A Sociophonetic Study on Tonal Variation

of

the Wúxī and Shànghǎi Dialects

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111

e-mail: lot@uu.nl
http://www.lotschool.nl

Cover illustration: a typical river scene in the Wú area. The photo is taken by Pengfei Zhang and edited by Jingwei Zhang.

ISBN: 978-94-6093-152-9
NUR 616

A Sociophonetic Study on Tonal Variation of the Wúxī and Shànghǎi Dialects

Een sociofonetisch onderzoek naar toonvariatie in de dialecten van Wúxī en Shànghǎi
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op vrijdag 17 oktober 2014 des ochtends te 12.45 uur

door

Jingwei Zhang

geboren op 14 mei 1985 te Wúxī, China

Promotor: Prof. dr. R.W.J. Kager
Co-promotor: Dr. H. Van de Velde

This dissertation was partly accomplished with financial support from the China Scholarship Council (CSC).

To my parents

致我的父母

Contents

Acknowledgements

Great, finally I get the award for my past five years PhD work, my dissertation. Doing a PhD is indeed a challenging process. I could not bring this dissertation into being without the help and support of many people. I would like to sincerely express my appreciations to all of them here.

First of all, I would like to express my heartfelt gratitude to my supervisors, René Kager and Hans Van de Velde, for their excellent supervision. René and Hans, you shared your knowledge, guidance, support, and enthusiasm along all the paths of research I have followed since I started this study. Thank you both for that. I have learned a tremendous amount from you and feel privileged to have had that opportunity. René, you are always available for questions and discussions. I am so grateful of your conscientious work. You worked on my writings a lot of the time, far into the night. Hans, I want to thank you in particular for what you have offered to me, not only in research, but also for the help when I moved to the new apartment, as well as for the knowledge and experience you shared with me and suggestions for my future career. It was very enjoyable when staying with you, Laurence and Guillaume in Erps-Kwerps. Laurence, thank you for inviting me to your Christmas party with your family, taking me to a Vadim Repin's concert and teaching me how to make a sandwich!

I am grateful to my reading committee for their careful reading and insightful comments on my manuscript: Hugo Quené, Wim Zonneveld, Vincent van Heuven, Richard V. Simmons and Daming Xu. Xu laoshi, you are especially thanked for introducing me into the linguistics and for recommending me to the UiL-OTS. You are always an experienced, far-sighted and trusted adviser to me.

I acknowledge China Scholarship Council for providing the funds that made it possible for me to study in Utrecht. The UiL OTS also offered me the funds for data collection, workshops and conferences, which I highly appreciate.

There are many people who helped me in various stages of my research. When designing my research proposal and tools for data collection, I received a lot of inspiration and advices from Yiya Chen, Rujie You, Xiaonong Zhu and Huan Tao, researchers who worked on the Wu dialects for many years. Thank you all for generously sharing your knowledge; I could not have accomplished this without your help in the early stages. I would like to thank Yueling Ping, Hongbo Lei, Jieming Wang, Xinrong Zhang and his family, Yiting Lin, Yi Tang, Yaojie Hu and my parents, for their assistance and contributions when I collected data in Wuxi and Shanghai. I am grateful for all the participants who attended the interviews in the hot summer of 2010. I would like to thank Hans Van de Velde, Hugo Quené, Mattis van den Bergh, Weiwei Zhang, James Walker, Anne-France Pinget, Sander van der Harst, Ao Chen and Qian Li for your kind support in statistics. Chuanyi Li, Chenjuan Jiang and Ao Chen helped me a lot in computer programming. My heartfelt thanks go to Huan Chen, who helped me with dull and repetitive work in segmentation and transcription. Thanks go to Rachel Gargiulo, Richard Thrift and Heidi Klockmann for the English proofreading and editing. I also want to thank Pengfei Zhang for offering me the beautiful photo I used on the cover picture of this book. Anqi, thanks a lot for your brilliant idea about the cover design and for all the coordination work.

The UiL OTS, where I have worked with pleasure every day, has been the domain of great people for all the years that I have been there, each with different characteristics and special qualities. Maaike Schoorlemmer has taken care of all the administration concerned with my work, in the most efficient and pleasant way imaginable. I especially thank you for offering me an opportunity to apply for the three months funding after my CSC contract; that was really helpful. I want to thank Rias van den Doel for your efforts to improve my English pronunciation. I also want to thank my former and current office-mates: Kiki Kushatanti, Rianne Schippers, Anna Volkova, Mulusew Asratie, Jolien Scholten, Björn't Hart, Anja Goldschmidt and Marta Castella, for all of you together make the office life full of joys and smiles in the warm attic in Trans 8. I also spent a wonderful time with the big Chinese group in the UiL OTS: Ao Chen, Liquan Liu, Anqi Yang, Xiaoli Dong, Fang Li, Zenghui Liu, Yipu Wei, Mengru Han, Luying Hou, Xin Li, Shuangshuang Hu, Qiong Li, Jun Zhang, Shanshan Fan, Yueru Ni, Yuning Sun and Haifan Lan. Thank you for your support and discussion on Chinese linguistics. I also will not forget the traditional Chinese festivals we spent together. Inside the UiL-OTS, I met colleagues and friends shared their thoughts and fun with me. We had nice chats and we had meetings, courses, coffee breaks, lunches, dinners, and parties together. I thank Sander van der Harst, Anne-France Pinget, Bert Le Bruyn, Maartje Schulpen, Stavroula Alexandropoulou, Alexis Dimitriadis, Ana Auilar Guevara, Marko

Simonovic, Assaf Toledo, Mike Olson, Heidi Klockmann, Anna Sara Romøren, Brigitta Keij, Nadya Goldberg, Ileana Grama, Anne van Leeuwen, Liv Persson and Anna Chernilovskaya.

Regarding my life in Utrecht outside the UiL-OTS, I also want to thank the Chinese friends I met in the Netherlands for their friendship and company: Qiulan Zhang, Yinhuan Kuang, Fangfang Cheng, Pu Hao, Qingyun Qian, Jinbao Gao, Xiaoqian Li, Yafei Liu, Can Cui, Huang Xu, Feng Jianxi, Li Shuai, Zhang Xinlin & Liu Xiaoman & Hehe, Miao Zhang, Meng Wang & Junting Zhang, Dandan Wang, Shaoyu Yin, Cong Zhang, Hao Zhang and Fang Chen. I also want to thank Qian Li, Yan Gu, Ting Zou, Junru Wu, who are the Chinese linguists not working at the UiL-OTS, for the helpful discussions and joyful times we spent together at the LOT winter and summer schools.

Hong Hu, my alumna at Nanjing University and my roommate in Utrecht. I had an enjoyable life with you in our cozy home in Zeist. Definitely, the life here would have been totally different without you. Your “ohayou” made the mornings full of energy. You are always trustworthy to share my small secrets and give me many valuable suggestions, just like my elder sister. I also thank your husband Xiang Zhang for bringing nice food from China and doing a lot of housework at our place when he visited you.

I would like to thank people I worked together with at Nanjing University for their contributions in a more or lesser extent to my development: Fang Liu, Yanhong Ge, Jie Xi, Xiaobing Fang, Juanjuan Xu, Weiqi Yu, Jing Zhang, Weihong Wang, Haiying Li and Yizheng Jiang.

I own many thanks to Jinfeng Liu for his full understanding, encouragement, support and love. Specially thanks for his great efforts to the editorial work on my dissertation.

Last but not least, I would like to thank my parents. They were always my strongest support.

爸爸妈妈，感谢你们的养育之恩。感谢这五年来，你们对我的理解与支持。我永远爱你们，我将这本论文献给你们。

Conventions

In this dissertation, Hànyǔ Pīnyīn (the alphabetical writing system for the Chinese language, 汉语拼音) is used to transcribe proper nouns (names of places, individual people, languages, nations, ethnic groups and other specific referents), titles of documents published in Chinese, and special academic terms in Chinese Dialectology and Chinese Phonology. An English translation and Guīfàn Hànzì (the Standard Chinese Characters, 规范汉字)[1] of Hànyǔ Pīnyīn may also be provided in parentheses as an annotation when a term is used for the first time. In addition, Guīfàn Hànzì's are used for the purpose of differentiating homophones in Mandarin Chinese.

Hànyǔ Pīnyīn is the Romanized system for the transcription of Chinese characters. It was adopted as the international standard for transcribing Chinese by the International Organization for Standardization[2] in 1982. The orthographic rules prescribed by the *Hànyǔ Pīnyīn Zhèngcífǎ Jīběn Guīzé* (*Basic rules for Hànyǔ Pīnyīn orthography*, 《汉语拼音正词法基本规则》, 2012)[3] are adopted in the spelling of Hànyǔ Pīnyīn in this dissertation. It is noteworthy that tone marks are added to comply with the *Hànyǔ Pīnyīn Zhèngcífǎ Jīběn Guīzé* as well as increasing attention for the tones.

>	becomes, turns into
BSD	the Bǎoshān dialect
BSL	Bǎoshān Luódiàn county
BSS	Bǎoshān Shuāngcǎodūn county
C	consonantal onset
DC	Dispersion Coefficient

---

[1] Guīfàn Hànzì is a concept defined by *Zhōnghuá Rénmín Gònghé Guó Tōngyòng Yǔyán Wénzì Fǎ (The Law of the People's Republic of China on the National Common Language and Characters*, 《中华人民共和国通用语言文字法》) (2000).

[2] "ISO 7098:1982 – Documentation – Romanization of Chinese", www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=13682, retrieved on May 20, 2014.

[3] Full text (in Chinese) can be found here: http://www.moe.gov.cn/ewebeditor/uploadfile/2012/08/21/20120821100233165.pdf, retrieved on May 20, 2014.

