceptual experiments. *Studies in Thai linguistics in honor of William J. Gedney*, 1-16.

Abramson, A. S., & Erickson, D. M. (1978). Diachronic tone splits and voicing shifts in Thai: Some perceptual data. *Haskins Laboratories Status Report on Speech Research, 53*, 85-96.

Abramson, A. S., Tiede, M. K., & Luangthongkum, T. (2015). Voice register in Mon: Acoustics and electroglottography. *Phonetica, 72*, 237-256.

Anderson, S. R. (1978). Tone features. In V. Fromkin (Ed.), *Tone: A lingusitic survey* (pp. 133-175). New York, NY: Academic Press.

Andruski, J. E., & Ratliff, M. (2000). Phonation types in production of phonological tone: The case of Green Mong. *Journal of the International Phonetic Association, 30*, 37-61.

Aronoff, M. & Fudeman, Kirsten. (2011). *What is morphology?* (2nd edition). Hoboken, NJ: Wiley-Blackwell.

Avelino, H. (2010). Acoustic and electroglottographic analyses of nonpathological, nonmodal phonation. *Journal of voice*, 24(3), 270-280.

Baart, J. (2010). *A field manual of acoustic phonetics.* Dallas, TX: SIL International.

Ballard, W. L. (1988). *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.

Bao, Z. (1990). *On the nature of tone* (Doctoral dissertation, Massachusetts Institute of Technology). Retrieved from http:// dspace.mit.edu/handle/1721.1/14143

Bao, Z. (1999). *The structure of tone.* New York, NY: Oxford University Press.

Barrie, M. (2006). Tone circle and contrast preservation. *Linguistic Inquiry, 37*, 131-141.

Benedict, P. K. (1948). Tonal systems in Southeast Asia. *Journal of the American Oriental Society, 68*, 184-191.

Bickford, A. C., & Floyd, R. (2006). *Articulatory phonetics: Tools for analyzing the world's languages* (4th ed.). Dallas, Texas: SIL International.

Blevins, J. (1996). The syllable in phonological theory. In J. A. Goldsmith (Ed.), *The handbook of phonological theory*. Retrieved from Blackwell Reference Online website: http://www.blackwellreference.com/subscriber/book.html?id=g9780631201267_9780631201267

Blumstein, S. E. (1991). The relation between phonetics and phonology. *Phonetica, 48*, 108-119.

Boersma, P. (2001). Praat: Doing phonetics by computer [Computer program]. *Glot International, 5*, 341-345.

Booj, G. (2012). *The Grammar of words: An introduction to linguistic morphology* (3rd edition). Oxford, UK: Oxford University Press.

Bowern, C. (2008). *Linguistic fieldwork a practical guide.* London, England: Palgrave Macmillan.

Brunelle, M. (2009). Tone perception in Northern and Southern Vietnamese. *Journal of Phonetics, 37*, 79-96.

Butler, C. (1985). *Statistics in linguistics.* Oxford, England: Blackwell.

Cahill, M. (2008). *More universals of tone* (Working Paper 2007-007). Dallas, TX: SIL International Electronic Working Papers, 1-10.

Chan, M. K.-m. (1985). *Fuzhou phonology: A non-linear analysis of tone and stress* (Doctoral Dissertation, University of Washington). Retrieved from http:// sunzi.lib.hku.hk/ER/detail/hkul/3633038

Chang, M.-c. L. (1992). The representation of tone and the parametric variations of tonal systems. *Annual Meeting of the Berkeley Linguistics Society*, *18*, pp. 41-53.

Chao, Y. (1930). ə sɪstəm əv "toun-lɛtəz" (A system of "tone letters"). *Le Maître Phonétique, 45*, 24–27.

Chelliah, S. L., & de Reuse, W. J. (2010). *Handbook of Descriptive Linguistic Fieldwork.* Berlin, Germany: Springer.

Chen, B. (1990). The diachronic evolution and applicaiton of alveolar and velar nasal coda in Zhangzhou dialect (漳州前后鼻韵尾演变状况及运用). *Journal of Zhangzhou Teachers College (*漳州师范学院学报*)*, 96-101.

Chen, B. (1998). A comparative study between Zhangzhou southern Min and Taiwanese (漳台闽南方言比较研究). *Journal of Zhangzhou Teachers College* (漳州师院学报), *4*, 15-25.

Chen, M. (1992). Tone rule typology. *Proceedings of the Eighteenth Annual Meeting of the Berkeley Linguistics Society: Special session on the typology of tone languages* (pp. 54-66). eLanguage: The Linguistic Society of America's digital publishing platform.

Chen, M. (2000). *Tone sandhi.* Cambridge, England: Cambridge University Press.

Chen, M. Y. (1987). The syntax of Xiamen tone sandhi. *Phonological Yearbook*, *4*, 109-149.

Chen, S.-w., Myers, J., & Tsay, J. (2010). Testing the allomorph selection hypothesis in Taiwanese tone sandhi. *The 22nd North American Conference on Chinese Linguistics & The 8th International Conference on Chinese Linguistics*, (pp. 283-300).

Chen, Z. (2007). *Southern Min dictionary of Zhangzhou variety* 闽南漳州腔辞典. Beijing, China: Zhonghua Shuju 中华书局.

Cheng, R. L. (1968). Tone sandhi in Taiwanese. *Linguistics*, *6*(41), 19-42.

Chk2011. (2012). Administrative Division Zhangzhou. Retrieved from https://zh.wikipedia.org/wiki/File:Administrative_Division_Zhangzhou.png

Chomsky, N., & Harris, H. (1968). *The sound pattern of English.* New York, NY: Harper & Row.

Clark, J., & Yallop, C. (1990). *An introduction to phonetics and phonology.* Oxford, England: Blackwell.

Clements, G. N. (1981). A hierarchical model of tone. In W. Dressler, O. Pfeiffer, & J. Rennison (Eds.), *Phonologica 1980 (Innsbrucker Beiträge zur Sprachwissenschaft*, vol. 36), Vol. 2, pp. 69-75). Innsbruck, Austria: Institue für Sprachwissenschaft

Clements, G. N. (1983). The hierarchical representation of tone features. In I. Dihoff (Ed.), *Current approaches to African linguistics* (pp. 145-176). Dordrecht, The Netherlands: Foris.

Clements, G. N. (1985). The geometry of phonological features. *Phonology, 2*(1), 225-252.

Cohen, A., & Collier, R. (1982). Declination: Construct or intrinsic feature of speech pitch? *Phonetica*, *39*, 254-273.

Cohn, A. C. (2007). Phonetics in phonology and phonology in phonetics. *Working Papers of the Cornell Phonetics Laboratory*, *16*, pp. 1-31.

Connell, B. (2002). Tone languages and the universality of intrinsic F0: Evidence from Africa. *Journal of Phonetics*, *30*, 101-129.

Dedering, U. (2010). *Zhangzhou is located in China.* Retrieved from https://en.wikipedia.org/wiki/Zhangzhou#/media/File:China_edcp_location_map.svg

DiCanio, C. T. (2009). The phonetics of register in Takhian Thong Chong. *Journal of the International Phonetic Association, 39*, 162-188.

DiCanio, C. T. (2012). Coarticulation between tone and glottal consonants in Itunyoso Trique. *Journal of Phonetics*, *40*, 162-176.

Ding, P. S. (2016). *Southern Min (Hokkien) as a migrating language: a comparative study of language shift and maintenance across national borders.* New York, United States: Springer.

Dong, T. (1959). *Four Southern Min varieties* 四个闽南方言. Taipei: Zhongyang Yanjiuyuan 中央研究院.

Donohue, C. (2007). The interaction of tones and vowels in Fuzhou. *Annual Meeting of the Berkeley Linguistics Society, 33*, pp. 97-104.

Donohue, C. (2013). *Fuzhou tonal acoustics and tonology.* Muenchen, Germany: LINCOM Europa.

Donohue, M. (1997). Tone systems in New Guinea. *Linguistic Typology, 1*, 347-368.

Donohue, M. (2005). Tone and the trans New Guinea languages. *Crosslinguistic studies of tonal phenomena* (pp. 33-53). Tokyo, Japan: Tokyo University of Foreign Studies: Research Institute for Languages and Cultures of Asia and Africa.

Dörnyei, Z. (2007). *Research methods in applied linguistics : Quantitative, qualitative, and mixed methodologies.* Oxford, England: Oxford University Press.

Du, X. (2013). Study of sound system in Zhangzhou Dialect Vocabularies (《漳州方言词汇》音系研究). *Fangyan*, 03, 279-285.

Duanmu, S. (1990). *A formal study of syllable, tone, stress and domain in Chinese languages* (Doctoral dissertation, Massachusetts Institute of Technology). Retrieved from http://dspace.mit.edu/handle/1721.1/13646

Duanmu, S. (2000). *The Phonology of Standard Chinese.* Oxford, England: Oxford University Press.

Duanmu, S. (2005). The tone-syntax interface in Chinese: Some recent controversies. *Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena, Historical Development, Tone-Syntax Interface, and Descriptive Studies*, (pp. 16-17).

Earle, M. (1975). *An acoustic phonetic study of North Vietnamese tones.* Santa Barbara, CA: Speech Communication Research Laboratories.

Esposito, C. M. (2006). *The effects of linguistic experience on the perception of phonation* (Doctoral dissertation, University of California, Los Angeles). Retrieved from http://linguistics.ucla.edu/research/phd-dissertations/

Esposito, C. M. (2010). The effects of linguistic experience on the perception of phonation. *Journal of Phonetics, 38,* 306-316.

Esposito, C. M. (2012). An acoustic and electroglottographic study of White Hmong tone and phonation. *Journal of Phonetics, 40*, 466-476.

Faytak, M., & Yu, A. C. (2011). 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.

FCCEC. (1998). *Fujian Province Gazette-Dialect Volume* 福建省志-方言志. Beijing, China: Fangzhi Chubanshe 方志出版社.

Foulkes, P., Scobbie, J. M., & Watt, D. (2010). 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.

Fox, A. (2000). *Prosodic features and prosodic structure: The phonology of suprasegmentals.* New York, NY: Oxford University Press.

Francis, A. L., Ciocca, V., Wong, V. K., & Chan, J. K. (2006). Is fundamental frequency a cue to aspiration in initial stops? *Journal of the Acoustical Society of America, 12*, 2884-2895.

Fromkin, V. A. (1972). Tone features and tone rules. *Studies in African Linguistics, 3,* 47-76.

Fu, B. (1995). *A system of tone features and its implications for the representation of tone* (Doctoral dissertation, Simon Fraser University). Retrieved from http:// summit.sfu.ca/item/6860

Gandour, J. (1977). On the interaction between tone and vowel length: Evidence from Thai dialects. *Phonetica, 34*, 54-65.

Gandour, J. T. (1978). The perception of tone. In V. A. Fromkin (Ed.), *Tone: A linguistic survey* (pp. 41-76). New York, NY: Academic Press.

Gao, R. (1999). Introduction to the sound system of Zhangzhou 漳州方言音系略说. In *Minnan dialect-studies of Zhangzhou variety* 闽南方言-漳州话研究 (pp. 109-116). Beijing, China: Zhongguo Wenlian Chubanshe 中国文联出版社.

Garellek, M., & Keating, P. (2011). The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association, 41*, 185-205.

Garellek, M., Keating, P., Esposito, C. M., & Kreiman, J. (2013). Voice quality and tone identification in White Hmong. *Journal of the Acoustical Society of America, 133,* 1078-1089.

Gick, B., Wilson, I., & Derrick, D. (2013). *Articulatory phonetics.* Hoboken, NJ: Wiley-Blackwell.

Goldsmith, J. (1976). *Autosegmental phonology* (Doctoral dissertation, Massachusetts Institute of Technology). Retrieved from http://hdl.handle.net/1721.1/16388

Goldsmith, J. A., Hume, E., & Wetzels, L. (2011). *Tones and features: Phonetic and phonological perspectives*. Berlin, Germany: De Gruyter Mouton.

Gordon, M., & Ladefoged, P. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics, 29*, 383-406.

Guo, J. (1999). Sound change of Zhangzhou variety during the past one hundred years (近一百年来漳州话语音变化). In B. Chen (Ed.), *Southern Min dialect- studies of Zhangzhou variety* (闽南方言一漳州话研究; pp. 117-126). Beijing, China: Zhongguo Wenlian Chubanshe (中国文联出版社).

Guo, J. (2014). *Zhangzhou Southern Min* 漳州闽南方言. Zhangzhou, China: Zhangzhou Library 漳州图书馆.

Halle, M., & Stevens, K. N. (1971). 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 1954 - 2002.* Berlin, Germany, de Gruyter

Harrington, J. (2010). 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.

Hashimoto, M. J. (1982). The so-called ‘original’ and ‘changed’ tones in Fukienese—A case study of Chinese tone morphophonemics. *Bulletin of the Institute of History and Philology, 53*, 645-659.

Haspelmath, M., & Sims, A.D (2002). *Understanding morphology* (2rd edition.). London, England: Arnold.

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

Hombert, J.-M. (1976). Effect of aspiration on the fundamental frequency of the following vowel. *Proceedings of the 2nd Annual Meeting of the Berkeley Linguistics*, (pp. 212-219).

Hombert, J.-M. (1978). Consonant types, vowel quality, and tone. In V. A. Fromkin (Ed.), *Tone: A linguistic survey* (pp. 77-112). New York, NY: Academic Press.

Hombert, J.-M., Ohala, J., & Ewan, W. G. (1979). Phonetic explanations for the development of tones. *Language, 55*, 37-58.

Hoole, P., & Hu, F. (2004). Tone-vowel interaction in standard Chinese. *International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages*, (pp. 89-92).

Hooper, J. B. (1972). The syllable in phonological theory. *Language, 48*, 525-540.

Hsieh, F.-f. (2005). Tonal chain-shifts as anti-neutralization-induced tone sandhi. *University of Pennsylvania Working Papers in Linguistics, 11*(1), 99-112.

Hsieh, H.-I. (1976). On the unreality of some phonological rule. *Lingua, 38*(1), 1-19.

Huang, D. (1999). A research into the Archic Chinese in different Southern Min varieties (闽南方音中的上古音残余). In B. Chen (Ed.), *Minnan dialect—Studies of Zhangzhou variety* (闽南方言－漳州话研究; (pp. 52-76). Beijing, China: Zhongguo Wenlian Chubanshe (中国文联出版社).

Huang, L. (2014). Research on the evolution of the Chang-chew dialect over the past 100 years through Chinese-English dictionaries of the vernacular or spoken language of Amoy (从《厦英大辞典》看漳州方言一百多年来的变化). (Master's thesis, Huaqiao University). Retrieved from http://60.16.24.71/KCMS/detail/detail.aspx?filename=1014079515.nh&dbcode=CMFD&dbname=CMFDTEMP

Huang, Y. (2009). A phonetic study on the acquisition of Chinese vowels in Minnan area (闽南方言区普通话元音习得的实验研究). *Proceedings of The 9th International Conference on Chinese Teaching -Linguistic Analysis 第九届华语文教学研讨会论文集.* Taiwan: Sino-Culture Press (世界华文出版社).

Huang, Y. (2010). *An acoustic-comparative study between oral vowels and nasalised vowels in Minnan Zhangzhou dialect* (闽南漳州方言口元音和鼻化元音的声学比较研究). (Master's thesis, Guangxi Normal University) Retrieved from http://59.57.115.123:85/KCMS/detail/detail.aspx?filename=2010155514.nh&dbcode=CMFD&dbname=CMFD2011.

Huang, Y., Donohue, M., Sidwell, P., & Rose, P. (2016). 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 Sciience & Technology Association .

Huckvale, M. (2012). Data processing: Digital analysis of speech audio signals. In N. Müller & M. J. Ball (Eds.), *Research methods in clinical linguistics and phonetics: A practical guide* (pp. 195-218). Chichester, England: Wiley and Sons.

Huffman, M. K. (1987). Measures of phonation type in Hmong. *Journal of the Acoustical Society of America, 81*, 495-504.

Hyman, L. M. (1975). *Phonology: theory and analysis.* New York, NY: Holt, Rinehart & Winston.

Hyman, L. M. (1978). Tone and/or accent. In D. Napoli (Ed.), *Elements of tone, stress and intonation* (pp. 1-20). Washington, DC: Georgetown University Press.

Hyman, L. M. (1993). Register tones and tonal geometry. In K. Snider & H. v. Hulst (Eds.), *The phonology of tone: The representation of tonal register* (pp. 75-108). Berlin, Germany: Mouton de Gruyter.

Hyman, L. M. (2011). 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.

Hyman, L. M. (2012). Coda constraints on tone. *UC Berkeley Phonology Lab Annual Reports,* 189-204.

Hyman, L.M., & Leben, W. R. (2000). Suprasegmental processes. In G. Booij, C. Lehmann, J. Mugdan, & S. Skopeteas (Eds), *Morphologie/Morphology: Ein internationales handbuch zur flexion und wortbildung (An International Handbook on Inflection and Word-formation)* (pp. 587-594). Berlin, Germany: Walter De Gruyter & Go.

Hyman, L. M., & Schuh, R. G. (1974). Universals of tone rules: Evidence from West Arica. *Linguistic Inquiry, 5*, 81-115.

Hyman, L. M., & VanBik, K. (2002). Tone and syllable structure in Hakha-Lai. *Annual Meeting of the Berkeley Linguistics Society, 28*, 15-28.

Hyman, L. M. (2006). Word prosodic typology. *Phonology*, *23*, 225-257.

Iunn, U.-G., Lau, K.-G., Tan-Tenn, H.-G., Lee, S.-A., & Kao, C.-Y. (2007). Modeling Taiwanese Southern-Min tone sandhi using rule-based mothods. *International Journal of Computational Linguistics and Chinese Language Processing*, *12*, pp. 349-370.

Johnson, K. (2008). *Quantitative methods in linguistics.* Hoboken, NJ: Wiley-Blackwell.

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

Kloepper, T. H., & Huson, D. H. (2008). Drawing explicit phylogenetic networks and their integration into SplitsTree. *BMC Evolutionary Biology* (1), 22.

Kong, Q.-m. (1987). Influence of tones upon vowel duration in Cantonese. *Language and Speech, 30*, 387-399.

Kuang, J. (2011). Production and perception maps of the multidimensional register contrast in Yi. *UCLA Working Papers: Phonetics*, 1-30.

Kuang, J. (2013). *Phonation in tonal contrasts* (Doctoral dissertation, University of California: Los Angeles). Retrieved from http://escholarship.org/uc/item/6n72p16n

Kuo, C.-H. (2013). *Perception and acoustic correlates of the Taiwanese tone sandhi group* (Doctoral dissertation, University of California, Los Angeles). Retrieved from http://escholarship.org/uc/item/30q6w11t#page-2

Ladd, D. R. (1984). Declination: A review and some hypotheses. *Phonology 1*, 53-74.

Ladefoged, P. (1971). *Preliminaries to linguistic phonetics.* Chicago, IL: University of Chicago Press.