F_0	fundamental frequency
FOR	Fraction of Range, a tonal normalization procedure
G	prenuclear onglide
GSDJPS	*General situation of dialects in Jiāngsū Province and Shànghǎi* (1960)
IPA	International Phonetic Alphabet
LD	ratio of log semitone distances, a tonal normalization procedure
LDA	linear discriminant analysis
LPOR	logarithmic proportion of range, a tonal normalization procedure
LVC	language variation and change
LZ	logarithmic Z-score, a tonal normalization procedure
M	mean
MC	Middle Chinese
NHD	the Nánhuì dialect
NHZ	Nánhuì Zhōupǔ county
NI	Normalization Index
NORMs	non-mobile older rural males
NT_{Px}	the normalized time of P_x
P_{max}	F_0 minimum of a tonal contour
P_{min}	F_0 maximum of a tonal contour
POR	proportion of range, a tonal normalization procedure
PTH	Pǔtōnghuà, the common language of China and a Mandarin variety
Px	the X point on the pitch contour
RMSD	Root Mean Square Deviation
SD	standard deviation
SHD	the Shànghǎi dialects, including the urban and suburban dialects
SHUD	the Shànghǎi urban dialect
SJC	Sōngjiāng County
SJD	the Sōngjiāng dialect
ST-AvgF_0	a semitone transformation relative to each speaker's average pitch in Hz
ST_{Px}	the semitone value of P_x on the pitch contour
TBU	tone bearing unit
T_{Px}	the real time of P_x
ULS	Urban Language Survey
V	nucleu
WXD	the Wúxī dialects, including the urban and suburban dialects
X	coda
X_{vertex}	the x-coordinate of the vertex on a parabola $y=ax^2+bx+c$
YRDA	Yangtze River Delta Area
σ_1	the first syllable
σ_2	the second syllable

List of Tables

List of Figures

Chapter 1. Introduction

This introductory chapter will set the stage for the investigation of tonal variation in the Wú dialects. The Wú area is undergoing a rapid urbanization process. In order to capture language change during this process, the wider topic of this study: language change in the context of urbanization will be first reviewed in Section 1.1. In Section 1.2, the general goals and scope of this study, a sociophonetic study into tonal variation in the Wú dialects, will be introduced. Finally, an outline of the study will be presented (Section 1.3).

1.1. Topic: urbanization and language change

Urbanization is one of the most radical worldwide social changes in the 20th and 21st centuries. Less than 5% of the world's population lived in cities a century ago. In 2008, for the first time in human history, this figure exceeded 50%. In the developed countries, the level of urbanization reached the 50% mark over half a century ago and is currently around 75%; but the growth of urban populations in these regions has been declining for decades. For instance, Europe has seen declining urban growth since 1950 and North America since 1970. Nowadays, urbanization is advancing mainly in the developing regions. Between 2007 and 2025, the annual rate of change in the urban population in developing countries is expected to be 2.3%, in contrast to the 0.5% in developed countries (Wang Guangtao, 2012).

Urbanization has been changing the modern social structure profoundly and has thus been of demonstrable consequence for language and language use. The impact of urbanization on language variation and change has received considerable attention from linguists over the past few decades. Sociolinguistic studies in the Labovian tradition originally studied linguistic variation in "socially diversified but geographically stable and mostly native populations, within economically and culturally well-established communities" (Nordberg, 1994), partly because these studies focused on the urban communities of the United States and Europe, which have completed the phase of accelerated development of urbanization. However, this does not mean that there were no variationist studies related to urbanization in the United States and Europe. Nordberg (1994) collected papers concerning linguistic consequences of migration and urbanization in four Nordic countries: Denmark, Finland, Norway and Sweden. In the American South, several variationist studies have also been conducted (Bailey, et al., 1996; Durian, 2007; Thomas, 1997; Tillery & Bailey, 2003). Communities in the South attracted attention mainly because they lagged behind the rest of the United Stated in the progress of urbanization by about 50 years (Tillery, et al., 2004). Therefore, in order to observe the urbanization effects on language in real time, it is more effective to conduct the variationist studies in the regions where the urbanization process is currently accelerating, i.e., the developing regions.

Cadora (1970) launched one of the first urban sociolinguistic studies in the developing regions. He described how a gradual process of ruralization and the later urbanization in Ramallah, an Arabic-speaking community of the Middle East, affected the phonological, morphological, and lexical changes there. Then followed systematic research on the urbanization of rural dialect speakers in Brazil (Bortoni-Ricardo, 1985), the absence of the Bahamian copula (Reaser, 2004), relations between language, migration, and urbanization in Palestine (Amara, 2005), the use of urban vernaculars in Zimbabwe (Makoni, et al., 2007) and language variation in a number of Arabic speech communities (Mille, et al., 2007). Similarly, Mille et al. (2007) summarized the importance of such studies conducted in the developing regions. On the one hand, the results gained in the developing regions can reveal to what extent general findings for Western cities apply to other parts of the world (universal trends); on the other hand, these results are also useful to solve the specific local linguistic issues induced in the process of industrialization, urbanization, and migration.

China, as the world's most populous and largest developing nation, has been accelerating its progress of industrialization and urbanization since the launch of the economic reform program in 1978. In 2011, the urban population of Mainland China increased from 17.9% of its total population in 1978 to 51.3%, which marked an urban majority for the first time in the nation's history. The spatial distribution of the urban population in China was concentrated and regionally imbalanced, mainly distributed over eastern China. According to research by the China National Development and Reform Commission（国家发展与改革委员会）, 10 major metropolitan regions were forming in China[4]. With the exception of regions

[4] http://politics.people.com.cn/GB/1026/5535974.html, retrieved on May 20, 2014.

across national boundaries, all the others are located in areas where Chinese dialects are spoken[5], rather than minority languages. In other words, China's urbanization process impacts, to a great extent, the dialect-speaking areas. In these areas, new kinds of social-geographic spaces have been created by numerous migrants. Local dialects are not only purely regional varieties but are changing their social functions and coming to reflect social stratification.

Along with the rapid urbanization process, Chinese linguists have documented subsequent dialect changes. Assuming a broad interpretation of sociolinguistics, a majority of them worked on the topics of language choice, language attitude and language identity (Xu, 2006). The general conclusion of such work is that, based on the decreasing number of dialect speakers and the non-"authentic" accent of the younger generation, local dialects are in competition with Pǔtōnghuà (PTH) but are gradually replaced by PTH. Some even concluded that Chinese dialects are endangered[6].

Trudgill (1999) pointed out that in order to study how changes diffuse over wider geographical areas, the researcher first needs to define what the exact "change" is. What is actually taking place in contemporary Chinese dialects? Are these dialects actually endangered? Is the "change" a complete replacement of local dialects by PTH or a contact-induced language change? If it is a contact-induced change, how is this change reflected in the segmental and suprasegmental aspects of pronunciation? To answer these questions, variationist sociolinguistic studies are needed.

Chinese dialectologists first examined these newly raised issues in the early 1980's (e.g., Bào, 1980) because there were very few Chinese linguists specializing in variationist sociolinguistics at that time. However, the dialectologists soon found that the traditional methodology and theories of dialectology were insufficient to

---

[5] The ten major metropolitan regions are: **Pearl River Delta** (mainly Cantonese), **Yangtze River Delta** (mainly the northern Wu dialects), **Bóhǎi Economic Rim** (mainly the Mandarin dialects, across boundaries with South Korea), **Cross-Strait Economic Zone** (the Southern Wu dialects, Hakka and the Mandarin dialects), **Beibu Gulf Economic Rim** (mainly Hakka, across boundary with Vietnam), **Central Plain** (mainly the Jin dialects), **Northeastern Cities** (mainly the Mandarin dialects, across boundaries with North Korea and Russia), **Chéngyú Megalopolis** (mainly the Mandarin dialects), **Greater Wǔhàn Megalopolis** (mainly the Mandarin dialects and the Gàn dialects), **Guānzhōng** (mainly the Mandarin dialects), **Greater Chángshā Metropolitan Region** (mainly the Gàn dialects and the Xiang dialects).

[6] The 1st and 2nd Seminar on Endangered Dialects were held at Zhongshan University in 2009 and 2011. More than fifty scholars participated in each seminar, including almost all the top scholars working on Chinese dialectology. Refer to: chinese.sysu.edu.cn/2012/Item/1870.aspx (in Chinese) and www.cich.org.cn/yjzx/zxxw/2013116/n06458450.html (in Chinese); both retrieved on May 20, 2014.

answer these questions (Zhou, 2010). The notion of sociolinguistics was formally introduced to Mainland China by Chén Yuán (陈原, Chén 1980; 1983) in the 1980s. Along with the development of sociolinguistic research in China, some innovative research approaches have emerged in order to deal with the new linguistic realities, among them, the "Urban Language Survey" (ULS), originated in China in the context of large-scale urbanization. The subject matter of the ULS is "any language-related matter in connection with the urban contexts or with the process of urbanization" (D. Xu 2006), so any methods or theories that can apply are included in the scope of the ULS. Therefore, the major feature of the ULS is its coordinateness. This feature is not only reflected in its diversified origins - urban dialectology, the sociology of language, and the ethnography of communication - but also in its research methods. The ULS integrates methods from dialectology, quantitative sociolinguistics, ethnomethodology and sociology (D. Xu, 2006). In the framework of ULS, a number of urban studies from a variationist sociolinguistic perspective sprang up in China (e.g., Fù, 2011), most of which were published in the *Journal of Chinese Sociolinguistics* and *Industrialization and the Re-structuring of Speech Communities in China and Europe* (van den Berg & Xu, 2010). These studies range over the areas of phonological variation (Yáng, 2005; Chéng, 2006; Qian, 2010; D. Xu, 2010), lexical variation (Gě, 2005; Liú, 2005) and syntactic variation (Qián, 2006). However, there are few studies on tone. Actually, in China and elsewhere, tonal variation is a less explored area in the whole field of variationist research compared to vowel and consonant variation (Stanford, 2008; Thomas, 2011). However, tone is a significant aspect of Chinese languages and dialects because all the varieties of the Chinese language family are tone languages (Yip, 2002: 171). Moreover, China has a long history of keeping records of tones, which can at least be dated back to *Qièyùn* (Spelling Rimes,《切韵》), the most authoritative Chinese rime dictionary published in 601 AD. The studies of Chinese tones were followed up in different linguistic fields, like historical Chinese phonology, Chinese dialectology and recently acoustic phonetics. The complexity of tonal variation makes it a suitable and interesting multidisciplinary area of study.