Ladefoged, P. (1999). Instrumental techniques for linguistic phonetic fieldwork. In T. H. Sciences, W. Hardcastle, J. Laver (Eds.), *Blackwell Reference Online*. Retrieved from http://www.blackwellreference.com/subscriber/tocnode.html?id=g9780631214786_chunk_g97806312147865

Ladefoged, P. (2003). *Phonetic data analysis: An introduction to fieldwork and instrumental techniques.* Hoboken, NJ: Wiley-Blackwell.

Ladefoged, P., & Disner, S. F. (2012). *Vowels and consonants* (3rd edition). Hoboken, NJ: Wiley-Blackwell.

Lai, Y., & Jongman, A. (2005). Effects of aspiration on fundamental frequency in Taiwanese syllables. *Journal of the Acoustical Society of America*, *117*, 2458-2459.

Lai, Y., Huff, C., Sereno, J., & Jongman, A. (2009). The raising effect of aspirated prevocalic consonants on F0 in Taiwanese. *Proceedings of the 2nd International Conference on East Asian Linguistics, Simon Fraser University Working Papers in Linguistics.*

Lass, R. (1984). *Phonology: An introduction to basic concepts*. Cambridge, England: Cambridge University Press.

Laver, J. (1980). *Phonetic description of voice quality.* Cambridge, England: Cambridge University Press.

Laver, J. (1994). *Principles of phonetics.* Cambridge, England: Cambridge University Press.

Lehiste, I. (1976). Suprasegmental features of speech. In N. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 225-239). New York, NY: Academic Press.

Levi, S. (2011). Glides. In *The Blackwell Companion to Phonology* (pp. 341-366). Hoboken, NJ: Wiley-Blackwell. doi:10.1002/9781444335262.wbctp0015

Levshina, N. (2015). *How to do Linguistics with R: data exploration and statistical analysis*. Amsterdam, Netherlands: John Benjamins Publishing Company.

Li, R., & Yao, R. (2008). *Southern Min dialects* (闽南方言). Fuzhou, China: Fujian Renmin Chubanshe (福建人民出版社).

Liang, N. (2014). *The Ph system comparison study of Southern Fukienese and Taiwanese* (闽台闽南方言音系比较研究). (Master's thesis, Fujian Normal University). Retrieved from http://cdmd.cnki.com.cn/Article/CDMD-10394-1015521194.htm

Liaon98. (2016). Hokkien distribution. Retrieved from https://en.wikipedia.org/wiki/Hokkien#/media/File:Hokkien_Map.svg

Lieber, R. (2010). *Introducing morphology*. Cambridge, England: Cambridge University Press.

Lieberman, P. (1967). *Intonation, perception and language.* Cambridge, MA: MIT Press.

Lin, B. (1992). Zhangzhou vocabularies 漳州方言词汇. *Fangyan 方言*, 1-3.

Lin, J.-w. (1994). Lexical government and tone group information. *Phonology*, *11*, 237-275.

Lorenz, F. (2013). *Basics of phonetics and english phonology* (2nd ed.). Berlin, Germany: Logos-Verlag.

Lotto, A. J., Holt, L. L., & Kluender, K. R. (1997). Effect of voice quality on perceived height of English vowels. *Phonetica, 54,* 76-93.

Ma, C. (1994). *Studies of Zhangzhou dialect* 漳州方言研究. Hongkong: Zongheng Chubanshe 纵横出版社.

Ma, C. (1997). A comparative study of the rhyme system between Guangyun and Zhangzhou dialect (1) (《广韵》与漳州方言韵系比较研究 (上)). *Journal of Fujian Teachers University* （福建师范大学学报-哲学社会科学版), *2*, 95-102.

Ma, C. (1997). A comparative study of the rhyme system between Guangyun and Zhangzhou dialect (2) (《广韵》与漳州方言韵系比较研究 (下)). *Journal of Fujian Teachers University* (福建师范大学学报-哲学社会科学版), *3*, 83-87.

Ma, C. (1999). A comparative study of rhyme systems between Guangyun and Zhangzhou dialect (《广韵》韵系与漳州方言韵系的比较研究). In B. Chen (Ed.), *Southern Min dialect—Studies of Zhangzhou variety* (闽南方言一漳州话研究, pp. 78-108). Beijing: Zhongguo Wenlian Chubanshe (中国文联出版社).

Ma, C. (2002). *Origins and evolutions of Fujian and Taiwan dialects (闽台方言的源流与嬗变).* Fuzhou, China: Fujian Renmin Chubanshe (福建人民出版社).

Ma, C. (2008). A comparative study of rhyming dictionaries of Fujian and Taiwan dialects 闽台闽南方言诸韵书音系比较研究. *Journal of Fujian Taiwan Cultural Research* (闽台文化研究), *4*, 103-116.

Ma, C. (2013). A comparative study of the phonological systems in four rhyming dictionaries of Zhangzhou variety in Fujian and Taiwan (闽台漳州腔四韵书音系比较研究-兼论台湾漳州腔韵书《增补汇音宝鉴》音系性质). *Fujian-Taiwan Culture Research (闽台文化研究)*.

Maddieson, I. (1990). The transcription of tone in the IPA. *Journal of the International Phonetic Association, 20*, 28-32.

Maddieson, I. (2013). Tone. In M. S. Dryer, & M. Haspelmath, *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved 05 01, 2017, from http://wals.info/chapter/13

Maeda, 1. (1976). *A characterisation of American English intonation.* Cambridge, MA: MIT Press.

Maran, L. (1971). Burmese and Jinghpaw: A study of tonal linguistic processes. *Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics, IV*.

Massaro, D. W., & Cohen, M. M. (1976). The contribution of fundamental frequency and voice onset time to the /zi/ – /si/distinction. *The Journal of the Acoustical Society of America, 60*, 704-717.

McCarthy, J. J. (1988). Feature geometry and dependency: A review. *Phonetica, 45*(2-4), 84-108.

Medhurst, W. H. (1838). *A dictionary of the Hok-këèn dialect of the Chinese language, according to the reading and colloquial idioms, Zhangpu* (福建方言字典). Macao, China: Honorable East India Company’s Press.

Mortensen, D. (2002). *Semper infidelis: Theoretical dimensions of tone sandhi chains in Jingpho and A-Hmao*. Unpublished manuscript, University of California, Berkeley.

Mortensen, D. R. (2013). Tonally conditioned vowel raising in Shuijingping. *Journal of East Asian Linguististics, 22*, 189-216.

Myers, J., & Tsay, J. (2003). A formal funcitonal model of tone. *Language and Linguistics, 4*, 105-138.

Odden, D. (2011). Features impinging on tone. In J. A. Goldsmith, E. Hume, & L. Wetzels (Eds.), *Tones and features. Phonetic and phonological perspectives* (pp. 81-107). Berlin, Germnay: De Gruyter Mouton.

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

Ohala, J. J. (2010). The relation between phonetics and phonology. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd ed., pp. 653-677). Hoboken, NJ: Wiley-Blackwell.

Ohala, J. J., & Ewan, W. G. (1972). Speed of pitch change. *Journal of the Acoustical Society of America, 53*, 345.

Pan, H.-h. (2005). Voice quality of falling tones in Taiwan Min. *Proceedings of Interspeech 2005*, Lisbon, Portugal, pp. 1401-1404.

Paterson III, H. J. (2015). Phonetic transcription of tone in the IPA. *Americas, 11*, 38.

Pierrehumbert, J. (1989). *A preliminary study of consequences of intonation for the voice source.* Stockholm, Sweden: Royal Institute of Technology, Speech Transmission Laboratory.

Pike, K. L. (1947). *Phonemics.* Ann Arbor: University of Michigan Press.

Pike, K. L. (1948). *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.

Ping, J.-K. (1995). *Fuzhou tone-vowel interaction*. Unpublished manuscripts, University of British Columbia, 89-113.

Rasinger, S. M. (2013). *Quantitative research in linguistics: An introduction.* London, England: Bloomsbury.

Ratliff, M. (2015). Tonoexodus, tonogenesis, and tone change. In P. Honeybone, & J. Salmons (Eds.), *The Oxford handbook of historical phonology* (pp. 245-261). Oxford, England: Oxford University Press.

Reetz, H., & Jongman, A. (2009). *Phonetics: Transcription, production, acoustics, and perception.* Hoboken, NJ: Wiley-Blackwell.

Rose, P. (1982). 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.

Rose, P. (1986). The normalization of tone. *Proceedings of the First Australian Conference on Speech Science and Technology,* Canberra (pp. 130-135).

Rose, P. (1987). Considerations in the normalisation of the fundamental frequency of linguistic tone. *Speech Communication*, 6 (4), 343-352.

Rose, P. (1988). On the non-equivalence of fundamental frequency and pitch in tonal description. In D. H. Bradley, & M. Mazaudon (Eds.), *Proceedings of Prosodic Analysis and Asian Linguistics: to Honour R.K.Sprigg* (pp. 55-82).Canberra, Australia: Pacific Linguistics.

Rose, P. (1990). Linguistic phonetic aspects of Shanghai tonal acoustics. In R. Seidl (Ed.), *Proceedings of the 3rd Australian International Conference on Speech Science and Technology* (pp. 388-393). Melbourne, Australia: Australian Speech Science and Technology Association.

Rose, P. (1993). A linguistic phonetic acoustic analysis of Shanghai tones. *Australian Journal of Linguistics,13*, 185-219.

Rose, P. (1994). Wenzhou tonal acoustic-depressor and register effects in Chinese tonology. In R. Togneri (Ed.), *Proceeding of the 5th Australian International Conference on Speech Science and Technology* (pp. 138-142). Perth, Australia: Australian Speech Science and Technology Association.

Rose, P. (1995). Acoustic analysis of disyllabic lexical tone sandhi in Zhenhai dialect (聲調音韻學研究 - 鎮海方言雙音節詞連讀變調的聲學語音學分析). In E. Zee (Ed.), *Studies of the Wu Dialects* 吳語研究 (pp. 239-260). New Asia Academic Bulletin 1. Chinese University of Hong Kong.

Rose, P. (1996). Between- and within-speaker variation in the fundamental frequency of Cantonese citation tones. In P. Davis, & N. Fletcher (Eds.), *Vocal fold physiology:*

Controlling complexity and chaos (pp. 307-324). Brighton, United Kingdom: Singular Press.

Rose, P. (2000). Hongkong Cantonese citation tones acoustics: a linguistic tonetic survey. In M. Barlow (Ed.), *Proceedings of the 8th Australian International Speech Science and Technology Conference* (pp. 198-203). Canberra: Australian Speech Science and Technology Association.

Rose, P. (2004). "Defying Explanation"? Accounting for Tones in Wenzhou dialect disyllabic lexical tone sandhi. In S. Cassidy (Ed.), *Proceedings of the 10th Australian International Conference on Speech Science and Technology*, (pp. 237-242). Canberra: Australian Speech Science and Technology Association.

Rose, P. (2011). Reverse-engineering tones and tone sandhi in Wenzhou dialect: Between-speaker differences and historical development (逆溯制造温州方言聲調和連讀變調: 兩個人差異與歷史發展). *International Bulletin of Chinese Linguistics, 5*, 29-48.

Rose, P. (2014). Transcribing tone—A likelihood-based quantitative evaluation of Chao’s tone letters. *Proceedings of Interspeech 2014*, (pp. 101-105). Singapore.

Rose, P. (2016). Comparing normalisation stratagies for citation tone F0 in three Chinese dialects. In C. Carignan & M. D. Tyler (Eds.), *Proceedings of the 16th Australasian International Conference on Speech Science and Technology*, (pp. 221-224). Sydney: Australian Speech Science and Technology Association.

Rose, P. (2016). 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.

Sagart, L. (1998). The origin of Chinese tones. *International Symposiumon on "Tonal Languages in the World"* (pp. 91-104). Tokyo, Japan: Institute for the Study of Languages and Cultures of Asia and Africa.

Samarin, W. J. (1967). *Field linguistics: A guide to linguistic field work.* New York, NY: Holt, Rinehart & Winston.

Schlegel, G. (1886). *Nederlansch-Chineesch-Woordenboek met de Trascriptie der Chineesche Karakters in het Tsiang-Tsiu Dialekt (荷华文语类参).* Leiden, The Netherlands: E.J. Brill.

Scholz, F., & Chen, Y. (2014). The independent effects of prosodic structure and information status on tonal coarticulation: Evidence from Wenzhou Chinese. In J. Caspers., Y. Chen., W. Heeren., J. Pacilly., N. Schiller & E. van Zanten (Eds). *Above and Beyond the Segments: Experimental Linguistics and Phonetics* (pp. 275-287). Amsterdam, Netherland: John Benjamins Publishing Company.

Schuh, R. G. (1978). Tone rules. In V. A. Fromkin (Ed.), *Tone: A linguistic survey* (pp. 221-256). New York, NY: Academic Press.

Scobbie, J. M. (2007). Interface and overlap in phonetics and phonology. In G. Ramchand & C. Reiss (Eds.), *The Oxford Handbook of Linguistic Interfaces.* Oxford, England: Oxford University Press

Shen, R., & Rose, P. (2016). Preservation of tone in right-dominant tone sandhi: A fragment of disyllabic tone sandhi in Maodian Wu Chinese. In C. Carignan & M. D. Tyler (Eds.), *Proceedings of the 16th Australasian International Conference on Speech Science and*

Technology, (pp. 345-348). Sydney: Australian Speech Science and Technology Association.

Shi, F. (1986). Tonal studies of disyllabic words in Tianjin dialect (天津方言双字组声调分析). *Studies in Languages and Linguistics 语言研究, X,* 77-90.

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

Steed, W. J. (2011). *Lishui Wu tone and tone sandhi—An acoustically-based description* (Doctoral dissertation: Australian National University). Retrieved from http://library.anu.edu.au/record=b2569834

Stevens, K. (1977). Physics of laryngeal behavior and larynx modes. *Phonetica, 34,* 264-279.

Stevens, K. N., & Klatt, D. H. (1974). Role of formant transitions in the voiced - voiceless distinction for stops. *Journal of the Acoustical Society of America, 55,* 653-659.

Styler, W. (2013). *Using Praat for linguistic research*. Boulder: University of Colorado at Boulder Phonetics Lab.

Thomas, G. (2008). An analysis of Xiamen tone circle. *Proceedings of the 27th West Coast Conference on Formal Linguistics* (pp. 422-430). Cascadilla Proceedings Project.

Thurgood, G. (2002). Vietnamese and tonogenesis: Revising the model and the analysis. *Diachronica, Graham; 19*, 333-363.

Thurgood, G. (2000). Voice quality differences and the origin of diphthongs. *Annual Meeting of the Berkeley Linguistics Society, 26,* pp. 295-303.

Tsay, J. (1991). Tone alternation in Taiwanese. In J. Ann & K. Yoshimura (Eds.), *Arizona Phonology Conference, 4*, pp. 76-87.

Tsay, J. (1994). *Phonological pitch.* (Doctoral dissertation, University of Arizona). Retrieved from http://arizona.openrepository.com/arizona/handle/10150/186900

Tsay, J. (1999). Bootstrapping into Taiwanese tone sandhi. *Chinese Languages and Linguistics V, Symposium Series of the Institute of History and Philology* (pp. 311-333).

Tsay, J., & Myers, J. (1996). Taiwanese tone sandhi as allomorph selection. *Proceedings of the Twenty-Second Annual Meeting of the Berkeley Linguistic Society: General Session and Parasession on the Role of Learnability in Grammatical Theory, 22,* pp. 395-405.

Wang, S. (1992). An experimental study on the productivity of Taiwanese tone sandhi. *Sea, 21*(1), 116-129.

Wang, W. S.-Y. (1967). Phonological features of tone. *International Journal of American Linguistics, 33,* 93-105.

Wedekind, K. (1983). A six-tone language in Ethiopia: Tonal analysis of Benč[4] non[4]. *Journal of Ethiopian Studies, 16,* 129-156.

Whalen, D., & Levitt, A. G. (1995). The universality of intrinsic F0 of vowels. *Journal of Phonetics, 23,* 349-366.

Woo, N. (1969). *Prosody and phonology* (Doctoral dissertation, Massachusetts Institute of Technology). Retrieved from https://dspace.mit.edu/handle/1721.1/14708

Woodrow, L. (2014). *Writing about quantitative research in applied linguistics.* London, England: Palgrave Macmillan.

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

Wu, X. (2012). Southern Min dialect and the great dictionary of Southern Min (闽南方言和《闽南方言大词典》). *Cishu Yanjiu* (辞书研究), 61-66.

Xie, X. (1818). *The phonology of common Zhangzhou speech (Huiji yasu tong shiwu yin* 汇集雅俗通十五音*).* Shanghai: Shanghai Chinese Classics (Shanghai guji chubanshe 上海古籍出版社).

Xu, C. X., & Xu, Y. (2003). Effects of consonant aspiration on Mandarin tones. *Journal of the International Phonetic Association, 33*(02), 165-181.

Xu, Y. (1994). Production and perception of coarticulated tones. *The Journal of the Acoustical Society of America*, 95(4), 2240-2253.

Xu, Y. (2006). Tone in connected discourse. *Encyclopedia of language and linguistics, 12,* 742-750.

Yang, X. (1999). How to distinguish initial consonants of Middle Chinese based on Zhangzhou speech (怎样利用漳州音辨别中古声母). In B. Chen (Ed.), *Southern Min dialect—Studies of Zhangzhou variety* (闽南方言一漳州话研究; pp. 136-143). Beijing, China: Zhongguo Wenlian Chubanshe (中国文联出版社).

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

Yang, X (2014). Studies of Zhangzhou Yinping tone and its dischronic evolution 漳州方言阴平调的调形特点与历史演变. *Journal of Minnan Normal University*, 03, 45-52.

Yin, X. (2009). Acoustic analysis of tonal patterns in Zhangzhou (漳州话声调格局的分析). *Journal of Chifeng University* 赤峰学院学报, *30*(6), 31-33.

Yip, M. (1980). *The tonal phonology of Chinese* (Doctoral dissertation, Massachusetts Institute of Technology). Retrieved from https://dspace.mit.edu/handle/1721.1/15971

Yip, M. (1989). Contour tones. *Phonology, 6,* 149-174.

Yip, M. (1993). Tonal register in East Asian languages. In K. Snider & H. v. Hulst (Eds.), *The phonology of tone: The representation of tonal register* (pp. 245-268). Berlin, Germany: Walter de Gruyter.

Yip, M. (2002). *Tone.* Cambridge, England: Cambridge University Press.

Yu, A. C. (2010). Tonal effects on perceived vowel duration. *Laboratory Phonology*, 151-168.

Yu, K. M., & Lam, H. W. (2014). The role of creaky voice in Cantonese tonal perception. *Journal of the Acoustical Society of America, 16,* 1320-1333.