1.2. Goals and scope

This study is a sociophonetic investigation into the tonal variation in the dialects spoken in Wúxī and Shànghǎi. This study has two general goals. The first is to present a description of tonal variation in Wú Chinese, the urban and suburban dialects spoken in Wúxī and Shànghǎi, including citation tones on monosyllables and sandhi tones on bisyllables. The second is to elicit internal, external and extra-linguistic constraints as well as their interplay within the variationist paradigm. Regarding the three-way taxonomy of internal, external and extra-linguistic factors, this study follows Jones & Esch (2002). An "internal" factor in this study is defined as an internal linguistic factor (e.g., vowel height, tonal duration), an "external" factor is defined as an external linguistic factor (e.g., language contact factors), and an "extra-linguistic" factor refers to sociopolitical and economic factors.

In addition to the general goals, two methodological issues concerning the acoustic analysis of tone are investigated in this study: (1) the definition of the domain of tone in the Wú dialects**,** which is particularly relevant for this study of tonal variation and change; and (2) the tonal normalization method which can best serve this type of variationist study. In the following, two general goals and two methodological issues will be briefly introduced. Since the methodological issues serve as a prerequisite for the descriptive goal, Goal I, they are introduced immediately following Goal I.

Goal I: To present a description of tonal variation in Wú Chinese (吴语): the urban and suburban dialects spoken in Wúxī and Shànghǎi, including citation tones on monosyllables and sandhi tones on bisyllables.

First the reasons for pursuing Goal I will be explained and then two relevant methodological issues regarding attaining this goal will be introduced. Goal I mentions several key terms: (1) tonal variation; (2) Wú Chinese; (3) urban and suburban dialects; (4) Wúxī and Shànghǎi; and (5) citation tones and sandhi tones. What follows is an explanation of each key word embedded in Goal I:

(1) Tonal variation

As discussed in Section 1.1, the study of tones is still in its infancy in variationist research compared to the study of vowels and consonants. Tone is also a suitable object to combine analytic methods from historical Chinese phonology, Chinese dialectology and acoustic phonetics. Hence this study aims at studying sound change in progress by observing lexical tones. Generational differences are used to reveal diachronic changes that happened in the (recent) past of a language.

(2) Wú Chinese

Linguistic change is often associated with social change in the view of sociolinguistics. Nowadays, China is undergoing rapid growth and urbanization, especially in the dialect speaking areas. The resultant urbanization stimulates the linguistic changes ongoing in Chinese dialects. Wú Chinese is, in terms of numbers of speakers, the second largest Chinese dialect group. It is not mutually intelligible with PTH and shows close resemblances to the Middle Chinese tonal system, preserving seven or eight tonal categories. Meanwhile, there are relatively rich research documents on Wú Chinese. For example, Chao (1928) transcribed tonal systems in Wú using a five-point scale with the assistance of a sliding pitch pipe. This is the earliest study of Chinese dialectology using a precise notation system for lexical tone.

(3) Shànghǎi and Wúxī

The goal of this dissertation is to examine the tonal changes in progress in Wú dialects in the context of urbanization. Places are selected which are: (i) undergoing massive urban growth but with different degrees of urbanization; (ii) contact places between the Mandarin dialect and PTH; and (iii) showing tonal differences within the same dialect group.

The two speech communities this study investigates are Shànghǎi and Wúxī. In terms of criterion (i), Shànghǎi and Wúxī are both located in the Yangtze River Delta Area (YRDA), the fifth largest urban agglomeration in the world by population.[7] Shànghǎi is the center of this metropolitan region and is undergoing a process of metropolitanization. With reference to the degree of urbanization, Wúxī is one of the five sub-centers in the YRDA. The other four sub-centers are Sūzhōu, Nánjīng, Hángzhōu and Níngbō. Since Shànghǎi and Wúxī are in different stages of urbanization, it is hypothesized that they also show differences in the subsequent tonal changes in progress. This is evaluated in Chapter 7.

In terms of criterion (ii), Shànghǎi and Wúxī are both located in the Northern Wú dialect-speaking area. The Wú dialect group is the second largest Chinese dialect group by population. It is not mutually intelligible with Mandarin or PTH. Meanwhile, the Northern Wú dialect-speaking area (mainly in southern Jiāngsū province and Shànghǎi) is bounded to the north by a Mandarin-speaking area, so it is a contact frontier with Mandarin. The speech communities in Shànghǎi and Wúxī offer a unique opportunity to look at linguistic outcomes in the context of urbanization and contact with Mandarin or PTH.

---

[7] According to *Demographia's World Urban Areas* (8th Annual Edition: Version 2, released on July 2012). It can be downloaded from: www.demographia.com/db-worldua.pdf, retrieved on 20 May, 2014.

Finally in terms of criterion (iii), according to the previous Chinese dialectology studies (e.g., Qián, 1992; Xú, Tāng & Qián, 1981), the citation and tone sandhi systems are relatively simple in the Shànghǎi urban dialect while quite complex in suburban dialects; whereas the Wúxī dialects are in-between. So the urban and suburban dialects spoken in Wúxī and Shànghǎi can show tonal differences within the same dialect group. More detailed reviews and discussions of the Wúxī and Shànghǎi dialects are provided in Section 2.4.

(4) Urban and suburban areas

In previous linguistic studies, there has been a dualistic distinction of urban versus rural. This reflects the tradition that dialectology focuses on rural speech while sociolinguistics focuses on urban linguistic realities. However, in the process of urbanization, there are a great deal of areas that cannot be called either city or countryside. These are situated geographically "in between" city and countryside and are semi-urban landscapes in configuration, function, and other characteristics according to the geographical definition (Tacoli, 1998). These urban fringes can play crucial roles in understanding variation in the context of urbanization because they are transition areas and these settings are full of linguistic variation and change. Therefore, beside the surveys in the urban communities of Shànghǎi and Wúxī, this dissertation also investigates the semi-urban areas, or more precisely the suburbs, to identify and document the unfolding tonal changes.

(5) Citation tone and sandhi tone

The tonal variation being described in Shànghǎi and Wúxī includes the variation in citation tones and sandhi tones.

In most Chinese dialects, including Wú, each morpheme has a fixed tone, which is a phonologically contrastive fundamental frequency (F_0) pattern over a syllable. Segmentally identical syllables are lexically distinct by carrying different tones. Most morphemes are monosyllabic. When a morpheme appears alone, its isolated tone is called a citation tone. When citation tones are juxtaposed in running speech, they undergo tonal contextual variations. The process of such tonal alternation is broadly called tone sandhi. Therefore, each morpheme of the Wú dialects has a citation tone and one or more sandhi tones. So this study intends to give a relatively complete picture of tonal variation of these two systems. The sandhi tones of bisyllables are used to investigate the tone sandhi system as a start.

There are two types of tone sandhi in connected speech: phonologically-determined and morphologically-determined (Chen 2000). Regarding the morphologically determined type, it has been a tradition in Chinese dialectal studies to distinguish between a "widely-used pattern" (guǎngyòngshì, 广用式) and a "narrowly-used pattern" (zhǎiyòngshì, 窄用式). A "widely-used pattern" does not have any restriction in the grammatical structure but a "narrowly-used pattern" is primarily limited to the verb-object structure and certain subject-verb, verb-modifier, and coordinated structures (Xǔ, et al., 1981; J. Zhang, 2007; Z. Zhang, 1998).

As a study of sound change, this dissertation is not devoted to discussing the morphologically-determined sandhi process but limited to the phonologically-determined sandhi process. For this reason, only data from the "widely-used pattern" will be used and compared (further discussed in Chapter 3).

To fulfill Goal I, the following two methodological issues also need to be resolved beforehand:

Issue I: the domain of tone in the variationist research

In tone languages, the tonal features are restricted to the voiced part of a syllable. But not all voiced parts of the syllable are tone-bearing. Therefore, the concept "domain of tone" is needed to define which entity of the voiced part is associated with the recognition of tone. In effect, to define the domain of tone is the foundation of any research concerning tonal description.

The domain of tone has been discussed widely but there is still a lack of consensus. Based on impressionistic accounts, Wang (1967), Chao (1968: 19, 25) and Duanmu (2000) claimed that the entire voiced portion of a Mandarin syllable is the domain of tone. To understand "the entire voiced portion", some knowledge about the structure of Chinese syllables is needed. In most Chinese dialects, including Mandarin and Wú, the maximum syllable consists of CGVX, which is a consonantal onset (C), a prenuclear onglide (G), the nucleus (V), and a coda (X) (Chen 2000: 4); and the minimal syllable consists of a single vowel (V) or a single consonant (C). CV, CGV and CVX forms are also common. The entire voiced portion in Mandarin includes voiced (C), i.e. /m/, /n/, /l/; prenuclear onglide (G), i.e., /i/, /u/, /y/; nucleus (V) and nasal coda (X), i.e., /n/, /ŋ/.

Conflicting views on "the entire voiced portion" were presented by Kratochvil (1970), Dow (1972), Howie (1974) and Lín (1995). Kratochvil (1970) argued that the voiced initial consonants should be excluded from the domain of tone while Dow (1972) even proposed to exclude the final nasal consonant. Howie (1974) claimed that the domain of tone in Mandarin is not the entire voiced part of the syllable but is confined to the nucleus and final nasal consonant. The initial voiced consonant and prenuclear onglide are not included in the domain of tone in Mandarin. The perceptual results of Lín (1995) also support the exclusion of initial voiced consonants or prenuclear onglides in monosyllables.

Goal I is to present a description of tonal variation. But the conflicting views presented above obstruct our understanding of tone, as well as further precise descriptions which meet the criteria of variationist research. So it is necessary to conduct my own investigation to define the domain of tone in the Wú dialects. Since Wú dialects have a segmental structure similar to Mandarin, three voiced portions - voiced onset (C), the prenuclear onglide (G) and the final nasal consonant (X) - are tested in Section 4.1 to see whether they are in the domain of tone for Wú.

Issue II: tonal normalization

Because of the inherent anatomical variation, studies of sound change require normalization of the acoustic presentations of different aspects of sounds, e.g., vowels, consonants and tones, in order to accurately model the perception of tones by the human auditory system.

Tone normalization, i.e., the normalization of the fundamental frequency associated with linguistic tone, has received relatively little attention in comparison with vowel normalization (Rose, 1987). In addition, previous tone normalization procedures mainly served for categorizing tones, but did not aim to preserve sociolinguistic variation. Therefore, it is necessary to re-evaluate the effectiveness of tone normalization procedures from the perspective of variationist studies.

In variationist studies, an effective technique for the normalization of vowels is one that: (1) preserves phonemic variation; (2) minimizes anatomical variation; and (3) preserves sociolinguistic variation (Adank, 2003; Adank, et al., 2004; van der Harst, 2011). These criteria assess vowel normalizations effectively for variationist studies, so they are also applied to assess tone normalizations in this study.

Goal II: To elicit the internal, external and extra-linguistic constraints as well as their interaction within the variationist paradigm.

To fulfill this goal, this study intends to investigate linguistic variables that are both linguistically constrained (either by internal or external factors) and socially constrained. For the citation tones, the variations and changes in the lax tone of the Yáng register (yángpíng, yángshǎng and yángqù[8]) meet this requirement because they are expected to be not only constrained by social factors due to urbanization but also internally motivated due to their links with the Zhuó Shǎng Guī Qù (浊上归去), one of the major phonological changes during the Middle Chinese (MC) era. Meanwhile, bisyllabic words initialed with those lax tones of the Yáng register are used as the scope of investigating tonal variables of sandhi patterns.

Zhuó Shǎng Guī Qù literally means that syllables with voiced initials of an MC tonal category Shǎng (rising, 上) merge with another MC tonal category Qù (falling, 去), resulting in a decrease of tonal categories. This tone merger started in the dialects of the northern parts of China during the mid-Tang dynasty (the 8th century), and then widely spread to all of the modern Chinese dialects (Ho, 1988). Yet the cause of this process is still unclear (ibid). Section 2.3 reviews this process in detail and shows that Zhuó Shǎng Guī Qù is a change in progress in the Wúxī dialects. So the data of the Wúxī dialects probably give us a unique chance to trace its dynamic process from sociophonetic perspectives.

The primary empirical task of variationist sociolinguistics is to correlate linguistic variation as the dependent variable with the independent variables (Chambers, 2003), including internal, external and extra-linguistic factors. The contexts underlying varying realizations of the same toneme are investigated to discover the significant factors that impact such variation. Here I propose some factors that could explain the variation from three aspects: internal, external and extra-linguistic motivations.

Since Zhuó Shǎng Guī Qù generally results in the simplification of tonal inventory, possible **internal constraints** are those that can lead to the reduction of tonal categories, i.e. tonemic reduction. The first possible motivation is tonal convergence, which is the tendency for a tonal category to merge with another category due to tonetic similarity. It is expected that similar tonal shapes may prompt tones that are originally from different MC categories to merge. This assumption is examined in the data of Shànghǎi and Wúxī dialects in Chapters 5 and 6 and evaluated in Chapter 7.

In vowel research, the articulatory process of vowel reduction, which is mostly

---

8 The precise meaning of "yángpíng", "yángshǎng" and "yángqù" can be found in the first three paragraphs of Section 2.3.1.

driven by economy of effort, contributes to phonemic reduction (van Bergem, 1989). "Economy of effort" refers to changes that simplify the articulation of sounds. This has been widely adopted as a major principle underlying phonetic reduction. To apply this principle to the study of tone, phonetic tonal reduction as a result of economy of effort may also contribute to the reduction of tonal categories. However, it is difficult to quantify and measure articulatory effort directly. If we cannot justify that tonal reduction found in our data is related to ease of articulation, it is difficult to conclude that the tonal reduction is internally caused by speakers striving for an economic way of articulation. Therefore, a proxy of articulatory effort is needed. The relation between articulatory effort and speaking style has been discussed in a large body of literature, most of it on segmental variation but – to my knowledge - none on tonal variation. Van der Harst (2011: 15-52) provides a detailed review of sociolinguistic and phonetic studies related to vowel space and speaking style. He finds that speaking style is associated with vowel space because different styles are tied to the amount of attention paid to speech. It is assumed that less attention is paid to informal speech and therefore it is produced with less articulatory effort. Reduced attention yields smaller vowel spaces, i.e., more vowel reduction, which is supported by Van der Harst's empirical data (ibid.). This model is summarized in Figure 1.1.

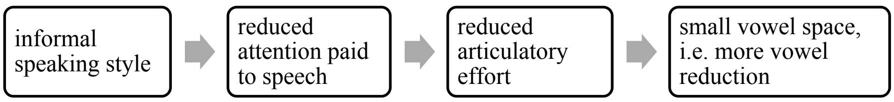


Figure 1.1 The relationship between speaking style, speech attention, articulatory effort and vowel space/vowel reduction

To apply this model to tone, if in an informal speaking style, more tonal reduction shows up in comparison to a formal speaking style, it can be concluded that the model in Figure 1.1 is also suitable for explaining tonal reduction. So this study intends to do an initial examination of the relationship between *speaking style* and tonal reduction. If style is a significant explanatory factor of tonal reduction, then tonal reduction is probably attention-induced and driven by the principle of ease of articulation.

Turning to **external constraints**, this study focuses on the contact-driven motivations in the context of rapid urbanization. During urbanization, dialects spoken in suburban areas inevitably come into contact with the dialects spoken in urban areas. This process is called regional contact. Furthermore, migration from other areas of China brings the Wú dialect speaking area in intensive contact with PTH, which is called supra-regional contact. Hence the possible contact-driven motivations for tonal variation are: (1) regional contact causing the convergence of suburban dialects towards urban dialects by eliminating marked tonal differences within Wú dialects; and (2) supra-regional contact causing the convergence of Wú dialects towards PTH by borrowing PTH tone.

The effect of regional contact can be assessed: (1) by comparison of dialects spoken in two cities, Shànghǎi and Wúxī; and (2) by comparison of suburban dialect(s) and urban dialects spoken in each city. So the factor of *region* should be considered in the research design.

The effect of supra-regional contact can be assessed: (1) by comparing speakers who can speak PTH with those who state they cannot; (2) by examining the effect of corresponding PTH tones on the tonal variation; and (3) by comparing the tonal variation in the shared lexical items of PTH and Wú with the tonal variation in purely Wú lexical items. In sum, three factors are indicators of supra-regional contact: speakers' *PTH competence*, the corresponding *PTH tone* of tested morphemes, and the *lexical attestedness* of each bisyllabic word.

Regarding **extra-linguistic constraints**, the urban versus suburban contrast is one of the most important foci of this study, as has been discussed in Section 1.1 and Goal I. Within each region, this study focuses on working class participants, because in suburban areas the majority of people are working class and the social stratification there is not obvious. Hence, an individual's socioeconomic status is not a factor**.** Then speakers' *sex* and *level of education*, two common social factors, are considered as the extra-linguistic factors. To briefly introduce these two factors, Labov (1990: 205) believed that the sex effect is the clearest and most consistent principle of all sociolinguistic principles. He formulated the following gender paradox: "women conform more closely than men to sociolinguistic norms that are overtly prescribed, but conform less than men when they are not" (Labov, 2001: 293). A large number of Chinese communities have however yielded different results as many of them did not find the significant effects of sex (Xu, to appear). This study intends to explore whether sex is an explanatory factor for tonal variation.

Regarding the effect of *education*, educated people normally have a better command of standard pronunciation and written language than less educated people, and they are more likely to use prestige forms and avoid stigmatized variants. However, the reality in Chinese communities is much more complex. Due to the wars in the first half of the 20th century and the political unrest afterwards,[9] most of the old participants did not have the opportunity to receive formal or good education. Only a few entered high school or college during the relatively stable period between 1949 and the early 1960s or after the College Entrance Examination was resumed in 1977. By contrast, all the young participants received compulsory education for at least nine years. A majority also attended high school or college. Due to this complex situation, it is difficult to make an assumption now whether the factor of education plays a role in tonal variation or what kind of role it plays. We attempt to explore its effect in our dataset.

1.3. Outline of the dissertation

This dissertation consists of seven chapters. This current chapter Chapter 1 has introduced the research topic: language change in the context of urbanization (Section 1.1), the goals and scope of this study (Section 1.2) and has provided the organization of the dissertation (Section 1.3).