Yuan, J., & Liberman, M. (2010). F0 declination in English and Mandarin broadcast news speech. *Eleventh Annual Conference of the International Speech Communication Association.*

ZCCEC. (1999). *Zhangzhou City Gazette-Dialect Volume* 漳州市志-方言 (Vol. 49). Beijing, China: Zhongguo Shehui Kexue Chubanshe 中国社会科学出版社.

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

Zhang, H. (1993). The syntactic condition of Taiwanese tone sandhi. *Proceedings of PASFoCoI (1993): Pacific Asia Conference on Formal and Computational Linguistics* (pp. 296-307).

Zhang, J. (2001). *The effects of duration and sonority on contour tone distribution—Typological survey and formal analysis* (Doctoral dissertation, University of California: Los Angeles). Retrieved from https://linguistics.drupal.ku.edu/sites/linguistics.ku.edu/files/docs/Zhang/zhang-routledge.pdf

Zhang, J. (2004). Contour tone licensing and contour tone representation. *Language and Linguistics*, *5*, 925-968.

Zhang, J. (2007). A directional asymmetry in Chinese tone sandhi systems. *Journal of East Asian Linguistics*, *16*, 259-302.

Zhang, J. (2010). Issues in the analysis of Chinese tone. *Language and Linguistics Compass*, *4*, 1137-1153.

Zhang, J. (2014). Tones, tonal phonology, and tone sandhi. In J. C. Huang, A. Y. Li, & A. Simpson (Eds.), *The handbook of Chinese linguistics* (pp. 443-464). Hoboken, NJ: Wiley-Blackwell.

Zhang, J., Lai, Y., & Sailor, C. (2007). Effects of phonetics and frequency on the productivity of Taiwanese tone sandhi. *The Annual Meeting of the Chicago Linguistic Society*, *43*, pp. 273-286.

Zhang, J., Lai, Y., & Sailor, C. (2006). Opacity, phonetics, and frequency in Taiwanese tone sandhi. *The Annual Meeting of the Chicago Linguistic Society*, *43*, pp. 3019-3038.

Zhang, J., & Liu, J. (2011). Tone sandhi and tonal coarticulation in Tianjin Chinese. *Phonetica*, 68(3), 161-191.

Zhou, C. (2006). *The great Southern Min dictionary* 闽南方言大词典. Fuzhou, China: Fujian Renmin Chubanshe 福建人民出版社.

Zhu, X. (2004). F0 normalization—How to deal with between-speaker tonal variations (基频归一化-如何处理声调的随机差异). *Linguistic Sciences* 语言科学, 3-19.

Zhu, X. (2012). Multiregisters and four levles: A new tonal model. *Journal of Chinese linguistics*, *40*, 1-17.

Appendices

Appendix A: Word Lists

Table A1. Monosyllabic token elicitation

Tone	Character	Example	Gloss
1	伊	ʔi1	she/he
	刀	tɵ1	knife
	医	ʔi1	medical
	甜	tĩ1	sweet
	鸡	ke1	chicken
	猪	ti1	pig
	天	tʰĩ1	sky
	胶	kɐ1	glue
	脚	kʰɐ1	foot
	诗	si1	poet
	东	tɐŋ1	east
	冰	piŋ1	ice
	山	swẽ1	mountain
	州	tsju1	state
	漳	tsjɐŋ1	Zhang
	箱	sjɔ̃1	box
	腔	kʰjɔ̃1	accent
	车	tsʰjɐ1	vehicle
	区	kʰi1	district
	边	pĩ1	side
2	图	tɔ2	image
	土	tʰɔ2	earth
	头	tʰɐw2	head
	平	pɛ̃2	flat
	时	si2	time
	柴	tsʰɐ2	firewood
	牛	ɠu2	cow
	平	pɛ̃2	flat
	移	ʔi2	move
	糊	kɔ2	paste
	茶	tɛ2	tea
	辞	si2	resign
	蹄	te2	hoof

Tone	Character	Example	Gloss
2	男	ɗɐm2	male
	台	tɐj2	station
	桥	kjɵ2	bridge
	门	ɓwĩ2	door
	龙	ɗjɔŋ2	dragon
	文	ɓun2	text
	姨	ʔi2	aunt
	南	ɗɐm2	south
3	主	tsu3	master
	椅	ʔi3	chair
	体	tʰe3	body
	死	si3	die
	短	te3	short
	米	ɓi3	rice
	苦	kʰɔ3	bitter
	考	kʰɵ3	test
	虎	ɦɔ3	tiger
	买	ɓe3	buy
	鼓	kɔ3	drum
	抢	tsʰjɔ̃3	grab
	好	ɦɵ3	good
	尾	ɓwe3	tail
	景	kiŋ3	view
	海	ɦɐj3	sea
	碾	kɐ3	grind
	酒	tsju3	wine
	马	ɓɛ3	horse
4	厝	tsʰu4	house
	告	kɵ4	report
	报	pɵ4	newspaper
	四	si4	four
	客	kʰɛ4	guest
	意	ʔi4	intension
	戏	ɦi4	drama
	教	kɐ4	teach
	气	kʰi4	gas
	雪	se4	snow
	肉	ɓɐ4	meat

Tone	Character	Example	Gloss
4	试	ts^{h}i4	test
	课	k^{h}ɵ4	class
	铁	t^{h}i4	iron
	八	pe4	eight
	案	ʔɐn4	case
	看	kwɐ̃4	watch
	破	p^{h}wɐ4	broken
	教	kɐw4	teach
5	事	su5	affair
	低	kɛ5	low
	卖	ɓe5	sell
	味	ɓi5	flavor
	字	zi5	character
	是	si5	yes
	易	ʔi5	play
	雨	ħɔ5	rain
	路	ɗɔ5	road
	面	ɓin5	surface
	面	ɓĩ5	noodle
	地	te5	land
	大	twɐ5	big
	电	tjɐn5	electricity
	视	si5	inspect
	硬	ɠɛ̃5	tough
	洞	tɔŋ5	hold
	院	ʔĩ5	hospitcal
6	击	kik6	hit
	级	kip6	grade
	法	ħwɐt6	law
	国	kɔk6	country
	激	kik6	excite
	色	sik6	colour
	一	ʔik6	one
	室	sit6	room
	七	ts^{h}it6	seven
	吉	kit6	lucky
	结	kɐt6	tie
	湿	sip6	wet

Tone	Character	Example	Gloss
6	角	kɐk6	horn
	亿	ʔik6	billion
	吸	kip6	suck
	骨	kut6	bone
	出	tsʰut6	out
7	佛	ħut7	Buddha
	席	sit7	seat
	日	zit7	day
	木	ɓɔk7	wood
	毒	tɔk7	poison
	浴	ʔik7	bathe
	热	zjɐt7	hot
	熟	sik7	cooked
	玉	ɠjɔk7	jade
	直	tit7	straight
	六	ɗɐk7	six
	十	tsɐp7	ten
8	活	ʔwɐ8	alive
	白	pɛ8	white
	舌	tsi8	tongue
	药	ʔjɵ8	medicine
	叶	ħjɵ8	leaf
	辣	ɗwɐ8	spicy
	薄	pɵ8	thin
	吃	tsjɐ8	eat

Table A2. Disyllabic token elicitation

S1	S2	Pattern	Character	Example	Gloss	Phrase
1	1	1+1	脚车	kʰɐ1.tsʰjɐ1	bicycle	NP
			猪脚	ti1.kʰɐ1	pig feet	NP
			鸡脚	ke1.kʰɐ1	chicken feet	NP
			乌鸡	ʔɔ1.ke1	black-boned chicken	NP
			猪心	ti1.sim1	pig heart	NP
			山猪	swɛ̃1.ti1	mountain pig	NP
			医生	ʔi1.siŋ1	doctor	NP
			中医	tjɔŋ1.ʔi1	Chinese medicine	NP
			家乡	kɛ1.ɦjɐŋ1	hometown	NP
			专家	tswɐn1.kɛ1	expert	NP
			家规	kɛ1.kwi1	family rules	NP
			冤家	ʔwɐn1.kɛ1	enemy, quarrel	NP; VP
1	2	1+2	脚球	kʰɐ1.kju2	football	NP
			奶牛	ɗiŋ1.ɠu2	cows	NP
			鸡头	ke1.tʰɐw2	chicken head	NP
			青茶	tsʰɛ̃1.tɛ2	raw tea	NP
			猪蹄	ti1.te2	pig feet	NP
			热茶	sjɵ1.tɛ2	hot tea	NP
			医疗	ʔi1.ɗjɐw2	medical treatment	NP
			专题	tswɐn1.te2	symposium	NP
			家庭	kɛ1.tiŋ2	household	NP
			天时	tʰĩ1.si2	climatic condition	NP
			家人	kɛ1.ɗɐŋ2	family member	NP
			当时	tɐŋ1.si2	at that time	AdvP
1	3	1+3	脚腿	kʰɐ1.tʰwi3	legs	NP
			黑马	ʔɔ1.ɓɛ3	black horse	NP
			鸡母	ke1.ɓɵ3	hen	NP
			斑马	pɐn1.ɓɛ3	zebra	NP
			猪屎	ti1.sɐj3	pig shit	NP
			生米	tsʰɛ̃1.ɓi3	uncooked rice	NP
			医保	ʔi1.pɵ3	medical insurance	NP
			家主	kɛ1.tsu3	the head of the family	NP
			家长	kɛ1.tjɔ̃3	parent of a child	NP
			尸体	si1.tʰe3	corpse	NP
			身体	sin1.tʰe3	body	NP

S1	S2	Pattern	Character	Example	Gloss	Phrase
1	4	1+4	脚印	kʰɐ1.ʔin4	footprint	NP
			新厝	sin1.tsʰu4	new house	NP
			申报	sin1.pɵ4	declare	VP
			生铁	sɛ̃1.tʰi4	raw iron	NP
			鸡嘴	ke1.tsʰwi4	chicken mouth	NP
			生客	tsʰɛ̃1.kʰɛ4	new guest	NP
			猪肺	ti1.ɦi4	pig lung	NP
			空气	kʰɔŋ1.kʰi4	gas; atmosphere	NP
			猪肉	ti1.ɓɐ4	pork	NP
			听课	tʰjɐ̃1.kʰɵ4	listen to a lecture	VP
			家教	kɛ1.kɐw4	family education	NP
			家世	kɛ1.si4	family background	NP
			通报	tʰɔŋ1.pɵ4	circulate a notice	VP
1	5	1+5	脚步	kʰɐ1.pɔ5	foot step	NP
			思路	su1.ɗɔ5	thinking	NP
			鸡蛋	ke1.ɗwĩ5	chicken egg	NP
			公路	kɔŋ1.ɗɔ5	public highway	NP
			猪肚	ti1.tɔ5	pork belly	NP
			生字	tsʰɛ̃1.zi5	unknown character	NP
			医患	ʔi1.ɦwɐn5	doctor and patient	NP
			医院	ʔi1.ʔĩ5	hospital	NP
			心事	sim1.su5	thoughts in heart	NP
			家务	kɛ1.ɓu5	housework	NP
			天地	tʰĩ1.te5	heven and earth; universe	NP
			家内	kɛ1.ɗɐj5	inside the family	NP
			工地	kɐŋ1.te5	construction site	NP
1	6	1+6	脚骨	kʰɐ1.kut6	foot bone	NP
			鸡角	ke1.kɐk6	rooster	NP
			交角	kɐw1.kɐk6	crossing angle	NP
			医德	ʔi1.tik6	medical ethics	NP
			家法	kɛ1.ɦwɐt6	family law	NP
			青色	tsʰɛ̃1.sik6	green colour	NP
			家室	kɛ1.sit6	spouse	NP
			中国	tjɔŋ1.kɔk6	China	NP

S1	S2	Pattern	Character	Example	Gloss	Phrase
1	7	1+7	脚蚀	kʰɐ1.sjɵt7	beriberi	NP
			归日	kwi1.zit7	all day long	AdvP
			鸡翅	ke1.sit7	chicken wings	NP
			雕佛	tjɐw1.ɦut7	carve a Buddha	VP
			医术	ʔi1.sut7	medical skill	NP
			家族	kɛ1.tsɔt7	family clan	NP
			真实	tsin1.sit7	truthful	AP
			家属	kɛ1.sjɔk7	family dependent	NP
			伸直	sin1.tit7	unbend; straighten	VP
1	8	1+8	奶白	ɗiŋ1.pɛ8	milk white	AP
			猪舌	ti1.tsi8	pig tongue	NP
			中药	tjɔŋ1.ʔjɵ8	Chinese medicine	NP
2	1	2+1	平脚	pɛ̃2.kʰɐ1	flat feet	NP
			牛奶	ɠu2.ɗiŋ1	cow milk	NP
			绒鸡	zjɔŋ2.ke1	fluffy chicken	NP
			牛刀	ɠu2.tɵ1	cattle knife	NP
			豪猪	ɦɵ2.ti1	porcupine	NP
			茶花	tɛ2.ɦwɐ1	camellia	NP
			名医	ɓjɐ̃2.ʔi1	famous doctor	NP
			提纲	te2.kɐŋ1	topic outline	NP
			头家	tʰɐw2.kɛ1	head of a family	NP
			时间	si2.kɐn1	time	NP
			成家	tsjɐ̃2.kɛ1	get married (set up a family)	VP
			时钟	si2.tsiŋ1	time clock	NP
2	2	2+2	牛头	ɠu2.tʰɐw2	cattle head	NP
			黄牛	ɦwi2.ɠu2	yellow cattle	NP
			牛皮	ɠu2.pʰwe2	cattle hide	NP
			红茶	ʔɐŋ2.tɛ2	black tea	NP
			茶楼	tɛ2.ɗɐw2	tea house	NP
			烧茶	ɦjɐ2.tɛ2	boil water	VP
			题材	te2.tsɐj2	theme	NP
			难题	ɗɐn2.te2	difficult question	NP
			时期	si2.ki2	time interval	NP
			临时	ɗim2.si2	temporary	AdvP
			时行	si2.kjɐ̃2	popular	AP
			同时	tɔŋ2.si2	at the same time	NP

S1	S2	Pattern	Character	Example	Gloss	Phrase
2	2	2+3	铜马	tɐŋ2.ɓɛ3	copper horse	NP
			牛屎	ɠu2.sɐj3	bull shit	NP
			骑马	kʰjɐ2.ɓɛ3	ride a horse	VP
			茶米	tɛ2.ɓi3	tea rice (dried tea leaves)	NP
			红米	ħɐŋ2.ɓi3	red rice	NP
			题海	te2.ħɐj3	question sea	NP
			财主	tsɐj2.tsu3	money owner (rich man)	NP
			时尾	si2.ɓwe3	end of a period	NP
			时点	si2.tjɐm3	time point	NP
			团体	tʰwɐn2.tʰe3	group	NP
2	4	2+4	牛嘴	ɠu2.tsʰwi4	cattle mouth	NP
			土厝	tʰɔ2.tsʰu4	mud house	NP
			铜铁	tɐŋ2.tʰi4	copper and iron	NP
			牛气	ɠu2.kʰi4	cattle breath	NP
			人客	ɗɐŋ2.kʰɛ4	human guest	NP
			茶店	tɛ2.tjɐm4	tea store	NP
			元气	ɠwɐn2.kʰi4	vitality	NP
			牛肉	ɠu2.ɓɐ4	cattle meat; beef	NP
			题库	te2.kʰɔ4	question bank	NP
			评课	pʰiŋ2.kʰɵ4	assess a lecture	VP
			时态	si2.tʰɐj4	tense and aspect	NP
			回报	ħwe2.pɵ4	pay back; reciprocate	VP
			时顿	si2.tun4	meal time	NP
2	5	2+5	牛肚	ɠu2.tɔ5	cattle belly	NP
			门路	ɓwĩ2.ɗɔ5	door and road (social connection)	NP
			牛市	ɠu2.tsʰi5	cattle market	NP
			头路	tʰɐw2.ɗɔ5	living method	NP
			茶树	tɛ2.tsʰju5	tea tree	NP
			文字	ɓun2.zi5	Chinese character	NP
			题字	te2.zi5	inscription; inscribe a word	NP; VP
			同事	tɔŋ2.su5	colleague	NP
			时阵	si2.tsun5	time period	NP
			田地	tʰɐn2.te5	field	NP
			时代	si2.ɗɐj5	era	NP
			农地	ɗɔŋ2.te5	farm land	NP

S1	S2	Pattern	Character	Example	Gloss	Phrase
2	6	2+6	牛角	ɠu2.kɐk6	cattle horn	NP
			茶色	tɛ2.sik6	tea colour	NP
			题刻	te2.kik6	inscription	NP
			龙骨	ɗjɔŋ2.kut6	dragon bone; spine bone	NP
			时速	si2.sjɔt6	hourly speed	NP
			颜色	ɠɐn2.sik6	colour	NP
			时刻	si2.kʰik6	moment	NP
			全国	tswɐn2.kɔk6	whole country	NP
2	7	2+7	头日	tʰɐw2.zit7	the first day	NP
			牛杂	ɠu2.tsɐp7	cattle offal	NP
			成佛	siŋ2.ɦut7	become a Buddha	VP
			题目	te2.ɓɐk7	topic	NP
			时食	si2.sit7	special food for a particular time	NP
			其实	ki2.sit7	in fact	AdvP
			时日	si2.zit7	time and day	NP
			垂直	swe2.tit7	perpendicular; plumb	AP; VP
2	8	2+8	牛舌	ɠu2.tsi8	cattle tongue	NP
			纯白	sun2.pɛ8	pure white	AP
			茶叶	tɛ2.ɦjɵ8	tea leaves	NP
			农药	ɗɔŋ2.ʔjɵ8	farm chemicals; pesticide	NP
3	1	3+1	软脚	ɗwĩ3.kʰɐ1	weak feet (paralysis)	NP
			马车	ɓɛ3.tsʰjɐ1	horse carriage	NP
			火鸡	ɦwe3.ke1	turkey	NP
			马脚	ɓɛ3.kʰɐ1	horse feet	NP
			野猪	ʔjɐ3.ti1	wild pig	NP
			米糕	ɓi3.kɵ1	rice cake	NP
			主张	tsu3.tjɔ̃1	propose; proposition	NP/VP
			祖家	tsɔ3.kɛ1	ancestral home	NP
			体操	tʰe3.tʰɐw1	gymnastics	NP
			酒家	tsju3.kɛ1	restaurant	NP
			体温	tʰe3.ʔun1	body temperature	NP

S1	S2	Pattern	Character	Example	Gloss	Phrase
3	2	3+2	水牛	tswi3.ɠu2	water buffalo	NP
			马头	ɓɛ3.tʰɐw2	horse head	NP
			早茶	tsɐ3.tɛ2	morning tea	NP
			马蹄	ɓɛ3.te2	horse feet	NP
			品茶	pʰiŋ3.tɛ2	taste tea	VP
			米糊	ɓi3.kɔ2	rice paste	NP
			考题	kʰɵ3.te2	examination question	NP
			主人	tsu3.dɐ̃ŋ2	main people (host)	NP
			准时	tsun3.si2	on time	AdvP
			体能	tʰe3.ɗiŋ2	physical fitness	NP
			小时	sjɵ3.si2	hour	NP
			体型	tʰe3.ɦiŋ2	body figure	NP
3	3	3+3	马尾	ɓɛ3.ɓwe3	horse tail	NP
			宝马	pɵ3.ɓɛ3	precious horse; BMW	NP
			马屎	ɓɛ3.sɐj3	horse shit	NP
			海马	ɦɐj3.ɓɛ3	sea horse	NP
			米粉	ɓi3.ɦun3	rice noodle	NP
			碾米	kɐ3.ɓi3	mill rice	VP
			主管	tsu3.kwɐn3	be in charge of/ person in charge	VP/NP
			买主	ɓe3.tsu3	buyer	NP
			体表	tʰe3.pjɵ3	body surface	NP
			好体	ɦɵ3.tʰe3	good manner	AP
			体检	tʰe3.kjɐm3	physical examination	NP
			坏体	pɐj3.tʰe3	bad manner	AP

S1	S2	Pattern	Character	Example	Gloss	Phrase
3	4	3+4	马戏	ɓɛ3.ħi4	horse drama; circus	NP
			马肉	ɓɛ3.ɓɐ4	horse meat	NP
			祖厝	tsɔ3.tsʰu4	ancestor's house	NP
			好铁	ħɵ3.tʰi4	good iron	NP
			马嘴	ɓɛ3.tsʰwi4	horse mouth	NP
			好客	ħɵ3.kʰɛ4	good customet	NP
			米店	ɓi3.tjɐm4	rice shop	NP
			氧气	ʔjɐŋ3.kʰi4	oxygen	NP
			粉肉	ħun3.ɓɐ4	tender meat	NP
			主要	tsu3.ʔjɐw4	principal	AP
			讲课	kɔŋ3. kʰɵ4	deliver a lecture	VP
			体制	tʰe3.tsi4	system of organisation	NP
			举报	ki3.pɵ4	tip off	VP
			体态	tʰe3.tʰɐj4	posture	NP
			警报	kiŋ3.pɵ4	alarm	NP;VP
3	5	3+5	马路	ɓɛ3.ɗɔ5	horse way (road)	NP
			跑路	tsɐw3.ɗɔ5	run on the road (flee)	NP
			马步	ɓɛ3.pɔ5	horse step	NP
			米饭	ɓi3.pwĩ5	rice meal	NP
			写字	sjɐ3.zi5	write characters; write	NP
			主任	tsu3.zim5	director	NP
			本事	pun3.su5	ability	NP
			体重	tʰe3.tɐŋ5	body weight	NP
			土地	tʰɔ3.te5	earth and land; territory	NP
			体谅	tʰe3.ɗjɐŋ5	understand	VP
			草地	tsʰɐw3.te5	grassland	NP
3	6	3+6	马骨	ɓɛ3.kut6	horse bone	NP
			顶角	tiŋ3.kɐk6	upper corner	NP
			米色	ɓi3.sik6	rice colour	NP
			主角	tsu3.kɐk6	leading actor	NP
			锁骨	sɵ3.kut6	collarbone	NP
			体贴	tʰe3.tʰjɐp6	considerate	AP
			彩色	tsʰɐj3.sik6	multicolour	NP
			体质	tʰe3.tsit6	physique	NP
			祖国	tsɔ3.kɔk6	native country	NP

S1	S2	Pattern	Character	Example	Gloss	Phrase
3	7	3+7	马术	ɓɛ3.sut7	horse skill (horsemanship)	NP
			每日	ɓwe3.zit7	every day	NP
			米业	ɓi3.ɠjɐp7	rice industry	NP
			礼佛	ɗe3.ɦut7	worship buddha	VP
			主席	tsu3.sit7	chairman	NP
			体力	tʰe3.ɗɐt7	physical strength	NP
			老实	ɗɐw3.sit7	honest	AP
			体育	tʰe3.ʔjɔk7	physical education	NP
			土直	tʰɔ3.tit7	blunt	AP
3	8	3+8	马舌	ɓɛ3.tsi8	horse tongue	NP
			米白	ɓi3.pɛ8	rice white	AP
			草药	tsʰɐw3.ʔjɵ8	herb medicine	NP
4	1	4+1	裤脚	kʰɔ4.kʰɐ1	trouser legs	NP
			赤脚	tsʰjɐ4.kʰɐ1	bare feet	NP
			厝边	tsʰu4.pĩ1	house side; nearby the house	NP
			铁钉	tʰi4.tiŋ1	iron nail	NP
			散鸡	swɐ̃4.ke1	free-range chicken	NP
			肉鸡	ɓɐ4.ke1	chicken	NP
			客厅	kʰɛ4.tʰjɐ̃1	guest hall	NP
			细猪	se4.ti1	little pig	NP
			肉猪	ɓɐ4.ti1	pig	NP
			气温	kʰi4.ʔun1	temperature	NP
			肉筋	ɓɐ4.kiŋ1	pork tendon	NP
			兽医	sju4.ʔi1	animous doctor	NP
			太医	tʰɐj4.ʔi1	imperial doctor	NP
			课间	kʰɵ4.kɐn1	course break	NP
			顾家	kɔ4.kɛ1	look after family; family-oriented	NP
			报销	pɵ4.sjɐw1	reimburse	VP
			建家	kjɐn4.kɛ1	set up a family	VP
			报恩	pɵ4.ʔin1	pay a debt of gratitude	VP

S1	S2	Pattern	Character	Example	Gloss	Phrase
4	2	4+2	看牛	kwẽ4.ɠu2	pasture cattle	VP
			铁牛	tʰi4.ɠu2	iron cattle	NP
			厝前	tsʰu4.tsin2	in front of house	NP
			铁门	tʰi4.ɓwĩ2	iron gate	NP
			散茶	swẽ4.tɛ2	tea in bulk	NP
			客房	kʰɛ4.pɐŋ2	guest room	NP
			泡茶	pʰɐw4.tɛ2	make tea	VP
			气瓶	kʰi4.pɐŋ2	gas cylinder	NP
			肉皮	ɓɐ4.pʰwe2	pig skin	NP
			课题	kʰɵ4.te2	project topic	NP
			课文	kʰɵ4.ɓun2	lessen text	NP
			过时	kwe4.si2	outdate	AP
			报名	pɵ4.ɓjẽ2	register; sign up	VP
			按时	ʔɐn4.si2	on time	AdvP
			报仇	pɵ4.sju2	revenge	VP
4	3	4+3	战马	tsjɐn4.ɓɛ3	battle horse	NP
			厝顶	tsʰu4.tiŋ3	house roof	NP
			铁桶	tʰi4.tʰɐŋ3	iron barrel	NP
			看马	kwẽ4.ɓɛ3	pasture horse	VP
			客鸟	kʰɛ4.tsjɐw3	guest bird	NP
			糙米	tsʰɵ4.ɓi3	brown rice	NP
			赤米	tsʰjɐ4.ɓi3	red rice	NP
			气管	kʰi4.kwĩ3	trachea	NP
			肉体	ɓɐ4.tʰe3	flesh body	NP
			厝主	tsʰu4.tsu3	house owner	NP
			课本	kʰɵ4.pun3	text book	NP
			抗体	kʰɔŋ4.tʰe3	antibody	NP
			报纸	pɵ4.tswɐ3	newspaper	NP
			固体	kɔ4.tʰe3	solid body; solid	NP
			报考	pɵ4.kʰɵ3	register for examination	VP

S1	S2	Pattern	Character	Example	Gloss	Phrase
4	4	4+4	厝架	tsʰu4.kɛ4	house sketch	NP
			厝契	tsʰu4.kʰe4	house deeds	NP
			建厝	kjɐn4.tsʰu4	build a house	VP
			拆厝	tʰjɐ4.tsʰu4	dismantle a house	VP
			铁线	tʰi4.swɐ̃4	iron wire	NP
			废铁	ɦwi4.tʰi4	abandoned iron	NP
			客桌	kʰɛ4.tɵ4	guest table	NP
			顾客	kɔ4.kʰɛ4	customer	NP
			气派	kʰi4.pʰɐj4	imposing style	NP
			客气	kʰɛ4.kʰi4	courteous	AP
			肉冻	ɓɐ4.tɐŋ4	meat jelly	NP
			醋肉	tsʰɔ4.ɓɐ4	sweet and sour pork	NP
			赤肉	tsʰjɐ4.ɓɐ4	lean meat	NP
			课卷	kʰɵ4.kwĩ4	text paper	NP
			晚课	ʔwɐ̃4.kʰɵ4	evening class	NP
			报告	pɵ4.kɵ4	report	NP;VP
			见报	kĩ4.pɵ4	appear in the newspaper	VP
			报案	pɵ4.ʔɐn4	report a case	VP
			快报	kʰwɐj4.pɵ4	bulletin	NP
4	5	4+5	厝瓦	tsʰu4.ɦjɐ5	house tile	NP
			过路	kwe4.ɗɔ5	pass by	VP
			铁路	tʰi4.ɗɔ5	railway	NP
			客户	kʰɛ4.ɦɔ5	customer	NP
			半路	pwɐ̃4.ɗɔ5	halfway	NP
			气味	kʰi4.ɓi5	odor	NP
			汉字	ɦɐn4.zi5	Chinese character	NP
			八字	pe4.zi5	horoscope	NP
			课外	kʰɵ4.ɠwɐ5	extracurricular	NP
			故事	kɔ4.su5	story	NP
			报道	pɵ4.tɵ5	news report; report	NP;VP
			性地	siŋ4.te5	personality	NP
			报幕	pɵ4.ɓɔ5	announce (a program)	NP;VP
			空地	kʰɐŋ4.te5	vacant land	NP

S1	S2	Pattern	Character	Example	Gloss	Phrase
4	6	4+6	厝角	tsʰu4.kɐk6	house corner	NP
			气色	kʰi4.sik6	complexion	NP
			铁笔	tʰi4.pit6	iron pen; pen	NP
			课室	kʰɵ4.sit6	class room	NP
			落骨	ɗɐw4.kut6	bone dislocation	NP
			肉骨	ɓɐ4.kut6	pork bone	NP
			报答	pɵ4.tɐp6	reciprocate	VP
			报国	pɵ4.kɔk6	serve the country	VP
			建国	kjɐn4.kɔk6	build a country	VP
4	7	4+7	半日	ɓwɐ̃4.zit7	half a day	NP
			客席	kʰɛ4.sit7	guest seat	NP
			气力	kʰi4.ɗit7	power	NP
			肉食	ɓɐ4.sit7	pork food	NP
			拜佛	pɐj4.ɦut7	worship Buddha	VP
			课业	kʰɵ4.ɠjɐp7	coursework	NP
			报复	pɵ4.ɦɔk7	revenge	VP
			证实	tsiŋ4.sit7	confirm	VP
			报业	pɵ4.ɠjɐp7	newspaper industry	VP
			正直	tsjɐ̃4.tit7	upright	AP
4	8	4+8	铁石	tʰi4.tsjɵ8	ironstone	NP
			漂白	pʰjɵ4.pɛ8	whiten	VP
			拆白	tʰjɐ4.pɛ8	straightforward	VP
			厝宅	tsʰu4.tʰɛ8	house	NP
			做药	tsɵ4.ʔjɵ8	medical use	VP
5	1	5+1	后脚	ʔɐw5.kʰɐ1	rear feet	NP
			路灯	ɗɔ5.tiŋ1	road lamp	NP
			咬鸡	kɐ5.ke1	gamecock	NP
			路边	ɗɔ5.pĩ1	road side	NP
			养猪	tsʰi5.ti1	feed pig	VP
			字音	zi5.ʔim1	character pronunciation	NP
			校医	ɦɐw5.ʔi1	school doctor	NP
			事先	su5.sjɐn1	beforehand	AdvP
			大家	twɐ5.kɛ1	mother in law	NP
			地方	te5.ɦɔŋ1	area	NP
			道家	tɵ5.kɛ1	Daoism	NP
			地基	te5.ki1	foundation	NP

S1	S2	Pattern	Character	Example	Gloss	Phrase
5	2	5+2	养牛	tsʰi5.ɠu2	feed cattle	VP
			路头	ɗɔ5.tʰɐw2	beginning of a road	NP
			用茶	ʔjɔŋ5.tɛ2	have tea	VP
			路途	ɗɔ5.tɔ2	journey	NP
			供茶	kjɔŋ5.tɛ2	serve tea (to deities)	VP
			字词	zi5.su2	character and word	NP
			问题	ɓwĩ5.te2	question	NP
			事前	su5.tsin2	before the event	AdvP
			饭时	pwĩ5.si2	meal time	NP
			地球	te5.kju2	Earth	NP
			五时	ɠɔ5.si2	Five o'clock (a special time on dragon boat day)	NP
			地图	te5.tɔ2	map	NP
5	3	5+3	养马	tsʰi5.ɓɛ3	pasture a horse	VP
			路尾	ɗɔ5.ɓwe3	end of a road	NP
			坐马	tse5.ɓɛ3	ride a horse	VP
			路警	ɗɔ5.kiŋ3	railway police	NP
			大米	twɐ5.ɓi3	rice	NP
			字典	zi5.tjɐm3	dictionary	NP
			地主	te5.tsu3	land owner	NP
			事理	su5.ɗi3	affair	NP
			字体	zi5.tʰe3	style of calligraphy	NP
			地理	te5.ɗi3	geography	NP
			具体	ki5.tʰe3	specific; concrete	AP
			地震	te5.tsin3	earthquake	NP

S1	S2	Pattern	Character	Example	Gloss	Phrase
5	4	5+4	旧厝	kɔ3.tsʰu4	old house	NP
			路线	ɗɔ5.swɐ̃4	routine	NP
			地铁	te5.tʰi4	subway	NP
			远客	ɦwĩ5.kʰɛ4	guest from afar	NP
			路费	ɗɔ5.ɦwi4	travelling expenses	NP
			运气	ʔun5.kʰi4	luckiness	NP
			字库	zi5.kʰɔ4	character bank	NP
			卖肉	ɓe5.ɓɐ4	sell meat	VP
			上课	tsjɔ̃5.kʰɵ4	attend a lecture	VP
			事故	su5.kɔ4	event	NP
			事迹	su5.tsjɐ4	deeds	NP
			汇报	ɦwe5.pɵ4	report	VP
			地税	te5.swe4	land tax	NP
			电报	tjɐn5.pɵ4	telegraph	NP
			地价	te5.kɛ4	land price	NP
5	5	5+5	路段	ɗɔ5.twɐ̃5	one part of a road	NP
			顺路	sun5.ɗɔ5	on the way	VP
			路面	ɗɔ5.ɓin5	road surface	NP
			电路	tjɐn5.ɗɔ5	electric circuit	NP
			字画	zi5.ʔwɐ5	calligraphy and painting	NP
			练字	ɗjɐn5.zi5	practice (writing) character	VP
			事后	su5.ʔɐw5	after the event	AdvP
			共事	kjɔŋ5.su5	work together	VP
			地洞	te5.tɔŋ5	burrow	NP
			内地	ɗɐj5.te5	inland	NP
			地位	te5.ʔwi5	social status	NP
			外地	ɠɐ5.te5	other places	NP
5	6	5+6	路节	ɗɔ5.tsɐt6	one part of a road	NP
			路宿	ɗɔ5.sjɵt6	roadside inn	NP
			下角	ʔɛ5.kɐk6	lower corner	NP
			字帖	zi5.tʰjɐp6	calligraphy copybook	NP
			大骨	ɗwɐ5.kut6	big bone	NP
			地壳	te5.kʰɐk6	Earth crust	NP
			面色	ɓin5.sik6	facial expression	NP
			地质	te5.tsit6	geology	NP
			外国	ɠwɐ5.kɔk6	foreign country	NP

S1	S2	Pattern	Character	Example	Gloss	Phrase
5	7	5+7	后日	ʔɐw5.zit7	the day after	NP
			路局	ɗɔ5.kjɔt7	railway administration	NP
			字目	zi5.ɓɐk7	character entry	NP
			念佛	ɗjɐm5.ɦut7	pray to Buddha	VP
			事实	su5.sit7	truth	NP
			地狱	te5.ɠɐk7	hell	NP
			地热	te5.zjɐt7	geothermal	NP
			硬直	ɠɛ̃5.tit7	loyal and honest	AP
5	8	5+8	路药	ɗɔ5.ʔjɵ8	roadside medicine	NP
			蛋白	ɗwĩ5.pɛ8	albumen; egg white	NP
			用药	ʔjɔŋ5.ʔjɵ8	take medicine	VP
6	1	6+1	角瓜	kɐk6.kwɐ1	horned melon	NP
			法医	ɦwɐt6.ʔi1	forensic doctor	NP
			骨胶	kut6.kɐ1	bone glue	NP
			出家	ts^hut6.kɛ1	to be a monk	VP
			色差	sik6.ts^hɐ1	colour difference	NP
			国家	kɔk6.kɛ1	nation; country	NP
6	2	6+2	色茶	sik6.tɛ2	colourful tea; flavoured tea	NP
			角头	kɐk6.t^hɐw2	corner (place)	NP
			湿茶	sip6.tɛ2	soak tea	VP
			出题	ts^hut6.te2	assign a question	VP
			骨头	kut6.t^hɐw2	bone	NP
			及时	kip6.si2	in time	AdvP
			色盲	sik6.ɓɔŋ2	colour blindness	NP
			吉时	kit6.si2	lucky time	NP
			国旗	kɔk6.ki2	national flag	NP
6	3	6+3	八马	pɐt6.ɓɛ3	Eight Horse	NP
			竹马	tik6.ɓɛ3	bamboo horse	NP
			角马	kɐk6.ɓɛ3	horned horse	NP
			失主	sip6.tsu3	owner of lost property	NP
			骨髓	kut6.ts^hwe3	bone marrow	NP
			国体	kɔk6.t^he3	national prestige	NP
			色水	sik6.tswi3	colour	NP
			国产	kɔk6.sɐn3	domestically-made	AP