Chapter 2 introduces the general background of Wú dialect areas and the speech communities of Wúxī and Shànghǎi. Furthermore, it looks at the historical tonal change Zhuó Shǎng Guī Qù (浊上归去) and the past tonal development recorded in Wúxī and Shànghǎi. The research questions are stated at the end of Chapter 2.

Chapter 3 discusses the methods employed in fieldwork. It describes the research design with respect to data collection, including the criteria for selecting participants, the procedures for recruiting participants, the recording procedures and the speech material for data elicitation. Chapter 3 also discusses the statistical modelling suite for the present data, and the research goals.

Chapter 4 deals with methodological issues. It mainly aims at solving two methodological issues mentioned above: the domain of tone in the Wú dialects, and tone normalization procedures. In addition, the criteria and procedures of tonal segmentation, F_0 measurement and correction, are also discussed.

---

[9] The political unrest includes the "Up to the Mountains and Down to the Countryside Movement" (上山下乡运动) in the late 1960s and early 1970s and the "Cultural Revolution" (文化大革命, 1966-1976).

Chapter 5 and Chapter 6 present and discuss the results of the Wúxī data and Shànghǎi data. These two chapters are structured in the same way, with both including the identification and analysis of linguistic variables in citation form and sandhi form.

Chapter 7 summarizes the findings of this study. The two general goals set in this chapter are looked at again. The significance of this present study and directions for future research are also addressed.

Chapter 2. General background and literature review

In this chapter, Section 2.1 discusses tonal notation system adopted in this study. Then the general background of Wú dialect areas is overviewed in Section 2.2. Section 2.3 introduces the tone system of Middle Chinese and the historical tonal change Zhuó Shǎng Guī Qù (浊上归去). Finally, Section 2.4 offers a description and discussion of the speech communities of Wúxī and Shànghǎi and the past tonal development recorded in Wúxī and Shànghǎi.

2.1. Tonal notation

The standard tool for auditory (impressionistic) research on tone is the iconic system documented in the International Phonetic Alphabet (IPA, Revised to 2005). As can be seen in Figure 2.1, the IPA provides two alternative sets of symbols for transcribing tone: (i) a diacritic marked on the top of a letter (e.g., ē) and (ii) a vertical stroke with a line preceding it (e.g., ˧). The latter set is also called “tone letters”. According to the *Handbook of the International Phonetic Association* (International Phonetic Association, 1999: 14), the set of diacritics is used for “languages in which tonal contrasts depend predominantly on the pitch height in each syllable” while tone letters are used for “languages in which lexical contrasts are predominantly dependent on the pitch movement on each syllable”. As Wú Chinese is a language variety using pitch movement to distinguish lexical contrasts, tone letters are commonly used. So, this study will not use the IPA diacritics for transcribing tone.

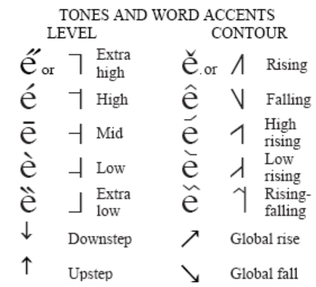


Figure 2.1 Chart for “tones & word accents” in the International Phonetic Alphabet (reproduced from www.langsci.ucl.ac.uk/ipa/tones.html)

In Figure 2.1, the left column of "LEVEL" shows the IPA system for transcribing tone divide the range of a speakers' voiced into four intervals, marked by five scales. This five-point scale adopts Chao's tonal notation system (1930, 1968). Chao's system not only includes graphical representations (˩, ˨, ˧, ˦, ˥) but also numerical representations 1 to 5 (1=lowest, 5=highest), which are called Chao's tone numerals. Combinations of two tone numerals are used to transcribe level tones (e.g., 11, 55), rising tones (e.g., 13, 35) and falling tones (e.g., 31, 53), where the first numeral indicates the pitch height of the starting point while the last numeral indicates the pitch height of the ending point. Combinations of three tone numerals are used to transcribe peaking tones (e.g., 131, 253) and dipping tones (e.g., 214, 425), where the middle numeral indicates the pitch height of the turning point. Previous tonal studies of Chinese predominantly use Chao's tone numerals rather than IPA tone letters; therefore, for the ease of comparison with other studies, Chao's tone numerals are adopted in the present study for the transcription of both citation tone and tone sandhi.

Besides Chao's tone numerals, previous studies related to the Wú dialects also have also used tone feature systems to tones, e.g., Chan & Ren (1989). I will follow the notations used in these studies and convert them into Chao's tone numerals on the basis of Zhang's (2006: 182-183) model. Chao's system is preferable to a tone feature system for three reasons. First, the majority of the literature related to the present study uses Chao's tone numerals to transcribe citation tone and tone sandhi. Second, tone feature systems are mainly used in phonological studies for the purpose of presenting the autosegmental processes of delinking and spreading, as well as observing the constraints against crossing association lines (e.g., Goldsmith 1976). As this is a sociophonetic study of language variation and change, the essential function of tone feature systems will be rarely used. Third, Chao's tone numerals provide more detailed information. It has five levels, enabling us to transcribe more variation in the phonetic realization of pitch height than the tone feature systems, which normally have only two or three levels, and consequently are less precise. When using Chao's tone numerals, the phonetic contrast of tone is indicated by square brackets "[]", while the phonemic contrast is indicated by slashes "//". In addition, in the presentation of tone sandhi forms, this study adopts Chen's (2000: F43) way of separating syllables. A dot separates tone numerals or H, M, L in the sandhi form. For instance, to distinguish LM.L (=LM+L) from L.ML

(=L+ML), or 13.1(=13+1) from 1.31(=1+31). When discussing any sandhi tones initialed with the same tone, e.g., T2, the expression T2.X is used. When discussing sandhi tones initialed with the same tone (e.g., T2), while the second syllable is from several consecutive tones (e.g., T3 to T6), the expression T2.T3-6 is used.

2.2. An introduction to the Wú Dialects

Wú (吴) dialect is a Chinese dialect group mainly spoken in the southern Jiāngsū (江苏) province, Shànghǎi (上海) Municipality, most of Zhèjiāng (浙江) province and part of Jiāngxī (江西) province (Norman, 1988). Wú has 77.2 million speakers and it is the second largest dialect group in China, according to the statistics published by *Ethnologue* (Lewis, 2009).

Wú is based on the geographical name for Wú Guó (the state of Wú, 吴国, 11th century BC - 473 BC), a vassal state of Xī Zhōu (Western Zhōu Dynasty, 西周, 1046 BC - 771 BC). Wú Guó was located at the mouth of the Yangtze River. There is a long history of language contact and culture contact between Wú Guó and its neighbor and later conqueror, Yuè Guó (the state of Yuè, 越国, ? - 222 BC, annexed Wú in 473 BC). The culture in the south of Jiāngsū Province and Zhèjiāng Province is generally called Wú Yuè Culture. Following the administrative division, Chao (1967) divided the Wú dialect into a "northern type group" and a "southern type group" (or more precise expressions "Northern Wú" and "Southern Wú" summarized by Norman (1988: 199). Northern Wú refers to the dialects mainly spoken in southern Jiāngsū and Shànghǎi, which overlap to a large extent with the ancient Wú Guó, while Southern Wú mainly covers the dialects spoken in Zhèjiāng, associated with the ancient Yuè Guó. Apart from the former administrative boundaries, the distinction between Northern Wú and Southern Wú is also based on the isoglosses of linguistic features. In the 1980s, Chinese dialectologists discussed the intra-Wú classification (Yán, 1995) and their consensus at that time[10] has been adopted in the *Language Atlas of China* (Wurm, et al., 1988), consisting of six subgroups across Northern and Southern Wú. The subgroup Tàihú Piàn basically corresponds to Northern Wú and the remaining five constitute Southern Wú.

Figure 2.2 presents the geographic locations of the Wú dialect group and its

[10] Hè (2012) proposed a new subgrouping of Southern Wú, but this issue is not the focus of this study and it will not be discussed in further detail.

subgroups. This map is based on Map B9 “Wu group” in *Language Atlas of China* (Wurm, et al., 1988) and Map 4 “The Wu Dialect” in *Introduction to Chinese Dialects* (Yan, 2006).

1. **Tàihú Piàn (Tàihú Sub-Group, 太湖片)**
   (1) Pílíng Xiǎopiàn (Pílíng Cluster, 毗陵小片)
   (2) Sūhùjiā Xiǎopiàn (Sūhùjiā Cluster, 苏沪嘉小片)
   (e.g., the Shànghǎi dialects, the Wúxī dialects)
   (3) Sháoxī Xiǎopiàn (Sháoxī Cluster, 苕溪小片)
   (4) Hángzhōu Xiǎopiàn (Hángzhōu Cluster, 杭州小片)
   (5) Línshào Xiǎopiàn (Línshào Cluster, 临绍小片)
   (6) Yǒngjiāng Xiǎopiàn (Yǒngjiāng Cluster, 甬江小片)
2. **Tāizhōu Piàn (Tāizhōu Sub-Group, 台州片)**
3. **Ōujiāng Piàn (Ōujiāng Sub-Group, 瓯江片)**
4. **Wùzhōu Piàn (Wùzhōu Sub-Group, 婺州片)**
5. **Chǔqú Piàn (Chǔqú Sub-Group, 处衢片)**
   (1) Chǔzhōu Xiǎopiàn (Chǔzhōu Cluster, 处州小片)
   (2) Lóngqú Xiǎopiàn (Lóngqú Cluster, 龙衢小片)
6. **Xuānzhōu Piàn (Xuānzhōu Sub-Group, 宣州片)**
   (1) Tàigāo Xiǎopiàn (Tàigāo Cluster, 太高小片)

Figure 2.2 The Wú dialect area, adopted from Wuem et al. (1988) and Yan (2006)

It is noteworthy that Northern Wú only overlaps with one subgroup - Tàihú Piàn - while Southern Wú is composed of five subgroups, which indicates that Northern Wú is much more homogeneous than Southern Wú. The relative homogeneity of Northern Wú is possibly caused by two factors: (1) the geography of the Yangtze delta plain and (2) its closer proximity to the Mandarin speaking area. Among Northern Wú, the dialects spoken in Shànghǎi and Wúxī, which are the focus of this dissertation, both belong to the Sūhùjiā Xiǎopiàn.