S1	S2	Pattern	Character	Example	Gloss	Phrase
6	4	6+4	吸铁	kʰip6.tʰi4	magnet iron	NP
			穷客	sɐp6.kʰɛ4	needy customer	NP
			角四	kɐk6.si4	forty cents	NP
			角铁	kɐk6.tʰi4	horned iron	NP
			福气	ɦɔk6.kʰi4	good fortune	NP
			缺课	kʰwɐt6.kʰɵ4	absence from class	VP
			骨架	kut6.kɛ4	bone sketch	NP
			恶报	ʔɔk6.pɵ4	retribution for evildoing	NP
			色素	sik6.sɔ4	pigment	NP
			福报	ɦɔk6.pɵ4	reward	NP
			国庆	kɔk6.kʰiŋ4	national celebration	NP
6	5	6+5	出路	tsʰut6.ɗɔ5	outlet	NP
			角度	kɐk6.tɔ5	angle	NP
			国事	kɔk6.su5	national affair	NP
			骨路	kut6.ɗɔ5	bone way	NP
			福地	ɦɔk6.te5	lucky place	NP
			色调	sik6.tjɐw5	tonality	NP
			湿地	sip6.te5	wetland	NP
6	6	6+6	缺角	kʰwɐt6.kɐk6	notch	NP/VP
			骨节	kut6.tsɐt6	bone joint	NP
			色笔	sik6.pit6	colourful pen	NP
			出色	tsʰut6.sik6	outstanding	AP
			国法	kɔk6.ɦwɐt6	national law	NP
			出国	tsʰut6.kɔk6	go aboard	VP
6	7	6+7	出日	tsʰut6.zit7	sun comes out	VP
			七佛	tsʰit6.ɦut7	Seven Buddha	NP
			骨力	kut6.ɗɐt7	bone power	NP
			色泽	sik6.tit7	colour and lustre	NP
			确实	kʰɐk6.sit7	indeed	AdvP
			结实	kjɐt6.sit7	durable	NP
			国籍	kɔk6.tsit7	nationality	NP
			笔直	pit6.tit7	straight as a pen	AP
6	8	6+8	角石	kɐk6.tsjɵ8	horned stone	NP
			国药	kɔk6.ʔjɵ8	national medicine	NP

S1	S2	Pattern	Character	Example	Gloss	Phrase
7	1	7+1	袤脚	pɐk7.kʰɐ1	bind feet	VP
			日间	zit7.kɐn1	daily time	NP
			佛珠	ɦut7.tsu1	Buddha beads	NP
			学医	ɦɐk7.ʔi1	study medicine	VP
			佛家	hut7.kɛ1	Buddhist	NP
			实心	sit7.sim1	solid	NP
			俗家	sjɔk7.kɛ1	secular home	NP
			直冲	tit7.tsʰjɔŋ1	rush ahead	VP
7	2	7+2	绑牛	pɐk7.ɠu2	bind cattle	VP
			日头	zit7.tʰɐw2	sun	NP
			熟茶	sik7.tɛ2	fermented tea	NP
			绿茶	ɗik7.tɛ2	green tea	NP
			佛门	ɦut7.ɓwĩ2	Buddha entrance (Buddhism)	NP
			别题	pɐt7.te2	another question	NP
			日时	zit7.si2	daily time	NP
			实权	sit7.kwɐn2	real power	NP
			一时	tsit7.si2	temporarily	AdvP
			直航	tit7.ɦɐŋ2	direct flight	NP
7	3	7+3	日子	zit7.tsi3	day	NP
			木马	ɓɔk7.ɓɛ3	wooden horse	NP
			糯米	tsut7.ɓi3	sticky rice	NP
			佛祖	ɦut7.tsɔ3	Buddha	NP
			业主	ɠjɐp7.tsu3	proprietor	NP
			立体	ɗit7.tʰe3	stereoscopic	NP; AP
			实践	sit7.tsjɐn3	practice	NP; VP
			集体	tsip7.tʰe3	collective group	NP
			直管	tit7.kwɐn3	in direct charge of	VP
7	4	7+4	入厝	zip7.tsʰu4	move to a new house	VP
			日记	zit7.ki4	diary	NP
			熟客	sik7.kʰɛ4	familiar customer	NP
			毒气	tɔk7.kʰi4	toxic gas	NP
			佛教	ɦut7.kɐw4	Buddhism	NP
			日报	zit7.pɵ4	daily newspaper	NP
			实际	sit7.tse4	actual	NP
			密报	ɓit7.pɵ4	secret report	NP
			直线	tit7.swɐ̃4	straight line	NP

S1	S2	Pattern	Character	Example	Gloss	Phrase
7	5	7+5	日用	zit7.ʔjɔŋ5	daily-use	NP
			绝路	tswɐt7.ɗɔ5	blind alley	NP
			别字	pɐt7.zi5	another character	NP
			佛像	ɦut7.tsʰjɐŋ5	Buddha statue	NP
			佛事	ɦut7.su5	Buddhist service	NP
			绿地	ɗik7.te5	green land	NP
			实在	sit7.tsɐj5	honestly	AdvP
			杂地	tsɐp7.te5	miscellaneous land	NP
			直系	tit7.ɦe5	direct line	NP
7	6	7+6	日出	zit7.tsʰut6	sunrise	NP
			直角	tit7.kɐk6	right angle	NP
			佛法	ɦut7.ɦwɐt6	Buddha doctrine	NP
			佛骨	ɦut7.kut6	Buddha bone	NP
			实质	sit7.tsit6	substance	NP
			佛国	ɦut7.kɔk6	Buddha country	NP
			直接	tit7.tsjɐp6	immediate	AP
7	7	7+7	日历	zit7.ɗit7	solar calendar	NP
			别日	pɐt7.zit7	another day	NP
			佛历	ɦut7.ɗit7	Buddha calendar	NP
			玉佛	ʛjɔk7.ɦut7	jade Buddha	NP
			实力	sit7.ɗik7	actual strength	NP
			落实	ɗɔk7.sit7	implement	VP
			直达	tit7.tɐt7	nonstop	AP
7	8	7+8	绿白	ɗik7.pɛ8	green in white	VP
			毒药	tɔk7.ʔjɵ8	toxicant	NP
8	1	8+1	白鸡	pɛ8.ke1	white chicken	NP
			白金	pɛ8.kim1	white gold	NP
			活猪	ʔwɐ8.ti1	live pig	NP
			药膏	ʔjɵ8.kɵ1	ointment	NP
8	2	8+2	白银	pɛ8.ʛin2	white silver	NP
			药材	ʔjɵ8.tsɐj2	medicinal material	NP
8	3	8+3	白马	pɛ8.ɓɛ3	white horse	NP
			白粉	pɛ8.ɦun3	white power	NP
			药酒	ʔjɵ8.tsju3	medical wine	NP

S1	S2	Pattern	Character	Example	Gloss	Phrase
8	4	8+4	白铁	pɛ8.tʰi4	white iron	NP
			白兔	pɛ8.tʰɔ4	white rabbit	NP
			腊肉	ɗɐ8.ɓɐ4	preserved meat	NP
			下课	ɗɵ8.kʰɵ4	finish class	VP
			药费	ʔjɵ8.ħwi4	medical expense	NP
8	5	8+5	石路	tsjɵ8.ɗɔ5	rocky road	NP
			白话	pɛ8.ʔwɐ5	empty promise	NP
			药量	ʔjɵ8.ɗjɐŋ5	medical dosage	NP
8	6	8+6	白色	pɛ8.sik6	white colour	AP
			药贴	ʔjɵ8.tʰjɐp6	medical sticker	NP
			白色	pɛ8.sik6	white	AP
8	7	8+7	白直	pɛ8.tit7	straightforward	NP
8	8	8+8	药物	ʔjɵ8.ɓwɵ8	medicine/drug	NP
			白白	pɛ8.pɛ8	white	AP
			薄薄	pɵ8.pɵ8	thin	AP
			辣辣	ɗwɐ8.ɗwɐ8	spicy	AP

Appendix B: Normalised and Statistical Data

Citation Tones

Table B1. Normalised F0 values of Zhangzhou citation tones across 10 sampling points

Tone	F0.10	F0.20	F0.30	F0.40	F0.50	F0.60	F0.70	F0.80	F0.90	F0.100
1	0.03	-0.03	0.02	0.15	0.33	0.56	0.82	1.09	1.31	1.47
2	-0.43	-0.58	-0.70	-0.78	-0.82	-0.83	-0.83	-0.82	-0.80	-0.79
3	2.10	2.10	1.98	1.73	1.33	0.87	0.32	-0.24	-0.76	-1.22
4	0.85	0.67	0.47	0.25	-0.01	-0.28	-0.58	-0.90	-1.20	-1.44
5	0.38	0.35	0.32	0.29	0.25	0.22	0.20	0.19	0.19	0.19
6	1.05	0.91	0.75	0.55	0.32	0.10	-0.15	-0.43	-0.75	-1.01
7	-0.45	-0.55	-0.62	-0.69	-0.74	-0.82	-0.92	-1.06	-1.19	-1.33
8	-0.38	-0.53	-0.65	-0.76	-0.84	-0.88	-0.90	-0.93	-0.96	-0.99

Table B2. Standard deviation values of the normalised F0 of Zhangzhou citation tones across 10 sampling points

Tone	F0.10	F0.20	F0.30	F0.40	F0.50	F0.60	F0.70	F0.80	F0.90	F0.100
1	0.39	0.36	0.40	0.46	0.54	0.61	0.65	0.68	0.70	0.75
2	0.38	0.32	0.31	0.33	0.38	0.38	0.37	0.37	0.38	0.38
3	0.63	0.58	0.57	0.55	0.61	0.66	0.65	0.69	0.66	0.63
4	0.56	0.49	0.46	0.46	0.44	0.45	0.47	0.46	0.48	0.53
5	0.35	0.34	0.35	0.36	0.37	0.37	0.37	0.37	0.36	0.37
6	0.57	0.51	0.51	0.52	0.54	0.55	0.60	0.66	0.71	0.78
7	0.43	0.37	0.33	0.33	0.37	0.39	0.39	0.42	0.51	0.61
8	0.49	0.38	0.35	0.35	0.35	0.33	0.31	0.31	0.34	0.42

Table B3. Result of pairwise *t*-test camparison of the normalised F0 values of Zhangzhou citation tones

Tone	T1.10	T1.100	T2.10	T2.100	T3.10	T3.100	T4.10	T4.100	T5.10	T5.100	T6.10	T6.100	T7.10	T7.100	T8.10
T1.100	0E+00	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T2.10	2E-31	0E+00	-	-	-	-	-	-	-	-	-	-	-	-	-
T2.100	9E-40	0E+00	1E+00	-	-	-	-	-	-	-	-	-	-	-	-
T3.10	0E+00	9E-20	0E+00	0E+00	-	-	-	-	-	-	-	-	-	-	-
T3.100	2E-163	0E+00	9E-49	1E-39	0E+00	-	-	-	-	-	-	-	-	-	-
T4.10	2E-111	5E-129	3E-239	9E-259	4E-244	0E+00	-	-	-	-	-	-	-	-	-
T4.100	9E-246	0E+00	1E-98	1E-85	0E+00	3E-09	0E+00	-	-	-	-	-	-	-	-
T5.10	2E-19	8E-273	3E-96	1E-109	0E+00	3E-281	4E-38	0E+00	-	-	-	-	-	-	-
T5.100	2E-33	2E-235	4E-122	6E-137	0E+00	151000379	6E-23	0E+00	4E-01	-	-	-	-	-	-
T6.10	7E-122	3E-59	2E-233	5E-250	1E-131	0E+00	4E-04	0E+00	5E-54	3E-38	-	-	-	-	-
T6.100	3E-97	0E+00	7E-23	2E-17	0E+00	4E-01	0E+00	6E-16	2E-181	4E-211	0E+00	-	-	-	-
T7.10	2E-30	0E+00	1E+00	1E+00	0E+00	4E-42	7E-223	1E-86	9E-91	9E-115	6E-222	7E-20	-	-	-
T7.100	1E-158	0E+00	2E-53	1E-44	0E+00	1E+00	0E+00	7E-03	5E-264	9E-300	0E+00	8E-04	3E-47	-	-
T8.10	6E-13	69545 1120	1E+00	1E+00	0E+00	1E-30	5E-120	5E-59	2E-45	6E-59	2E-134	7E-17	1E+00	1E-35	-
T8.100	9E-40	0E+00	3E-04	3E-02	0E+00	1E-07	3E-181	7E-25	2E-87	2E-105	1E-192	5E-02	2E-03	3E-11	3E-04

Table B4. Normalised duration values of Zhangzhou citation tones

Tone	Normalised	S.d
1	1.13	0.15
2	1.20	0.17
3	0.88	0.14
4	0.83	0.13
5	1.19	0.15
6	0.55	0.12
7	1.01	0.16
8	1.23	0.17

Table B5. Result of pairwise *t*-test camparison of the normalised duration values of Zhangzhou citation tones

```
      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
```

Phrase-initial Tones

Table B6. Normalised F0 values of phrase-initial tones across 64 tonal combinations in Zhangzhou

Pattern	F0.10	F0.20	F0.30	F0.40	F0.50	F0.60	F0.70	F0.80	F0.90	F0.100
1+1	0.21	0.13	0.1	0.11	0.12	0.14	0.16	0.18	0.19	0.17
1+2	0.34	0.27	0.27	0.28	0.31	0.34	0.37	0.4	0.42	0.4
1+3	0.11	0.03	0.01	0.01	0.02	0.02	0.03	0.04	0.05	0.06
1+4	0.34	0.24	0.2	0.18	0.17	0.18	0.19	0.2	0.21	0.18
1+5	0.2	0.13	0.1	0.1	0.1	0.1	0.11	0.11	0.11	0.09
1+6	0.3	0.2	0.16	0.13	0.13	0.13	0.13	0.13	0.13	0.11
1+7	0.39	0.33	0.31	0.32	0.34	0.37	0.41	0.44	0.45	0.42
1+8	0.29	0.29	0.3	0.31	0.33	0.35	0.39	0.43	0.45	0.42
2+1	0.12	0.08	0.06	0.05	0.05	0.05	0.06	0.08	0.07	0.05
2+2	0.2	0.2	0.22	0.25	0.27	0.29	0.32	0.35	0.35	0.33
2+3	0.07	0	-0.02	-0.02	-0.01	0	0.02	0.04	0.06	0.07
2+4	0.2	0.17	0.16	0.16	0.17	0.17	0.19	0.2	0.2	0.17
2+5	0.11	0.09	0.09	0.09	0.09	0.09	0.1	0.1	0.1	0.07
2+6	0.21	0.19	0.17	0.16	0.16	0.16	0.16	0.16	0.14	0.12
2+7	0.4	0.38	0.39	0.4	0.42	0.44	0.47	0.5	0.5	0.47

Pattern	F0.10	F0.20	F0.30	F0.40	F0.50	F0.60	F0.70	F0.80	F0.90	F0.100
2+8	0.23	0.2	0.2	0.22	0.24	0.26	0.29	0.33	0.36	0.33
3+1	-0.1	-0.09	-0.01	0.12	0.29	0.48	0.67	0.83	0.97	1.02
3+2	0.02	0.01	0.09	0.24	0.45	0.69	0.97	1.22	1.42	1.51
3+3	-0.3	-0.29	-0.21	-0.09	0.06	0.23	0.41	0.57	0.7	0.77
3+4	-0.16	-0.16	-0.07	0.06	0.23	0.45	0.68	0.92	1.09	1.16
3+5	-0.12	-0.13	-0.05	0.1	0.27	0.46	0.65	0.82	0.93	0.99
3+6	-0.06	-0.07	0	0.12	0.29	0.49	0.69	0.86	0.98	1.04
3+7	-0.18	-0.13	-0.01	0.18	0.42	0.69	0.97	1.23	1.44	1.54
3+8	-0.23	-0.18	-0.06	0.13	0.36	0.62	0.92	1.22	1.43	1.56
4+1	1.76	1.67	1.56	1.42	1.25	1.05	0.82	0.6	0.38	0.16
4+2	2.13	2.09	2.04	1.92	1.75	1.52	1.26	0.98	0.7	0.45
4+3	1.84	1.74	1.62	1.45	1.24	1.02	0.78	0.55	0.35	0.21
4+4	2.01	1.95	1.86	1.72	1.51	1.27	1	0.71	0.45	0.22
4+5	1.89	1.8	1.67	1.51	1.31	1.08	0.85	0.63	0.43	0.26
4+6	1.91	1.85	1.77	1.63	1.44	1.21	0.96	0.67	0.41	0.17
4+7	1.96	1.94	1.9	1.78	1.59	1.36	1.09	0.8	0.51	0.27
4+8	2.33	2.27	2.17	2.03	1.83	1.59	1.3	1.03	0.72	0.45
5+1	0.18	0.05	-0.06	-0.17	-0.28	-0.39	-0.49	-0.58	-0.66	-0.72
5+2	0.42	0.33	0.24	0.13	0	-0.13	-0.27	-0.4	-0.51	-0.61
5+3	0.03	-0.1	-0.2	-0.3	-0.41	-0.52	-0.61	-0.68	-0.73	-0.77
5+4	0.37	0.26	0.15	0.03	-0.12	-0.27	-0.43	-0.57	-0.69	-0.79
5+5	0.15	0.03	-0.08	-0.19	-0.31	-0.42	-0.52	-0.6	-0.66	-0.7
5+6	0.25	0.16	0.07	-0.03	-0.16	-0.31	-0.46	-0.59	-0.69	-0.78
5+7	0.36	0.31	0.25	0.16	0.05	-0.08	-0.2	-0.31	-0.42	-0.52
5+8	0.27	0.24	0.18	0.09	-0.02	-0.15	-0.28	-0.41	-0.51	-0.62
6+1	1.86	1.81	1.76	1.7	1.64	1.57	1.5	1.4	1.33	1.31
6+2	1.98	1.94	1.91	1.87	1.84	1.8	1.75	1.7	1.66	1.61
6+3	1.66	1.61	1.55	1.48	1.43	1.38	1.32	1.26	1.21	1.2
6+4	1.81	1.77	1.73	1.68	1.64	1.6	1.55	1.51	1.46	1.42
6+5	1.86	1.81	1.75	1.7	1.65	1.6	1.54	1.47	1.43	1.4
6+6	2.03	1.94	1.85	1.75	1.66	1.57	1.47	1.38	1.3	1.25
6+7	2.18	2.18	2.17	2.15	2.12	2.07	2.02	1.95	1.9	1.86
6+8	2	1.92	1.87	1.83	1.79	1.76	1.75	1.72	1.67	1.66
7+1	0.45	0.34	0.25	0.17	0.09	0.01	-0.06	-0.13	-0.2	-0.25
7+2	0.46	0.43	0.38	0.34	0.28	0.23	0.17	0.1	0.04	0
7+3	0.16	0.09	0.04	-0.02	-0.08	-0.14	-0.21	-0.28	-0.34	-0.4
7+4	0.28	0.24	0.19	0.15	0.1	0.05	-0.01	-0.07	-0.15	-0.16
7+5	0.32	0.25	0.17	0.1	0.04	-0.04	-0.11	-0.18	-0.26	-0.31
7+6	0.39	0.31	0.25	0.18	0.11	0.03	-0.05	-0.13	-0.19	-0.24