The most distinctive phonological feature of the Wú dialects is the retention of a three-way contrast of Middle Chinese (MC) initials (Chao, 1967). In addition to voiceless aspirated and unaspirated initials (including stops and affricates), a distinctive set of MC voiced initials are retained, having the phonetic nature of voiced, murmured or breathy voice (further discussed in Chapter 3 and Chapter 4). But Norman (2003) claimed that this feature cannot be treated as the defining characteristic of the Wú dialects since voiced (or murmured) initials are also found in other dialect groups such as the Xiāng (湘) dialects. According to Norman (2003), the most distinctive feature of Wú is its tonal pattern.

> The feature that most clearly sets Wú off from Mandarin is its tonal pattern; typically, all tonal categories are divided into two registers. The more conservative Wú dialects maintain a distinction between the upper and lower shàng[11] tones – for example, Wúxī, Chángshú and Wēnzhōu. More innovative dialects (Shànghǎi, Chángzhōu, and Sūzhōu) have merged the yángshàng with yángqù. But even a very innovative dialect like Shànghǎi still maintains the register system very faithfully in that only in very rare cases do upper and lower register tones merge. The tonal conservatism of Wú serves to divide it from the those Xiāng dialects in which voiced obstruent initials are found but in which there is no yángshàng category nor is there a register distinction for the rù category which is usual in Wú dialects.[12]
>
> (Norman 2003: 78)

Before discussing the tonal issues of Wú more precisely, the classification, including a labeling scheme for the MC tonal categories, will be introduced, along with relevant research results on tonal development that are generally accepted in the historical phonology of Chinese.

---

[11] "Shàng" refers to "shǎng" in other places of Norman (2003), including his own work of Norman (1988).
[12] Terms like "register", "upper and lower shàng", "yángshàng", "yángqù" and "rù category" are all terms used in the Chinese dialectology and historical phonology. The statement of their meaning can be found in the first three paragraphs of Section 2.3.1.

2.3. Linguistic Variables: from the phenomenon Zhuó Shǎng Guī Qù

2.3.1. Tone system in MC

MC is considered to be the ancestor of most modern Chinese dialects. *Qièyùn* (*Spelling Rimes*, 《切韵》), the most authoritative Chinese rime dictionary published in 601 CE, was a record of early MC. Based on *Qièyùn*, MC possessed four tonal categories: Píng (level, 平), Shǎng (rising, 上), Qù (falling, 去) and Rù (entering[13], 入), which were collectively called Sìshēng (four tones, 四声). Here "tonal category" means the tonal property of monosyllables that are perceived as having the same pitch contour and pitch range. Tonal category is a distinct feature of lexical meaning in addition to their segmental features. In *Qièyùn*, each monosyllable in a stressed form is classified into one of the four tonal categories. The exact phonetic values of these four tonal categories are unknown but their text descriptions can be found. Mei (1970: 104) reconstructs the MC tonal categories around the 8th century as follows:

> Píng: long, level and low, with a higher and a lower allotone
> Shǎng: short, level, and high, its lower allotone having merged with the departing tone
> Qù: slightly drawn out and hence longish
> Rù: short

The first three tones, i.e., Píng, Shǎng and Qù, are also collectively referred to as Shūshēng (lax tone, 舒声) because they have longer duration than Rù which sounds abruptly due to the glottal stop in coda.

Each of these tonal categories can be divided into two registers. Register refers to the effects of initial consonants on the tones: in general, syllables with voiceless initials in MC are traditionally classified as Yīn (upper, 阴), implying higher pitch, while those with voiced initials are classified as Yáng (lower, 阳), implying lower pitch. It is difficult to verify when the register distinction evolved, but going back to *Qièyùn*, its chief editor Lù Fǎyán (陆法言) used two sets of concepts in the preface: (1) qīng (clear, 清) versus zhuó (muddy, 浊), and (2) qīng (light, 轻) versus zhòng (heavy, 重) as two distinctive elements of MC phonological system. It was believed that these concepts were associated with the register. For instance, Bernhard Karlgen

[13] "Entering tone" is the Chinese calque of "Rù Shēng" (入声). In tonology, this term is more commonly used as "checked tone", referring to a tone realized on a syllable that ends in a stop consonant, such as p, t, k, and glottal stop.

advocated that a register distinction already existed at the time of the *Qièyùn* (Norman, 1988: 53). It is certain that by the late Táng Dynasty (618 - 907 CE), the original four tonal categories were split into two registers (Mei, 1970). When the yùntú (rime table, 韵图, a Chinese phonological model tabulating the syllables of the *Qièyùn* system with the initials on the horizontal grids, the finals and tones on the vertical grids) became available in the Sòng Dynasty (960 - 1279 CE), the terms of "qīng" (清) and "zhuó" became interpretable. From the way these terms were used in the tables, it became clear that they corresponded roughly to the modern phonetic concepts of "unvoiced"[14] and "voiced" (W. S. Wang, 1982). The eight combinations of two registers and four tonal categories were called bādiào (八调). In the word "bādiào", "bā" means eight and "diào" literally means tone, like "shēng" (声) in "Sìshēng" (four tones). "Shēng" refers to the modern linguistic concept of tonal category, but it is uncertain whether the concept "diào" referred to the phonemic tonal category as well, or just to an allotone. Wang (1982) claimed that there were eight tone shapes conditioned by the distinction between voiceless and voiced initials due to physiological factors - in other words, the coarticulation was influenced by the initial contrast. So the term of "diào" defines at least the distinction of tonal shapes and its precise meaning will be discussed from the dynamic view in Section 2.3.2. In sum, Sìshēng Bādiào (four tones and eight tonal shapes, 四声八调) is the essential feature of MC tone system. Table 2.1 shows the four tonal categories and their combinations with register, where both the English and Chinese names are provided. Some linguists (e.g., Norman 1988; Chen 2000) use the numerical labeling scheme for these categories in Table 2.1 and this dissertation will adopt its tone labeling scheme accordingly.

Table 2.1 MC tone system and labeling scheme

Register	**Tonal category**			
	Píng	**Shǎng**	**Qù**	**Rù**
Yīn	T1 yīnpíng upper level	T3 yīnshǎng upper rising	T5 yīnqù upper falling	T7 yīnrù upper entering
Yáng	T2 yángpíng lower level	T4 yángshǎng lower rising	T6 yángqù lower falling	T8 yángrù lower entering

[14] "Unvoiced" is a term used by Wang (1982). It equals "voiceless", describing sounds being pronounced without the larynx vibrating.

The above eight groups do not suffice to explain the tonal development from MC to modern dialects. In some modern dialects, tonal categories are not only conditioned by initial contrast of voicing, but also by the subdivisions of articulation manner within the Yīn and Yáng register. Yīn not only includes voiceless unaspirated obstruents but also aspirated stops. Aspiration is responsible for the tonal subdivision of words with ancient voiceless initials in 67 Chinese dialects (Ho, 1988). So the tones of the Yīn register were split into two categories. The symbols “A” and “B” are followed by the tonal categories of the Yīn register to mark words with voiceless unaspirated obstruents and aspirated stops, respectively, e.g., T1A, T1B, T3A, T3B. Yáng is used in connection with the feature of initial sonorancy. In many Chinese dialects including Wú, it has been proven that words with obstruent initials and words with sonorant initials have undergone different paths of tonal development: sonorant initials sometimes condition the same tonal changes as the voiced initials, and sometimes condition the same changes as the voiceless initials (Chang, 1975; Ho, 1988; Norman, 1988). Accordingly, the symbols “A” and “B” following the tonal categories in the Yáng register are used to refer to words initialed with voiced obstruents and sonorants, respectively, e.g., T2A, T2B, T4A, T4B. So far, there are 16 possible tonal combinations in total. The design of the word list used for eliciting reading data will be based on these 16 combinations (more details will be given in Chapter 3).

2.3.2. The development of MC tones into modern dialects

It is generally accepted that the register was not originally an independent feature occurring before the devoicing of initial consonants (Ting, 1996; W. S. Wang, 1967; Sagart, 1998). Register splits become phonemic after the voicing distinction of initials is eliminated through historical change, either by voiced initials becoming voiceless or voiceless initials becoming voiced. At that point, register becomes a distinct feature and the number of tonal categories increases to maximally eight. Register splits can also become phonemic after the parallelism in tonal contours between high and low allophones is disrupted by other phonetic processes (Sagart, 1998).

Modern Chinese dialects have from three to ten distinctive tones for their tonal systems. In contrast with MC, the increases in number were mainly caused by tonal split from the four traditional tonal categories conditioned by register. Therefore, “tone split” was traditionally accepted as a perspective to describe the tonal development, although the whole process must involve a reintegration or merging at a later stage, resulting in fewer tones.