Pattern	F0.10	F0.20	F0.30	F0.40	F0.50	F0.60	F0.70	F0.80	F0.90	F0.100
7+7	0.45	0.39	0.34	0.29	0.24	0.24	0.12	0.05	-0.01	-0.05
7+8	0.48	0.43	0.39	0.34	0.28	0.23	0.16	0.08	0.03	-0.03
8+1	0.24	0.12	0.01	-0.09	-0.21	-0.34	-0.46	-0.57	-0.66	-0.74
8+2	0.54	0.42	0.29	0.19	0.07	-0.06	-0.2	-0.32	-0.44	-0.54
8+3	0.06	-0.09	-0.21	-0.34	-0.47	-0.6	-0.69	-0.75	-0.8	-0.84
8+4	0.31	0.2	0.09	-0.02	-0.16	-0.31	-0.47	-0.61	-0.72	-0.82
8+5	0.26	0.12	-0.01	-0.13	-0.26	-0.4	-0.51	-0.6	-0.67	-0.71
8+6	0.37	0.24	0.14	0.02	-0.12	-0.27	-0.42	-0.56	-0.7	-0.77
8+7	0.38	0.26	0.17	0.07	-0.04	-0.18	-0.31	-0.42	-0.51	-0.6
8+8	0.48	0.39	0.3	0.2	0.08	-0.05	-0.18	-0.29	-0.4	-0.49

Table B7. Standard deviation values of the normalised F0 of Zhangzhou phrase-initial tones across 10 sampling points across 64 disyllabic combinations

Pattern	F0.10	F0.20	F0.30	F0.40	F0.50	F0.60	F0.70	F0.80	F0.90	F0.100
1+1	0.4	0.38	0.37	0.37	0.37	0.38	0.38	0.38	0.38	0.38
1+2	0.4	0.34	0.34	0.34	0.34	0.35	0.37	0.37	0.37	0.37
1+3	0.35	0.32	0.31	0.31	0.31	0.31	0.32	0.33	0.33	0.33
1+4	0.38	0.34	0.34	0.34	0.34	0.35	0.35	0.35	0.34	0.33
1+5	0.34	0.32	0.32	0.31	0.31	0.31	0.31	0.31	0.31	0.31
1+6	0.38	0.31	0.29	0.28	0.28	0.29	0.3	0.3	0.31	0.31
1+7	0.42	0.38	0.37	0.38	0.38	0.39	0.4	0.4	0.39	0.37
1+8	0.33	0.3	0.33	0.35	0.36	0.38	0.39	0.41	0.41	0.39
2+1	0.34	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.3	0.3
2+2	0.36	0.33	0.32	0.32	0.33	0.33	0.34	0.34	0.34	0.35
2+3	0.34	0.3	0.29	0.28	0.28	0.28	0.28	0.29	0.29	0.3
2+4	0.33	0.3	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29
2+5	0.32	0.29	0.29	0.29	0.28	0.28	0.28	0.28	0.28	0.29
2+6	0.29	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.28
2+7	0.35	0.32	0.32	0.33	0.33	0.34	0.34	0.34	0.34	0.37
2+8	0.37	0.31	0.3	0.3	0.3	0.3	0.3	0.31	0.31	0.32
3+1	0.37	0.33	0.35	0.39	0.43	0.45	0.45	0.42	0.39	0.41
3+2	0.4	0.37	0.41	0.46	0.5	0.54	0.55	0.55	0.53	0.54
3+3	0.37	0.35	0.38	0.42	0.45	0.5	0.51	0.5	0.49	0.47
3+4	0.42	0.38	0.39	0.42	0.46	0.5	0.52	0.53	0.55	0.54
3+5	0.4	0.36	0.37	0.36	0.39	0.42	0.45	0.43	0.43	0.43
3+6	0.38	0.33	0.35	0.38	0.42	0.44	0.46	0.45	0.42	0.42
3+7	0.4	0.35	0.36	0.41	0.43	0.46	0.48	0.51	0.5	0.51
3+8	0.33	0.3	0.34	0.4	0.47	0.52	0.55	0.53	0.53	0.57
4+1	0.6	0.52	0.47	0.44	0.43	0.43	0.46	0.52	0.68	0.64

Pattern	F0.10	F0.20	F0.30	F0.40	F0.50	F0.60	F0.70	F0.80	F0.90	F0.100
4+2	0.61	0.54	0.5	0.47	0.45	0.45	0.48	0.53	0.59	0.65
4+3	0.61	0.53	0.49	0.48	0.49	0.51	0.53	0.56	0.59	0.63
4+4	0.63	0.59	0.55	0.52	0.52	0.56	0.6	0.64	0.68	0.75
4+5	0.5	0.44	0.41	0.39	0.39	0.42	0.44	0.47	0.48	0.51
4+6	0.63	0.56	0.52	0.49	0.48	0.49	0.51	0.56	0.6	0.68
4+7	0.73	0.7	0.65	0.6	0.56	0.55	0.56	0.63	0.71	0.77
4+8	0.61	0.54	0.53	0.52	0.52	0.51	0.52	0.53	0.57	0.65
5+1	0.5	0.43	0.39	0.37	0.36	0.35	0.35	0.36	0.39	0.42
5+2	0.45	0.37	0.34	0.32	0.32	0.33	0.34	0.36	0.36	0.37
5+3	0.47	0.42	0.4	0.38	0.37	0.35	0.35	0.31	0.32	0.34
5+4	0.54	0.45	0.4	0.38	0.36	0.35	0.33	0.34	0.34	0.37
5+5	0.44	0.36	0.32	0.3	0.3	0.3	0.31	0.31	0.32	0.33
5+6	0.47	0.42	0.38	0.46	0.41	0.34	0.33	0.36	0.36	0.41
5+7	0.51	0.43	0.39	0.37	0.36	0.36	0.36	0.37	0.4	0.44
5+8	0.4	0.37	0.34	0.32	0.3	0.29	0.29	0.3	0.33	0.34
6+1	0.54	0.51	0.49	0.46	0.44	0.43	0.45	0.5	0.56	0.61
6+2	0.83	0.83	0.83	0.83	0.83	0.82	0.83	0.83	0.83	0.85
6+3	0.57	0.55	0.53	0.51	0.5	0.47	0.46	0.44	0.43	0.48
6+4	0.62	0.59	0.57	0.57	0.53	0.49	0.47	0.46	0.46	0.49
6+5	0.53	0.51	0.49	0.48	0.47	0.49	0.46	0.48	0.48	0.48
6+6	0.5	0.48	0.46	0.47	0.49	0.52	0.57	0.64	0.7	0.75
6+7	0.85	0.82	0.8	0.8	0.79	0.8	0.81	0.83	0.84	0.85
6+8	0.69	0.56	0.52	0.5	0.48	0.47	0.46	0.47	0.47	0.48
7+1	0.74	0.58	0.56	0.54	0.51	0.49	0.49	0.49	0.52	0.52
7+2	0.6	0.57	0.55	0.54	0.52	0.51	0.51	0.51	0.52	0.53
7+3	0.66	0.58	0.6	0.58	0.57	0.56	0.55	0.54	0.54	0.56
7+4	0.46	0.41	0.37	0.34	0.34	0.34	0.34	0.33	0.34	0.46
7+5	0.56	0.51	0.47	0.46	0.49	0.45	0.47	0.47	0.48	0.49
7+6	0.58	0.53	0.51	0.51	0.49	0.48	0.47	0.47	0.49	0.51
7+7	0.5	0.46	0.43	0.41	0.4	0.75	0.39	0.41	0.4	0.41
7+8	0.66	0.56	0.52	0.49	0.47	0.46	0.47	0.49	0.51	0.55
8+1	0.44	0.36	0.32	0.31	0.29	0.29	0.29	0.3	0.32	0.36
8+2	0.83	0.71	0.51	0.49	0.48	0.47	0.46	0.43	0.42	0.41
8+3	0.48	0.42	0.39	0.37	0.37	0.35	0.35	0.35	0.35	0.43
8+4	0.46	0.42	0.4	0.38	0.38	0.36	0.34	0.34	0.33	0.36
8+5	0.48	0.39	0.36	0.36	0.34	0.34	0.33	0.32	0.35	0.34
8+6	0.45	0.36	0.33	0.3	0.3	0.29	0.3	0.3	0.32	0.34
8+7	0.41	0.32	0.26	0.24	0.25	0.24	0.25	0.27	0.32	0.35
8+8	0.62	0.54	0.49	0.47	0.47	0.46	0.47	0.47	0.49	0.51

Table B8. Result of pairwise *t*-test camparison of the normalised F0 values of Zhangzhou phrase-initial tones across 64 tonal combinations at 10% sampling point

	1+1	1+2	1+3	1+4	1+5	1+6	1+7
1+2	0.00280	-	-	-	-	-	-
1+3	0.10967	4.3e-10	-	-	-	-	-
1+4	0.00175	1.00000	7.7e-11	-	-	-	-
1+5	1.00000	0.00087	0.25397	0.00050	-	-	-
1+6	0.42535	1.00000	1.0e-05	1.00000	0.20437	-	-
1+7	2.1e-05	1.00000	9.9e-13	1.00000	5.5e-06	0.49753	-
1+8	1.00000	1.00000	0.02632	1.00000	1.00000	1.00000	1.00000

	2+1	2+2	2+3	2+4	2+5	2+6	2+7
2+2	0.36933	-	-	-	-	-	-
2+3	1.00000	0.00084	-	-	-	-	-
2+4	0.23034	1.00000	0.00026	-	-	-	-
2+5	1.00000	0.10686	1.00000	0.05805	-	-	-
2+6	0.18473	1.00000	0.00050	1.00000	0.05399	-	-
2+7	1.8e-14	1.1e-07	< 2e-16	2.6e-08	9.2e-16	6.5e-06	-
2+8	0.40791	1.00000	0.00741	1.00000	0.17332	1.00000	0.00466

	3+1	3+2	3+3	3+4	3+5	3+6	3+7
3+2	0.02204	-	-	-	-	-	-
3+3	4.9e-07	< 2e-16	-	-	-	-	-
3+4	1.00000	2.4e-06	0.00039	-	-	-	-
3+5	1.00000	0.00106	1.0e-05	1.00000	-	-	-
3+6	1.00000	0.96709	4.4e-09	0.17862	1.00000	-	-
3+7	1.00000	3.5e-06	0.03602	1.00000	1.00000	0.07809	-
3+8	0.52369	0.00016	1.00000	1.00000	1.00000	0.07431	1.00000

	4+1	4+2	4+3	4+4	4+5	4+6	4+7
4+2	9.9e-14	-	-	-	-	-	-
4+3	1.00000	3.7e-08	-	-	-	-	-
4+4	5.7e-07	0.15680	0.00639	-	-	-	-
4+5	0.18845	7.8e-06	1.00000	0.19560	-	-	-
4+6	0.09811	0.00068	1.00000	1.00000	1.00000	-	-
4+7	0.00467	0.04291	0.78870	1.00000	1.00000	1.00000	-
4+8	2.9e-15	0.14695	4.8e-11	5.1e-05	4.1e-09	2.2e-07	1.7e-05

	5+1	5+2	5+3	5+4	5+5	5+6	5+7
5+2	2.0e-06	-	-	-	-	-	-
5+3	0.00909	< 2e-16	-	-	-	-	-
5+4	0.00012	1.00000	3.6e-15	-	-	-	-
5+5	1.00000	4.8e-08	0.09127	3.7e-06	-	-	-
5+6	1.00000	0.00647	2.6e-05	0.12574	0.96214	-	-
5+7	0.00353	1.00000	2.9e-11	1.00000	0.00026	0.60833	-
5+8	1.00000	1.00000	0.00702	1.00000	1.00000	1.00000	1.00000

	6+1	6+2	6+3	6+4	6+5	6+6	6+7
6+2	1.00000	-	-	-	-	-	-
6+3	0.17456	0.00016	-	-	-	-	-
6+4	1.00000	0.27945	0.83816	-	-	-	-
6+5	1.00000	1.00000	0.17365	1.00000	-	-	-
6+6	0.95667	1.00000	2.5e-05	0.05643	0.62133	-	-
6+7	0.00050	0.08732	1.4e-11	1.1e-06	0.00015	0.94291	-
6+8	1.00000	1.00000	0.08737	1.00000	1.00000	1.00000	1.00000

	7+1	7+2	7+3	7+4	7+5	7+6	7+7
7+2	1.00000	-	-	-	-	-	-
7+3	0.00011	9.6e-06	-	-	-	-	-
7+4	0.18715	0.05785	1.00000	-	-	-	-
7+5	0.91128	0.37451	0.27830	1.00000	-	-	-
7+6	1.00000	1.00000	0.00959	1.00000	1.00000	-	-
7+7	1.00000	1.00000	0.00033	0.30824	1.00000	1.00000	-
7+8	1.00000	1.00000	0.04900	1.00000	1.00000	1.00000	1.00000

	8+1	8+2	8+3	8+4	8+5	8+6	8+7
8+2	0.07514	-	-	-	-	-	-
8+3	1.00000	0.00016	-	-	-	-	-
8+4	1.00000	0.35960	0.07205	-	-	-	-
8+5	1.00000	0.20186	1.00000	1.00000	-	-	-
8+6	1.00000	1.00000	0.08931	1.00000	1.00000	-	-
8+7	1.00000	1.00000	0.02844	1.00000	1.00000	1.00000	-
8+8	0.04112	1.00000	1.1e-05	0.28194	0.19475	1.00000	1.00000

Table B9. Result of pairwise *t*-test camparison of the normalised F0 values of Zhangzhou phrase-initial tones across 64 tonal combinations at 100% sampling point

	1+1	1+2	1+3	1+4	1+5	1+6	1+7
1+2	7.7e-13	-	-	-	-	-	-
1+3	0.00901	< 2e-16	-	-	-	-	-
1+4	1.00000	3.4e-12	0.00041	-	-	-	-
1+5	0.26096	< 2e-16	1.00000	0.02813	-	-	-
1+6	1.00000	< 2e-16	1.00000	0.57707	1.00000	-	-
1+7	5.3e-13	1.00000	< 2e-16	2.5e-12	< 2e-16	< 2e-16	-
1+8	5.7e-06	1.00000	2.9e-12	2.1e-05	2.6e-10	2.1e-08	1.00000

	2+1	2+2	2+3	2+4	2+5	2+6	2+7
2+2	< 2e-16	-	-	-	-	-	-
2+3	1.00000	< 2e-16	-	-	-	-	-
2+4	0.00011	5.7e-08	0.00231	-	-	-	-
2+5	1.00000	< 2e-16	1.00000	0.00171	-	-	-
2+6	0.77501	3.0e-11	1.00000	1.00000	1.00000	-	-
2+7	< 2e-16	0.00016	< 2e-16	< 2e-16	< 2e-16	< 2e-16	-
2+8	2.1e-11	1.00000	5.3e-10	0.00075	3.9e-10	3.3e-06	0.02278

	3+1	3+2	3+3	3+4	3+5	3+6	3+7
3+2	< 2e-16	-	-	-	-	-	-
3+3	4.6e-07	< 2e-16	-	-	-	-	-
3+4	0.01330	1.0e-15	< 2e-16	-	-	-	-
3+5	1.00000	< 2e-16	6.8e-06	0.00063	-	-	-
3+6	1.00000	< 2e-16	1.8e-07	0.13547	1.00000	-	-
3+7	< 2e-16	1.00000	< 2e-16	1.6e-15	< 2e-16	< 2e-16	-
3+8	2.0e-13	1.00000	< 2e-16	1.3e-07	5.0e-15	7.4e-12	1.00000

	4+1	4+2	4+3	4+4	4+5	4+6	4+7
4+2	5.7e-07	-	-	-	-	-	-
4+3	1.00000	0.00012	-	-	-	-	-
4+4	1.00000	8.8e-05	1.00000	-	-	-	-
4+5	1.00000	0.00717	1.00000	1.00000	-	-	-
4+6	1.00000	3.8e-05	1.00000	1.00000	1.00000	-	-
4+7	1.00000	0.08034	1.00000	1.00000	1.00000	1.00000	-
4+8	0.00199	1.00000	0.02869	0.03208	0.23914	0.00841	0.63238

	5+1	5+2	5+3	5+4	5+5	5+6	5+7
5+2	0.02713	-	-	-	-	-	-
5+3	1.00000	0.00011	-	-	-	-	-
5+4	1.00000	1.7e-06	1.00000	-	-	-	-
5+5	1.00000	0.20514	1.00000	0.19487	-	-	-
5+6	1.00000	4.7e-05	1.00000	1.00000	0.65333	-	-
5+7	9.9e-07	0.45833	4.5e-10	1.3e-12	2.2e-05	2.5e-10	-
5+8	1.00000	1.00000	0.16062	0.03785	1.00000	0.08617	1.00000

	6+1	6+2	6+3	6+4	6+5	6+6	6+7
6+2	0.0011	-	-	-	-	-	-
6+3	1.0000	7.6e-08	-	-	-	-	-
6+4	1.0000	0.0967	0.0277	-	-	-	-
6+5	1.0000	0.0658	0.1199	1.0000	-	-	-
6+6	1.0000	1.3e-05	1.0000	0.4113	1.0000	-	-
6+7	1.5e-11	0.0141	< 2e-16	5.9e-09	8.6e-09	1.7e-14	-
6+8	0.0834	1.0000	0.0014	0.9478	0.6854	0.0110	1.0000

	7+1	7+2	7+3	7+4	7+5	7+6	7+7
7+2	8.0e-05	-	-	-	-	-	-
7+3	0.10529	1.4e-13	-	-	-	-	-
7+4	1.00000	0.04127	0.00012	-	-	-	-
7+5	1.00000	6.1e-08	1.00000	0.16006	-	-	-
7+6	1.00000	0.00028	0.11201	1.00000	1.00000	-	-
7+7	0.01714	1.00000	1.0e-08	1.00000	0.00013	0.03355	-
7+8	0.33472	1.00000	0.00041	1.00000	0.03489	0.42793	1.00000

	8+1	8+2	8+3	8+4	8+5	8+6	8+7
8+2	0.25213	-	-	-	-	-	-
8+3	1.00000	0.00513	-	-	-	-	-
8+4	1.00000	0.00337	1.00000	-	-	-	-
8+5	1.00000	0.98430	1.00000	1.00000	-	-	-
8+6	1.00000	0.22373	1.00000	1.00000	1.00000	-	-
8+7	1.00000	1.00000	0.02453	0.01589	1.00000	0.88972	-
8+8	0.00027	1.00000	4.7e-07	4.2e-09	0.00995	0.00192	1.00000

Table B10. Normalised duration values of Zhangzhou phrase-initial tones across phrase-final tones

Tone	V1	V2	V3	V4	V5	V6	V7	V8
1	1.04	1.05	0.95	1.01	1.04	0.96	1.00	1.08
2	1.04	1.08	0.98	0.97	1.01	0.91	0.99	1.08
3	0.98	1.03	0.98	0.98	1.03	0.90	0.99	1.01
4	0.93	0.96	0.90	0.90	1.01	0.82	0.98	0.90
5	0.97	1.03	0.92	0.97	1.03	0.88	1.00	1.10
6	0.35	0.34	0.36	0.33	0.36	0.34	0.40	0.40
7	0.35	0.41	0.36	0.35	0.42	0.36	0.46	0.42
8	1.03	1.04	1.02	0.98	1.10	0.94	1.01	0.98

Table B11. Standard deviation values of the normalised duration of Zhangzhou phrase-initial tones across phrase-final tones

Tone	V1	V2	V3	V4	V5	V6	V7	V8
1	22.25	19.88	19.50	17.16	21.04	19.83	20.81	24.41
2	25.30	22.62	21.00	19.56	22.39	24.58	20.15	19.66
3	17.83	18.20	19.82	18.44	19.35	17.68	18.98	17.85
4	19.11	19.45	19.68	18.66	21.34	18.06	18.51	18.03
5	19.36	19.90	19.21	17.76	21.34	18.48	19.26	18.30
6	9.98	8.40	9.96	7.90	9.13	8.61	11.49	12.00
7	9.73	14.90	11.47	8.98	15.83	9.87	14.29	9.71
8	19.31	17.02	17.54	16.64	22.14	15.23	17.87	21.87

Table B12. Result of pairwise *t*-test camparison of the normalised duration values of Zhangzhou phrase-initial tones across 64 tonal combinations

	1+1	1+2	1+3	1+4	1+5	1+6	1+7
1+2	1.00000	-	-	-	-	-	-
1+3	1.2e-05	2.6e-07	-	-	-	-	-
1+4	1.00000	0.70891	0.00399	-	-	-	-
1+5	1.00000	1.00000	2.7e-05	1.00000	-	-	-
1+6	0.00153	8.4e-05	1.00000	0.13630	0.00284	-	-
1+7	1.00000	0.21005	0.21564	1.00000	1.00000	1.00000	-
1+8	1.00000	1.00000	0.00019	0.68718	1.00000	0.00246	0.26150

	2+1	2+2	2+3	2+4	2+5	2+6	2+7
2+2	0.66699	-	-	-	-	-	-
2+3	0.14305	1.3e-05	-	-	-	-	-
2+4	0.00305	1.5e-08	1.00000	-	-	-	-
2+5	1.00000	0.01339	1.00000	0.29064	-	-	-
2+6	6.1e-08	3.0e-14	0.01830	0.18437	3.6e-05	-	-
2+7	0.45798	0.00029	1.00000	1.00000	1.00000	0.03894	-
2+8	1.00000	1.00000	0.00869	0.00041	0.34063	7.5e-08	0.02490

	3+1	3+2	3+3	3+4	3+5	3+6	3+7
3+2	0.40940	-	-	-	-	-	-
3+3	1.00000	0.07103	-	-	-	-	-
3+4	1.00000	0.15963	1.00000	-	-	-	-
3+5	0.24110	1.00000	0.03743	0.08489	-	-	-
3+6	0.00013	1.0e-10	0.00085	5.4e-05	3.0e-11	-	-
3+7	1.00000	0.92037	1.00000	1.00000	0.58634	0.00017	-
3+8	1.00000	1.00000	1.00000	1.00000	1.00000	0.00141	1.00000

	4+1	4+2	4+3	4+4	4+5	4+6	4+7
4+2	1.00000	-	-	-	-	-	-
4+3	0.67654	0.00378	-	-	-	-	-
4+4	1.3e-07	5.4e-12	0.01592	-	-	-	-
4+5	2.4e-05	0.04762	5.6e-11	< 2e-16	-	-	-
4+6	9.0e-07	2.1e-10	0.01456	1.00000	< 2e-16	-	-
4+7	0.43623	1.00000	0.00043	3.0e-12	1.00000	4.5e-11	-
4+8	1.00000	0.19467	1.00000	0.63861	2.4e-05	0.36281	0.03489

	5+1	5+2	5+3	5+4	5+5	5+6	5+7
5+2	0.03862	-	-	-	-	-	-
5+3	0.25024	2.4e-07	-	-	-	-	-
5+4	1.00000	0.01804	0.25003	-	-	-	-
5+5	0.02512	1.00000	9.4e-08	0.01104	-	-	-
5+6	4.5e-05	2.8e-13	0.57584	3.2e-05	7.1e-14	-	-
5+7	1.00000	1.00000	0.00094	1.00000	1.00000	2.7e-08	-
5+8	9.1e-05	0.26478	9.0e-09	4.9e-05	0.29098	5.9e-13	0.02128

	6+1	6+2	6+3	6+4	6+5	6+6	6+7
6+2	1.00000	-	-	-	-	-	-
6+3	1.00000	1.00000	-	-	-	-	-
6+4	1.00000	1.00000	1.00000	-	-	-	-
6+5	1.00000	1.00000	1.00000	1.00000	-	-	-
6+6	2.7e-07	2.4e-10	2.0e-07	1.9e-11	2.8e-07	-	-
6+7	0.01512	0.00022	0.01636	4.6e-05	0.02037	0.36153	-
6+8	1.00000	0.24316	1.00000	0.15238	1.00000	1.00000	1.00000

	7+1	7+2	7+3	7+4	7+5	7+6	7+7
7+2	0.00021	-	-	-	-	-	-
7+3	1.00000	0.01398	-	-	-	-	-
7+4	1.00000	4.6e-05	1.00000	-	-	-	-
7+5	1.2e-06	1.00000	0.00016	1.7e-07	-	-	-
7+6	1.00000	0.06804	1.00000	1.00000	0.00149	-	-
7+7	1.5e-13	0.00193	6.6e-11	8.2e-15	0.13488	3.7e-09	-
7+8	0.03146	1.00000	0.24524	0.01916	1.00000	0.43016	1.00000

	8+1	8+2	8+3	8+4	8+5	8+6	8+7
8+2	1.00000	-	-	-	-	-	-
8+3	1.00000	1.00000	-	-	-	-	-
8+4	1.00000	1.00000	1.00000	-	-	-	-
8+5	0.65930	1.00000	0.42300	0.00041	-	-	-
8+6	0.41961	0.68294	1.00000	1.00000	0.00081	-	-
8+7	1.00000	1.00000	1.00000	1.00000	0.13203	1.00000	-
8+8	0.92078	1.00000	1.00000	1.00000	0.00036	1.00000	1.00000

Phrase-final Tones

Table B13. Normalised F0 values of phrase-final tones across 64 tonal combinations in Zhangzhou

Pattern	F0.10	F0.20	F0.30	F0.40	F0.50	F0.60	F0.70	F0.80	F0.90	F0.100
1+1	0.04	-0.10	-0.12	-0.06	0.04	0.19	0.36	0.54	0.68	0.79
2+1	-0.10	-0.22	-0.23	-0.18	-0.08	0.06	0.21	0.36	0.49	0.59
3+1	-0.05	-0.31	-0.40	-0.40	-0.34	-0.25	-0.13	-0.01	0.10	0.19
4+1	-0.30	-0.44	-0.48	-0.46	-0.38	-0.26	-0.13	0.01	0.14	0.23
5+1	-0.39	-0.43	-0.41	-0.30	-0.15	0.04	0.26	0.47	0.63	0.77
6+1	-0.27	-0.44	-0.49	-0.48	-0.41	-0.29	-0.15	-0.02	0.09	0.17
7+1	-0.25	-0.35	-0.36	-0.28	-0.15	0.03	0.25	0.46	0.65	0.79
8+1	-0.30	-0.34	-0.32	-0.23	-0.09	0.09	0.31	0.52	0.68	0.77
1+2	-0.53	-0.79	-0.99	-1.11	-1.21	-1.26	-1.26	-1.26	-1.27	-1.27
2+2	-0.67	-0.92	-1.11	-1.23	-1.31	-1.32	-1.32	-1.33	-1.34	-1.35
3+2	-0.34	-0.67	-0.93	-1.11	-1.23	-1.30	-1.32	-1.34	-1.35	-1.36

Pattern	F0.10	F0.20	F0.30	F0.40	F0.50	F0.60	F0.70	F0.80	F0.90	F0.100
4+2	-0.82	-1.02	-1.17	-1.26	-1.36	-1.40	-1.41	-1.41	-1.41	-1.41
5+2	-0.91	-1.09	-1.21	-1.28	-1.34	-1.36	-1.35	-1.34	-1.35	-1.37
6+2	-0.73	-0.96	-1.11	-1.23	-1.31	-1.37	-1.37	-1.39	-1.40	-1.39
7+2	-0.83	-1.02	-1.13	-1.20	-1.26	-1.28	-1.29	-1.27	-1.28	-1.27
8+2	-0.89	-1.03	-1.12	-1.16	-1.20	-1.23	-1.21	-1.20	-1.18	-1.16
1+3	1.43	1.45	1.38	1.20	0.90	0.52	0.08	-0.38	-0.84	-1.20
2+3	1.30	1.34	1.27	1.10	0.81	0.44	0.02	-0.43	-0.87	-1.25
3+3	1.53	1.42	1.26	1.03	0.72	0.37	-0.05	-0.48	-0.84	-1.17
4+3	1.25	1.22	1.10	0.91	0.64	0.30	-0.11	-0.52	-0.89	-1.21
5+3	1.33	1.45	1.43	1.25	0.93	0.54	0.08	-0.39	-0.79	-1.16
6+3	1.60	1.50	1.34	1.12	0.84	0.47	0.03	-0.40	-0.82	-1.19
7+3	1.58	1.63	1.58	1.39	1.08	0.70	0.24	-0.28	-0.75	-1.16
8+3	1.17	1.31	1.33	1.20	0.97	0.61	0.21	-0.22	-0.64	-1.03
1+4	0.38	0.19	-0.01	-0.22	-0.44	-0.69	-0.92	-1.12	-1.32	-1.50
2+4	0.37	0.15	-0.06	-0.27	-0.51	-0.73	-0.96	-1.16	-1.36	-1.54
3+4	0.79	0.51	0.23	-0.03	-0.31	-0.59	-0.85	-1.11	-1.34	-1.53
4+4	0.17	-0.01	-0.21	-0.40	-0.59	-0.80	-0.99	-1.19	-1.37	-1.52
5+4	0.04	-0.12	-0.27	-0.45	-0.63	-0.83	-1.02	-1.22	-1.39	-1.54
6+4	0.67	0.36	0.07	-0.21	-0.46	-0.70	-0.95	-1.19	-1.39	-1.52
7+4	0.23	0.08	-0.12	-0.31	-0.52	-0.74	-0.97	-1.17	-1.37	-1.54
8+4	0.10	-0.02	-0.15	-0.30	-0.50	-0.69	-0.92	-1.14	-1.33	-1.46
1+5	0.12	0.03	-0.02	-0.05	-0.07	-0.09	-0.09	-0.09	-0.09	-0.09
2+5	0.13	0.03	-0.03	-0.07	-0.10	-0.11	-0.11	-0.12	-0.11	-0.10
3+5	0.25	0.04	-0.08	-0.16	-0.22	-0.25	-0.27	-0.27	-0.27	-0.27
4+5	-0.08	-0.18	-0.24	-0.28	-0.32	-0.34	-0.35	-0.35	-0.34	-0.34
5+5	-0.16	-0.19	-0.21	-0.23	-0.26	-0.27	-0.27	-0.27	-0.26	-0.25
6+5	0.16	-0.04	-0.15	-0.23	-0.28	-0.32	-0.34	-0.35	-0.35	-0.34
7+5	0.03	-0.05	-0.11	-0.15	-0.18	-0.19	-0.20	-0.20	-0.20	-0.19
8+5	-0.31	-0.27	-0.26	-0.28	-0.32	-0.35	-0.36	-0.35	-0.35	-0.34
1+6	0.55	0.39	0.21	0.01	-0.22	-0.48	-0.77	-1.07	-1.35	-1.59
2+6	0.67	0.48	0.26	0.04	-0.20	-0.45	-0.73	-1.03	-1.31	-1.53
3+6	0.87	0.65	0.42	0.17	-0.08	-0.39	-0.71	-1.05	-1.34	-1.59
4+6	0.41	0.26	0.08	-0.11	-0.32	-0.55	-0.82	-1.09	-1.34	-1.56
5+6	0.29	0.18	0.03	-0.15	-0.37	-0.61	-0.88	-1.16	-1.42	-1.64
6+6	0.60	0.39	0.17	-0.08	-0.32	-0.58	-0.85	-1.19	-1.45	-1.70
7+6	0.45	0.31	0.18	0.01	-0.21	-0.47	-0.74	-1.02	-1.31	-1.59
8+6	0.43	0.32	0.17	-0.03	-0.26	-0.51	-0.82	-1.11	-1.41	-1.68
1+7	-0.62	-0.81	-1.00	-1.13	-1.25	-1.34	-1.39	-1.46	-1.53	-1.59
2+7	-0.63	-0.84	-1.02	-1.17	-1.28	-1.35	-1.39	-1.44	-1.52	-1.58
3+7	-0.22	-0.48	-0.76	-0.96	-1.14	-1.31	-1.40	-1.52	-1.59	-1.66
4+7	-0.90	-1.05	-1.18	-1.32	-1.41	-1.50	-1.55	-1.60	-1.66	-1.70
5+7	-0.92	-1.06	-1.17	-1.27	-1.33	-1.40	-1.50	-1.56	-1.60	-1.66

Pattern	F0.10	F0.20	F0.30	F0.40	F0.50	F0.60	F0.70	F0.80	F0.90	F0.100
6+7	-0.53	-0.75	-0.92	-1.07	-1.19	-1.27	-1.36	-1.45	-1.53	-1.61
7+7	-0.82	-0.95	-1.07	-1.16	-1.24	-1.30	-1.37	-1.42	-1.49	-1.55
8+7	-0.93	-1.05	-1.16	-1.26	-1.35	-1.42	-1.45	-1.46	-1.50	-1.53
1+8	-0.49	-0.76	-0.96	-1.08	-1.20	-1.24	-1.25	-1.26	-1.29	-1.31
2+8	-0.50	-0.77	-0.95	-1.12	-1.25	-1.29	-1.32	-1.32	-1.33	-1.34
3+8	-0.21	-0.43	-0.79	-1.03	-1.18	-1.27	-1.34	-1.38	-1.41	-1.45
4+8	-0.72	-1.00	-1.16	-1.30	-1.43	-1.46	-1.49	-1.50	-1.53	-1.55
5+8	-0.92	-1.05	-1.14	-1.24	-1.30	-1.35	-1.36	-1.35	-1.35	-1.36
6+8	-0.47	-0.69	-0.89	-1.02	-1.14	-1.24	-1.26	-1.28	-1.28	-1.31
7+8	-0.75	-0.96	-1.11	-1.22	-1.27	-1.31	-1.31	-1.32	-1.34	-1.33
8+8	-0.69	-0.86	-0.98	-1.04	-1.08	-1.11	-1.10	-1.11	-1.11	-1.12

Table B14. Standard deviation values of the normalised F0 of Zhangzhou phrase-final tones across 10 sampling points across 64 disyllabic combinations

Pattern	F0.10	F0.20	F0.30	F0.40	F0.50	F0.60	F0.70	F0.80	F0.90	F0.100
1+1	0.41	0.38	0.37	0.39	0.42	0.45	0.48	0.51	0.53	0.56
2+1	0.33	0.32	0.33	0.35	0.38	0.41	0.43	0.44	0.47	0.53
3+1	0.58	0.40	0.36	0.37	0.39	0.43	0.45	0.45	0.45	0.46
4+1	0.39	0.37	0.38	0.39	0.41	0.43	0.45	0.47	0.47	0.49
5+1	0.36	0.34	0.36	0.40	0.45	0.51	0.56	0.60	0.61	0.63
6+1	0.35	0.28	0.30	0.33	0.36	0.39	0.41	0.41	0.43	0.46
7+1	0.40	0.37	0.38	0.41	0.46	0.51	0.57	0.64	0.64	0.68
8+1	0.30	0.30	0.32	0.34	0.38	0.42	0.47	0.51	0.55	0.60
1+2	0.53	0.46	0.41	0.41	0.42	0.41	0.41	0.44	0.47	0.49
2+2	0.53	0.44	0.40	0.40	0.42	0.41	0.43	0.46	0.49	0.50
3+2	0.71	0.60	0.52	0.49	0.50	0.44	0.48	0.49	0.51	0.53
4+2	0.53	0.48	0.45	0.59	0.46	0.44	0.45	0.48	0.49	0.51
5+2	0.44	0.35	0.34	0.36	0.39	0.41	0.43	0.47	0.50	0.52
6+2	0.51	0.44	0.41	0.46	0.45	0.42	0.46	0.50	0.53	0.54
7+2	0.52	0.42	0.39	0.40	0.40	0.42	0.45	0.47	0.50	0.51
8+2	0.45	0.42	0.42	0.44	0.48	0.54	0.59	0.61	0.59	0.62
1+3	0.57	0.52	0.51	0.51	0.50	0.56	0.58	0.56	0.57	0.65
2+3	0.49	0.45	0.45	0.45	0.48	0.51	0.62	0.59	0.61	0.68
3+3	0.52	0.49	0.49	0.49	0.51	0.52	0.56	0.57	0.58	0.64
4+3	0.67	0.66	0.66	0.65	0.65	0.64	0.63	0.61	0.60	0.62
5+3	0.59	0.56	0.55	0.53	0.53	0.55	0.57	0.55	0.60	0.80
6+3	0.64	0.59	0.56	0.53	0.50	0.50	0.54	0.53	0.55	0.60
7+3	0.61	0.54	0.50	0.50	0.52	0.56	0.62	0.62	0.66	0.66
8+3	0.78	0.77	0.74	0.69	0.66	0.65	0.65	0.62	0.60	0.57
1+4	0.45	0.39	0.36	0.37	0.37	0.43	0.47	0.47	0.50	0.55
2+4	0.51	0.45	0.44	0.44	0.43	0.44	0.44	0.45	0.56	0.55