In the view of split, Chang (1975) presents a detailed investigation of the tonal development based on data of modern Chinese dialects and developed a tonal taxonomy. He examined hundreds of Chinese dialects and pointed out that:

> Tonal splits conditioned by voiceless versus voiced initials have failed to affect some tonal categories in some dialects. Obviously, a tonal split needs not involve all tones simultaneously; tonal splits may have taken place in succession, and spread from one tone to another.
>
> (Chang 1975: 643)

He classified the dialects based on the occurrences of splits in different tonal categories: (1) dialects that underwent ancient *Píng* split, (2) dialects that underwent ancient *Shǎng* and *Qù* split and (3) dialects that underwent splits of yángshǎng and yángqù tones. In the same section, dialects that underwent subsequent mergers were also listed; they were classified into (4) dialects with the merger of T4 and T6, and (5) dialects with the merger of T4, T5 and T6, because these mergers happened only after the tonal splits. Thus, the splitting view is useful for tonal description and taxonomy (Wáng, 2009). But, when considering the tonal development issue, the key questions regarding the splitting view are when and how the tonal split took place, since the merger is taken to have arisen from the termination of splitting. Moreover, the answer to the above two questions cannot be generalized because the answers should be different for different dialects - in other words, they are dialect-specific.

Wang (1967) first proposed a hypothesis of "tone-shape merge", i.e., tone merger from eight contours to fewer. This hypothesis was based on the consensus of eight tonal contours of MC rather than four tonal categories and based on the same materials from Chinese dialectology and historical phonology. However, the merger hypothesis traced the same tonal development from the opposite angle, i.e., in terms of successive mergers instead of splits. The "tone shape-merge" hypothesis was proven conclusively by Wang's follow up study (Wang, 1982) and by Pān (1982). It allows us to conclude that the fewer tone shapes occur in a dialect's tonal system, the more advanced it is. The modern Mandarin tonal systems are thus the most advanced, since they have undergone the most mergers; the Wú dialects and Cantonese dialects are relatively conservative due to their very close resemblance to the MC tonal system, preserving seven or eight tonal categories. In sum, the merging view is more theoretically relevant to studies from the dynamic perspective. Moreover, the merging view is suitable for explaining the tonal development from proto-Wú to some of the modern Wú dialects (Ting, 1996) as well as the Cantonese case.

As stated in Section 2.3.1, the Wú dialects have preserved voicing contrast in initials and had four ancient tones split into two subcategories: syllables of the Yīn register generally have a higher tone than those from Yáng register. However, the pitch contours of the paired subcategories were no longer the same. Otherwise stated, the register split not only conditioned the high-low contrast of pitch levels but also the differences of pitch contour conditioned the register split. For instance, the *Píng* tone of Wú was reconstructed as T1-/55/, T2-/22/ (Ting, 1996), two level allotones having high-low contrast. In the present-day Wúxī urban dialect, quite a few T1 words still have high level tone /55/, while T2 was developed into /213/ (Xú, 2007). From Chao's (1928) data collected in the 1920s, more than one third of the Wú dialects (12/33) had eight tonal categories, and more than one third had seven. Sixty years later, in the 1980s, nearly all of the 33 dialects had tone mergers for their citation tone system (Qián, 1992: 441 - 442). This dissertation aims to study the dynamic progress of tonal change in the Wú dialects, so it will adopt the hypothesis of "tone shape-merger" because of both its research subject and dynamic view.

Regarding the terms of the types of "tone shape-merge", Pān (1982) generalized two types of merger, type X and type Y:

Type X is tonal shape merger within the four original tonal categories, that is, the merger of T1-T2, T3-T4, T5-T6, or T7-T8. Type X mergers would occur under two conditions: (1) the neutralization of initial voicing contrast and (2) similar tonal shapes within each category. A typical example of Type X merger is the formation of the Qù category in modern Mandarin. Mandarin syllables of T5 and T6, namely syllables initialed with voiceless consonants and voiced consonants of Qù, merged their register splits together and became one tone.

Type Y is a merger across the four original tonal categories, but within the register of Yīn or Yáng, which occurred among odd (T1, T3, T5 and T7) or even (T2, T4, T6 and T8) numbered tones. Type Y mergers occurred before the devoicing of voiced obstruent initials. Type Y mergers can often be found in the Wú dialects, which preserved the register distinction faithfully. A typical Type Y merger is the phenomenon Zhuó Shǎng Guī Qù, which is detailed below.

2.3.3. Zhuó Shǎng Guī Qù

"Zhuó Shǎng Guī Qù" literally means that syllables with voiced initials of Shǎng (T4) were merged into Qù (T5 or T6), resulting in a decrease of tonal categories. The phenomenon of Zhuó Shǎng Guī Qù is one of the major phonological changes during the MC era. This tone merger started in the dialects in the northern parts of China during the mid-Tang dynasty (the 8th century), then it widely spread to all of the modern Chinese dialects (Ho, 1988). But dialects beyond the confines of Mandarin did not all fully comply with the same phonological rules as Mandarin did.

The standard form of Zhuó Shǎng Guī Qù (merger of T4A with T6) is found in Guānhuà Fāngyán (the Mandarin dialect group, 官话方言) and its phonological rules have been influencing all the Chinese dialects, including Wú. In Guānhuà Fāngyán, syllables of Shǎng merging into Qù are all initialed with voiced obstruent consonants (T4A), while syllables with sonorant initials of Shǎng (T4B) developed to Yīnshǎng (T3). Normally, syllables with voiced obstruent and sonorant initials changed together, as in *Píng*, *Qù* and *Rù*. It is only in the Shǎng that syllables with sonorant initials changed together with syllables with voiceless initials. It is still uncertain what the forces creating this exception were. Regarding the origin of Shǎng, Mei (1970) hypothesized that it was developed out of an earlier glottal stop ending. The final glottal stop in Shǎng caused the assimilation of its preceding vowel(s) and initial consonants. The effects were reflected in the creaky voice of the vowel(s) and the glottalization of sonorant initials, e.g., /m/ > /ʔm/. Though this hypothesis has been accepted by many linguists (e.g., Pulleyblank, 1978; Sagart, 1998; Zhengzhang, 2003: 208), it still lacks empirical support.

Lǐ Róng (1966) pointed out that in the Wēnlǐng (温岭) dialect (a southern Wú dialect), MC sonorant initials of Shǎng were articulated in two distinct manners: Type 1, nasal and lateral initials have a glottal stop [ʔ] as the onset, such as /ʔm/, /ʔn/, and /ʔl/; Type 2, nasal and lateral initials were followed by a voiced aspiration: /$m^ɦ$/, /$n^ɦ$/, and /$l^ɦ$/. Type 1 conditions a higher pitch, as with the T3 syllables and Type 2 conditions a lower pitch, as with the T4A syllables. These materials presented by Lǐ (1966) are inspiring. On the one hand, it represents empirical evidence in modern dialects that the T4B change is affected by the characteristics of initials; on the other hand, it reminds us of the resemblance between the contemporary Wú dialects and MC. In the Wú dialects, there were not only initial glottal stops (e.g., the Wēnlǐng dialect, the Sōngjiāng 松江 dialect), but also final glottal stops (e.g., the Wēnzhōu 温州 dialect, the Huángyán 黄岩 dialect). Meanwhile, the Wú dialects have their own type of Zhuó Shǎng Guī Qù in addition to the standard form in Mandarin, as summarized by Ho (1988), which he called the Wú type.

In the majority of Wú dialects, T4A and T4B morphemes merged into T6, e.g., the dialects of Bǎoshān (宝山), Wúxiàn (吴县) and Wújiāng (吴江) (all investigated in the 1920s by Chao). However, a few Wú dialects also adopted the Mandarin type by merging T4B with T6 and T4B with T3, e.g., dialects of Jiāngyīn (江阴, investigated in the 1920s), Shànghǎi (investigated in the 1950s) and Hángzhōu (杭州, investigated in the 1920s). Still others had the same T4A and T4B and conservatively kept them independent of T6, e.g., dialects of Wēnzhōu (investigated in the 1950s), Wúxī (无锡, investigated in the 1920s) and Sōngjiāng (investigated in the 1920s). Ho (1988) deduced that the T4 situation in one Wú dialect was subjected to its contact degree with Mandarin: the more contact, the higher possibility of using the Mandarin type.

These classifications are mainly based on data collected by Chao Yuen Ren in the 1920s. Chao investigated a total of 33 dialects, including the Wúxī urban dialect and five dialects from the Shànghǎi area: the city of Shànghǎi, Sōngjiāng center of Sōngjiāng County (hereafter SJC), Nánhuì Zhōupǔ county (南汇周浦, hereafter NHZ), Bǎoshān Shuāngcǎodūn county (宝山双草墩, hereafter BSS), and Bǎoshān Luódiàn county (宝山罗店, hereafter BSL). At that time, Chao found that the Shànghǎi urban dialect has six/seven tones, T4 and T6 have already merged and T2 was merging with them. In the suburban area of Shànghǎi, two sites in Bǎoshān both had T4 merged with T6. Sōngjiāng's T4B merged with T6 while keeping T4A independent; Zhōupǔ had distinct T2, T4 and T6. Wúxī had separate T2, T4 and T6, like Zhōupǔ, which systematically correspond with the eight tonal categories of MC. Sixty years later, Qián (1992) investigated the same survey points as Chao had, and found that most of the Wú dialects showed an on-going merger of T4 with T6. However, the Wúxī urban dialect deviated from most other Wú dialects by merging T2 and T6 before merging T4 with T6. A recent acoustic study of the Wúxī urban dialect shows its T2 and T6 have merged completely while T4 was in the process of merging with T2/6 (Xú, 2007). After reviewing the Type Y merger of the Yáng register (merger of T2, T4 and T6) in Shànghǎi (see more details in Table 2.3 and Table 2.4), it was found that Wúxī differs from all the Shànghǎi dialects in its merging path and represents a special type. Wúxī provides a unique chance to trace the dynamic process of Zhuó Shăng Guī Qù from a sociophonetic perspective. Meanwhile, though Zhuó Shăng Guī Qù was observed in the suburban Shànghăi in the 1920s, it was not fully completed in the old variety of the 1980s (Qián 1992, see more details in Table 2.10). In the new variety of the 1980s, Zhuó Shăng Guī Qù was completed but T2 in the Băoshān dialect and Sōngjiāng dialect were undergoing mergers towards T4/6. So investigations into Type Y merger of the Yáng register in Shànghăi urban and suburban varieties can also provide greater insights into the phenomenon of Zhuó Shăng Guī Qù.