Pattern	F0.10	F0.20	F0.30	F0.40	F0.50	F0.60	F0.70	F0.80	F0.90	F0.100
3+4	0.58	0.50	0.42	0.41	0.41	0.39	0.40	0.40	0.44	0.52
4+4	0.52	0.48	0.41	0.41	0.39	0.40	0.42	0.43	0.47	0.49
5+4	0.47	0.39	0.36	0.34	0.33	0.33	0.35	0.37	0.41	0.46
6+4	0.60	0.46	0.40	0.38	0.37	0.37	0.36	0.42	0.46	0.59
7+4	0.65	0.38	0.32	0.30	0.31	0.29	0.34	0.39	0.43	0.49
8+4	0.54	0.47	0.44	0.40	0.37	0.36	0.38	0.38	0.47	0.51
1+5	0.32	0.31	0.31	0.31	0.31	0.31	0.31	0.32	0.32	0.33
2+5	0.48	0.29	0.29	0.29	0.30	0.31	0.32	0.32	0.32	0.33
3+5	0.36	0.32	0.31	0.31	0.31	0.31	0.31	0.31	0.33	0.34
4+5	0.33	0.31	0.30	0.31	0.31	0.31	0.32	0.32	0.32	0.33
5+5	0.32	0.29	0.29	0.30	0.31	0.31	0.31	0.32	0.32	0.33
6+5	0.31	0.29	0.29	0.30	0.31	0.37	0.38	0.42	0.45	0.47
7+5	0.46	0.29	0.28	0.28	0.28	0.28	0.28	0.28	0.29	0.31
8+5	0.28	0.27	0.27	0.27	0.27	0.28	0.29	0.31	0.32	0.30
1+6	0.43	0.37	0.34	0.34	0.36	0.40	0.48	0.52	0.61	0.72
2+6	0.49	0.38	0.35	0.36	0.38	0.42	0.45	0.51	0.60	0.68
3+6	0.42	0.36	0.37	0.35	0.37	0.40	0.48	0.58	0.69	0.74
4+6	0.44	0.39	0.38	0.39	0.43	0.48	0.52	0.56	0.60	0.65
5+6	0.48	0.45	0.44	0.44	0.42	0.43	0.49	0.52	0.55	0.58
6+6	0.54	0.47	0.42	0.42	0.41	0.42	0.55	0.58	0.70	0.72
7+6	0.42	0.36	0.36	0.38	0.41	0.44	0.48	0.53	0.58	0.66
8+6	0.35	0.28	0.26	0.29	0.30	0.37	0.44	0.54	0.59	0.68
1+7	0.57	0.62	0.46	0.45	0.44	0.45	0.46	0.50	0.54	0.56
2+7	0.51	0.43	0.40	0.38	0.39	0.39	0.39	0.41	0.46	0.50
3+7	0.91	0.81	0.59	0.54	0.51	0.47	0.46	0.52	0.55	0.60
4+7	0.56	0.53	0.49	0.48	0.47	0.48	0.48	0.51	0.53	0.56
5+7	0.42	0.39	0.40	0.43	0.45	0.44	0.47	0.49	0.51	0.55
6+7	0.57	0.49	0.45	0.44	0.42	0.41	0.43	0.46	0.50	0.54
7+7	0.47	0.41	0.40	0.41	0.42	0.43	0.47	0.50	0.50	0.55
8+7	0.45	0.42	0.44	0.45	0.48	0.49	0.49	0.49	0.52	0.52
1+8	0.67	0.55	0.49	0.55	0.51	0.46	0.46	0.49	0.51	0.50
2+8	0.67	0.56	0.67	0.61	0.46	0.46	0.49	0.50	0.52	0.53
3+8	1.08	1.22	0.82	0.69	0.62	0.56	0.53	0.56	0.56	0.58
4+8	0.65	0.57	0.61	0.70	0.54	0.52	0.56	0.58	0.60	0.64
5+8	0.40	0.36	0.37	0.40	0.45	0.42	0.41	0.46	0.48	0.52
6+8	0.62	0.60	0.57	0.54	0.51	0.42	0.44	0.42	0.41	0.45
7+8	0.46	0.40	0.38	0.35	0.32	0.37	0.38	0.41	0.44	0.47
8+8	0.47	0.43	0.37	0.38	0.38	0.40	0.40	0.44	0.46	0.47

Table B15. Result of pairwise *t*-test camparison of the normalised F0 values of Zhangzhou phrase-final tones across 64 tonal combinations at 10% sampling point

	1+1	2+1	3+1	4+1	5+1	6+1	7+1
2+1	0.0023	-	-	-	-	-	-
3+1	0.3864	1.0000	-	-	-	-	-
4+1	< 2e-16	3.1e-08	3.8e-12	-	-	-	-
5+1	< 2e-16	6.8e-14	< 2e-16	0.3274	-	-	-
6+1	3.8e-12	0.0013	5.9e-06	1.0000	0.1637	-	-
7+1	1.6e-11	0.0054	2.7e-05	1.0000	0.0186	1.0000	-
8+1	4.3e-10	0.0018	2.3e-05	1.0000	1.0000	1.0000	1.0000

	1+2	2+2	3+2	4+2	5+2	6+2	7+2
2+2	0.0944	-	-	-	-	-	-
3+2	0.0030	3.9e-10	-	-	-	-	-
4+2	1.3e-08	0.0447	< 2e-16	-	-	-	-
5+2	6.4e-13	5.2e-05	< 2e-16	1.0000	-	-	-
6+2	0.0028	1.0000	3.1e-12	1.0000	0.0280	-	-
7+2	9.0e-08	0.0490	< 2e-16	1.0000	1.0000	1.0000	-
8+2	0.0017	0.4043	3.7e-08	1.0000	1.0000	1.0000	1.0000

	1+3	2+3	3+3	4+3	5+3	6+3	7+3
2+3	0.40425	-	-	-	-	-	-
3+3	1.00000	0.00056	-	-	-	-	-
4+3	0.00420	1.00000	3.0e-07	-	-	-	-
5+3	1.00000	1.00000	0.00421	1.00000	-	-	-
6+3	0.16368	2.3e-05	1.00000	1.4e-08	0.00018	-	-
7+3	0.28123	3.9e-05	1.00000	1.9e-08	0.00032	1.00000	-
8+3	0.04538	1.00000	0.00047	1.00000	1.00000	3.4e-05	5.8e-05

	1+4	2+4	3+4	4+4	5+4	6+4	7+4
2+4	1.0000	-	-	-	-	-	-
3+4	< 2e-16	< 2e-16	-	-	-	-	-
4+4	6.8e-06	3.4e-05	< 2e-16	-	-	-	-
5+4	1.2e-13	9.9e-13	< 2e-16	0.0228	-	-	-
6+4	3.5e-08	4.9e-09	0.4013	< 2e-16	< 2e-16	-	-
7+4	0.0702	0.1632	< 2e-16	1.0000	0.0026	8.3e-15	-
8+4	3.2e-05	9.7e-05	< 2e-16	1.0000	1.0000	< 2e-16	1.0000

	1+5	2+5	3+5	4+5	5+5	6+5	7+5
2+5	1.00000	-	-	-	-	-	-
3+5	0.00340	0.00970	-	-	-	-	-
4+5	1.3e-09	2.0e-10	< 2e-16	-	-	-	-
5+5	1.0e-15	< 2e-16	< 2e-16	0.41992	-	-	-
6+5	1.00000	1.00000	0.68989	4.8e-11	< 2e-16	-	-
7+5	0.22897	0.10963	1.8e-08	0.02905	5.3e-06	0.01211	-
8+5	6.3e-15	1.7e-15	< 2e-16	0.00025	0.10379	< 2e-16	1.2e-08

	1+6	2+6	3+6	4+6	5+6	6+6	7+6
2+6	0.33869	-	-	-	-	-	-
3+6	2.7e-10	0.00039	-	-	-	-	-
4+6	0.05831	3.7e-07	< 2e-16	-	-	-	-
5+6	3.3e-07	6.3e-15	< 2e-16	0.12051	-	-	-
6+6	1.00000	1.00000	1.2e-06	0.00378	1.2e-08	-	-
7+6	1.00000	0.00037	2.1e-15	1.00000	0.02945	0.15738	-
8+6	1.00000	0.05649	3.6e-07	1.00000	1.00000	0.96031	1.00000

	1+7	2+7	3+7	4+7	5+7	6+7	7+7
2+7	1.00000	-	-	-	-	-	-
3+7	2.9e-09	3.1e-09	-	-	-	-	-
4+7	8.1e-05	0.00039	< 2e-16	-	-	-	-
5+7	2.2e-05	0.00011	< 2e-16	1.00000	-	-	-
6+7	1.00000	1.00000	2.6e-05	9.3e-08	2.3e-08	-	-
7+7	0.04702	0.10920	< 2e-16	1.00000	1.00000	0.00040	-
8+7	0.01043	0.02068	3.0e-14	1.00000	1.00000	0.00021	1.00000

	1+8	2+8	3+8	4+8	5+8	6+8	7+8
2+8	1.00000	-	-	-	-	-	-
3+8	0.40326	0.22224	-	-	-	-	-
4+8	0.75930	0.51230	2.6e-05	-	-	-	-
5+8	0.00654	0.00279	3.9e-08	1.00000	-	-	-
6+8	1.00000	1.00000	1.00000	0.96668	0.01501	-	-
7+8	1.00000	1.00000	0.00081	1.00000	1.00000	1.00000	-
8+8	1.00000	0.87365	4.6e-05	1.00000	0.69784	1.00000	1.00000

Table B16. Result of pairwise *t*-test camparison of the normalised F0 values of Zhangzhou phrase-final tones across 64 tonal combinations at 100% sampling point

	1+1	2+1	3+1	4+1	5+1	6+1	7+1
2+1	0.00085	-	-	-	-	-	-
3+1	< 2e-16	2.8e-14	-	-	-	-	-
4+1	< 2e-16	5.8e-14	1.00000	-	-	-	-
5+1	1.00000	0.00481	< 2e-16	< 2e-16	-	-	-
6+1	< 2e-16	6.9e-12	1.00000	1.00000	< 2e-16	-	-
7+1	1.00000	0.00752	< 2e-16	< 2e-16	1.00000	< 2e-16	-
8+1	1.00000	0.27211	3.7e-15	2.4e-14	1.00000	7.0e-14	1.00000

	1+2	2+2	3+2	4+2	5+2	6+2	7+2
2+2	1.000	-	-	-	-	-	-
3+2	1.000	1.000	-	-	-	-	-
4+2	0.042	1.000	1.000	-	-	-	-
5+2	1.000	1.000	1.000	1.000	-	-	-
6+2	0.398	1.000	1.000	1.000	1.000	-	-
7+2	1.000	1.000	1.000	0.075	1.000	0.532	-
8+2	1.000	0.937	0.668	0.097	0.508	0.243	1.000

	1+3	2+3	3+3	4+3	5+3	6+3	7+3
2+3	1.0	-	-	-	-	-	-
3+3	1.0	1.0	-	-	-	-	-
4+3	1.0	1.0	1.0	-	-	-	-
5+3	1.0	1.0	1.0	1.0	-	-	-
6+3	1.0	1.0	1.0	1.0	1.0	-	-
7+3	1.0	1.0	1.0	1.0	1.0	1.0	-
8+3	1.0	0.5	1.0	1.0	1.0	1.0	1.0

	1+4	2+4	3+4	4+4	5+4	6+4	7+4
2+4	1	-	-	-	-	-	-
3+4	1	1	-	-	-	-	-
4+4	1	1	1	-	-	-	-
5+4	1	1	1	1	-	-	-
6+4	1	1	1	1	1	-	-
7+4	1	1	1	1	1	1	-
8+4	1	1	1	1	1	1	1

	1+5	2+5	3+5	4+5	5+5	6+5	7+5
2+5	1.00000	-	-	-	-	-	-
3+5	3.6e-07	2.4e-06	-	-	-	-	-
4+5	5.9e-16	1.1e-14	0.35682	-	-	-	-
5+5	4.3e-06	2.4e-05	1.00000	0.08632	-	-	-
6+5	4.7e-11	3.7e-10	1.00000	1.00000	0.46808	-	-
7+5	0.06223	0.16531	0.73208	0.00011	1.00000	0.00281	-
8+5	1.4e-05	4.2e-05	1.00000	1.00000	1.00000	1.00000	0.12254

	1+6	2+6	3+6	4+6	5+6	6+6	7+6
2+6	1.00	-	-	-	-	-	-
3+6	1.00	1.00	-	-	-	-	-
4+6	1.00	1.00	1.00	-	-	-	-
5+6	1.00	1.00	1.00	1.00	-	-	-
6+6	1.00	0.71	1.00	1.00	1.00	-	-
7+6	1.00	1.00	1.00	1.00	1.00	1.00	-
8+6	1.00	1.00	1.00	1.00	1.00	1.00	1.00

	1+7	2+7	3+7	4+7	5+7	6+7	7+7
2+7	1.00	-	-	-	-	-	-
3+7	1.00	1.00	-	-	-	-	-
4+7	1.00	0.93	1.00	-	-	-	-
5+7	1.00	1.00	1.00	1.00	-	-	-
6+7	1.00	1.00	1.00	1.00	1.00	-	-
7+7	1.00	1.00	1.00	0.27	1.00	1.00	-
8+7	1.00	1.00	1.00	0.83	1.00	1.00	1.00

	1+8	2+8	3+8	4+8	5+8	6+8	7+8
2+8	1.0000	-	-	-	-	-	-
3+8	1.0000	1.0000	-	-	-	-	-
4+8	0.1237	0.2310	1.0000	-	-	-	-
5+8	1.0000	1.0000	1.0000	0.8091	-	-	-
6+8	1.0000	1.0000	1.0000	0.3765	1.0000	-	-
7+8	1.0000	1.0000	1.0000	0.6122	1.0000	1.0000	-
8+8	0.5700	0.0730	0.0019	4.1e-08	0.0754	1.0000	0.7583

Table B17. Normalised duration values of Zhangzhou phrase-final tones across phrase-initial tones

Tone	V1	V2	V3	V4	V5	V6	V7	V8
1	1.38	1.17	1.11	0.97	1.42	0.82	0.97	1.27
2	1.39	1.21	1.13	0.99	1.40	0.82	1.01	1.23
3	1.40	1.11	1.11	0.96	1.40	0.81	0.97	1.13
4	1.36	1.07	1.09	0.88	1.40	0.81	0.92	1.18
5	1.39	1.20	1.14	0.99	1.47	0.87	1.04	1.32
6	1.48	1.15	1.15	0.94	1.47	0.88	0.99	1.24
7	1.38	1.26	1.15	0.99	1.47	0.89	1.15	1.29
8	1.33	1.18	1.12	0.94	1.59	0.87	1.01	1.33

Table B18. Standard deviation values of the normalised duration of Zhangzhou phrase-final tones across phrase-initial tones

Tone	V1	V2	V3	V4	V5	V6	V7	V8
1	20.28	32.75	17.09	18.46	20.56	18.76	24.52	28.86
2	18.31	29.80	17.68	18.97	21.85	18.60	19.59	30.17
3	20.27	29.40	17.95	20.19	20.16	18.13	24.58	35.58
4	22.90	33.16	17.92	20.97	22.87	18.25	21.83	28.99
5	19.83	30.82	18.02	20.05	25.70	19.61	22.58	32.76
6	21.59	34.34	18.08	18.91	24.81	18.36	24.27	31.16
7	19.25	33.81	18.75	19.84	22.33	18.36	21.94	26.74
8	18.34	32.17	16.47	17.50	21.66	22.71	24.26	27.39

Table B19. Result of pairwise *t*-test camparison of the normalised duration values of Zhangzhou phrase-final tones across 64 tonal combinations

	1+1	2+1	3+1	4+1	5+1	6+1	7+1
2+1	1.00000	-	-	-	-	-	-
3+1	1.00000	1.00000	-	-	-	-	-
4+1	1.00000	1.00000	0.59575	-	-	-	-
5+1	1.00000	1.00000	1.00000	1.00000	-	-	-
6+1	2.8e-05	0.00097	0.00441	4.6e-08	0.00083	-	-
7+1	1.00000	1.00000	1.00000	1.00000	1.00000	0.00088	-
8+1	1.00000	0.35943	0.19311	1.00000	0.39082	1.7e-06	1.00000

	1+2	2+2	3+2	4+2	5+2	6+2	7+2
2+2	1.0000	-	-	-	-	-	-
3+2	1.0000	0.0223	-	-	-	-	-
4+2	0.0037	3.2e-06	1.0000	-	-	-	-
5+2	1.0000	1.0000	0.0714	2.4e-05	-	-	-
6+2	1.0000	1.0000	1.0000	0.1446	1.0000	-	-
7+2	0.1071	1.0000	6.5e-05	9.7e-10	1.0000	0.0223	-
8+2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

	1+3	2+3	3+3	4+3	5+3	6+3	7+3
2+3	1.0000	-	-	-	-	-	-
3+3	1.0000	1.0000	-	-	-	-	-
4+3	1.0000	0.3400	1.0000	-	-	-	-
5+3	1.0000	1.0000	1.0000	0.0271	-	-	-
6+3	1.0000	1.0000	1.0000	0.0429	1.0000	-	-
7+3	0.5360	1.0000	0.4196	0.0051	1.0000	1.0000	-
8+3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

	1+4	2+4	3+4	4+4	5+4	6+4	7+4
2+4	1.00000	-	-	-	-	-	-
3+4	1.00000	1.00000	-	-	-	-	-
4+4	4.3e-08	1.7e-11	9.6e-07	-	-	-	-
5+4	1.00000	1.00000	1.00000	1.2e-11	-	-	-
6+4	1.00000	1.00000	1.00000	0.00047	0.99574	-	-
7+4	1.00000	1.00000	1.00000	7.0e-10	1.00000	0.72625	-
8+4	1.00000	0.60582	1.00000	0.09217	0.45702	1.00000	0.33447

	1+5	2+5	3+5	4+5	5+5	6+5	7+5
2+5	1.0000	-	-	-	-	-	-
3+5	1.0000	1.0000	-	-	-	-	-
4+5	1.0000	1.0000	1.0000	-	-	-	-
5+5	0.8845	0.0167	0.0109	0.0175	-	-	-
6+5	0.8091	0.0240	0.0166	0.0275	1.0000	-	-
7+5	0.6203	0.0143	0.0096	0.0157	1.0000	1.0000	-
8+5	4.7e-06	4.5e-08	2.9e-08	4.4e-08	0.0029	0.0118	0.0103

	1+6	2+6	3+6	4+6	5+6	6+6	7+6
2+6	1.00000	-	-	-	-	-	-
3+6	1.00000	1.00000	-	-	-	-	-
4+6	1.00000	1.00000	1.00000	-	-	-	-
5+6	0.24175	0.11899	0.02752	0.01492	-	-	-
6+6	0.01126	0.00495	0.00094	0.00044	1.00000	-	-
7+6	0.01730	0.00778	0.00155	0.00075	1.00000	1.00000	-
8+6	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000

	1+7	2+7	3+7	4+7	5+7	6+7	7+7
2+7	1.00000	-	-	-	-	-	-
3+7	1.00000	1.00000	-	-	-	-	-
4+7	0.41648	0.00180	0.32454	-	-	-	-
5+7	0.20448	1.00000	0.27503	7.0e-06	-	-	-
6+7	1.00000	1.00000	1.00000	0.04587	1.00000	-	-
7+7	2.2e-10	6.0e-06	4.4e-10	< 2e-16	0.00033	7.8e-08	-
8+7	1.00000	1.00000	1.00000	0.22312	1.00000	1.00000	0.00145

	1+8	2+8	3+8	4+8	5+8	6+8	7+8
2+8	1.00000	-	-	-	-	-	-
3+8	0.20143	1.00000	-	-	-	-	-
4+8	1.00000	1.00000	1.00000	-	-	-	-
5+8	1.00000	1.00000	0.01266	0.15010	-	-	-
6+8	1.00000	1.00000	1.00000	1.00000	1.00000	-	-
7+8	1.00000	1.00000	0.21694	1.00000	1.00000	1.00000	-
8+8	1.00000	0.54630	0.00038	0.00570	1.00000	1.00000	1.00000

Appendix C: Sound Files and Raw Data (attached USB)