For this reason, this dissertation will use Shànghăi and Wúxī as the data points for the investigation on Zhuó Shăng Guī Qù. Ongoing changes in respect to T2, T4 and T6 are examined in both the urban and suburban areas. Taking into consideration the size of the city (defined by area and population) and the degree of linguistic heterogeneity, in Shànghăi four sites were investigated (one urban and three suburban) while in Wúxī two sites were investigated (one urban and one suburban). Sōngjiāng, Băoshān and Nánhuì were chosen as Shànghăi suburban sites because they have been documented in Chinese dialectology. Huàzhuāng (华庄) was chosen as a Wúxi suburban site because it is adjacent to Tài Lake. Its location isolates it from the contact with other dialects except the urban variety. Section 2.4 gives an overview of these sites and the documented tonal developments.

2.4. Speech Communities and past tonal developments

2.4.1. The speech community of Shànghăi

In view of intersecting geographical and linguistic dimensions, a few concepts should be clarified. Shànghăi area refers to the administrative areas of Shànghăi at present and in the past, including the urban areas as well as county regions within its jurisdiction. Taking Shànghăi as a single speech community, the Shànghăi dialect (SHD) refers to all the dialect varieties spoken in the Shànghăi area that share a number of common phonological features (Chen 2000: 47-50). Shànghăi is administered as a municipality of China with province-level status and administratively divided into 17 county-level divisions: 16 districts and 1 county in 2011. The city proper is bisected by the Huángpŭ (黄浦) River. The historic center of the city, the Pŭxī (Huángpŭ West Bank, 浦西) area, is located on the western side of the Huángpŭ River, while the newly developed Pŭdōng (Huángpŭ East Bank, 浦

东) is on its eastern bank. Eight districts in the Pǔxī area are collectively referred to as urban Shànghǎi. These eight districts are Huángpǔ District[15], Xúhuì District (徐汇区), Chángníng District (长宁区), Jìng'ān District (静安区), Pǔtuó District (普陀区), Zháběi District (闸北区), Hóngkǒu District (虹口区) and Yángpǔ District (杨浦区), region “a” - “h” in Figure 2.4. Following the tradition of Chinese dialectology (Xǔ & Tāng, 1988) , the variety spoken in these eight districts is defined as Shànghǎi urban dialect (hereafter SHUD). Suburban Shànghǎi includes the remaining eight districts and one county which were all counties before the 1980s. These regions are Pǔdōng New District (浦东新区), Bǎoshān District (宝山区), Mǐnháng District (闵行区), Jiādìng District (嘉定区), Jīnshān District (金山区), Sōngjiāng District (松江区), Qīngpǔ District (青浦区), Fèngxián District (奉贤区) and Chóngmíng County (崇明县). The language varieties spoken in suburban Shànghǎi are called by their place names. For instance, the dialect spoken in Pǔdōng New District is called Pǔdōng vernacular (浦东话) or the Pǔdōng dialect (浦东方言). It should be pointed out that Nánhuì District (南汇区, Nánhuì County until 2001) was merged into Pǔdōng New District in 2009, but the dialect spoken in the old Nánhuì area is still called the Nánhuì dialect. Language varieties spoken in the suburban Shànghǎi are collectively called Shànghǎi suburban dialects.

Shànghǎi is the largest city in the world in terms of population[16], located alongside the Yangtze River Delta Area (YRDA, c.f. Figure 2.4) at the mouth of the Yangtze River. Knowledge of the historical development of language in Shànghǎi is essential to understand the present day language situation. “Shànghǎi” first appeared in *Sòng Huìyào Jígǎo* (*Compendium of the institutional history in Song Dynasty,* 《宋会要辑稿》) in 1077 as the name of a regional trade center on the side of Sōngjiāng County (called Huátíng County at that time). In 1291, Shànghǎi became a county town of Sōngjiāng Prefecture (松江府) and a large population center. At that time, the Shànghǎi dialect was a branch of the Sōngjiāng dialect, Wú Chinese. As late as the third quarter of the 19th century, the Sōngjiāng dialect was still commonly spoken in what is nowadays the Shànghǎi urban area (Lu, 2004). Due to its geographical situation, Shànghǎi started playing a more important role as a trading port in the lower Yangtze River region in the 18th century and got international attention.

---

[15] Huángpǔ District merged Nánshì District (南市区) in 2000 and Lúwān District (卢湾区) in 2011.
[16] http://www.geohive.com/earth/cy_notagg.aspx, retrieved on May 20, 2014.

Figure 2.3 Map of Yangtze River Delta Area (YRDA)

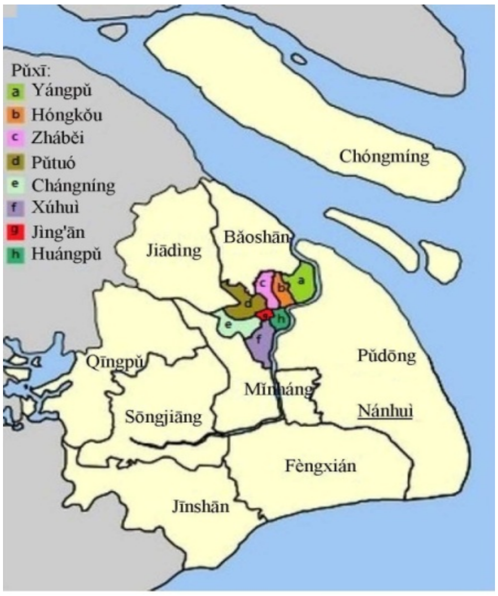


Figure 2.4 Map of Shànghǎi

Being one of the first five treaty ports of China to open after the First Opium War (1839 - 1842), the city grew explosively as a trade hub. Concessions and foreign settlements had been established in Shànghǎi by Great Britain, the United States and France and remained until 1943. Its population increased suddenly and dramatically with businessmen and refugees immigrating during this semi-colonial period. Its population increased from 0.5 million in 1852 to 4.5 million in 1950 (Chen 2000: 45). In this period, the migrants were mainly from nearby Jiāngsū and Zhèjiāng province. It is estimated that there were 3.5 million people in the city in 1934: 25% natives, 39% from Jiāngsū Province, 19% from Zhèjiāng, and the remaining 17% from other parts of the country (Chen 2000: 3). The early migrants of Shànghǎi were mostly from the Wú dialect speaking area, so the Shànghǎi dialect was highly influenced by its neighboring Wú dialects, especially the Sūzhōu and Níngbō dialects, and gradually changed into a new variety of Wú, a lingua franca with mixed natures. However, there had not been a large-scale migration to the surrounding suburban areas; dialects in these areas changed slowly. So SHUD and the surrounding suburban dialects diverged in the early 20th century with the former changing fast and the latter being conservative. In the mid-twentieth century, the population migration was controlled by the household registration policy and SHD remained relatively stable during that period (Chen 2000: 8).

Since the "reform and opening-up policy" implemented in 1979, millions of Chinese have been moving to the cities in search of a better life and China has been urbanized at a pace and scale never before seen in world history. Entering the 21st century, China is in a new stage of growth as an urban society. According to the 2010 National Population Census, nearly half of the population (49.7%, 665.57 million) lives in urban areas, compared to 37% in 2000[17]. Urbanization has become the new engine in promoting the social and economic progress in China following the engine of industrialization.

The processes of industrialization and urbanization are largely responsible for bringing about increased mobility of populations and the expansion of cities, which inevitably result in linguistic diversities through language contact and the formation of bi/multilingual communities. The sociolinguistic complexity associated with urbanization can arise in two important ways: (1) *Urbanization Consolidation* - where in-bound migrants such as rural dwellers or non-indigenous foreigners take

---

[17] Ma Jiantang, 2011, Press Release on Major Figures of the 2010 National Population Census, http://www.stats.gov.cn/english/NewsEvents/201104/t20110428_26448.html, retrieved on May 20, 2014.

up residence in significant numbers in urban centers; (2) *Urbanization Spread* - urban lifestyles and economic conditions spread to encompass a different population in rural areas or possibly in a foreign community (Tsou, 2012). These two patterns can both be observed in Shànghǎi and might be reflected in the language use.

From the 1980s on, the migration movement has returned to Shànghǎi, but with a new trend: immigrants do not only originate from Wú dialect areas, but large numbers also come from Mandarin speaking areas. These new migrants have brought the Shànghǎi dialect into contact with Mandarin and pushed it closer to Mandarin (You, 2010). Meanwhile, with the large scale of urban renewal constructions and redevelopment projects, millions of local Shànghǎi residents have moved from the inner city to its borders and newly developed suburbs. Shànghǎi's rapid transit system also contributes significantly to urbanization spread. Subway and light railway lines have extended to every core urban district as well as neighboring suburban districts. Therefore, the demographic situation, in both the urban and suburban Shànghǎi has changed considerably.

The total population of Shànghǎi was 23 million persons as of 2010, including a floating population of 8.97 million. The floating population increased by 5.92 million in the last 10 years, which accounts for 89% of the total population growth in Shànghǎi. This indicates that at least one third of the Shànghǎi population is migrants, not native speakers of the Shànghǎi dialect. The increasing mobility of population has promoted the usage of Pǔtōnghuà in Shànghǎi as well as China's government intervention by way of language policy. Pǔtōnghuà has been changed from a seldom-used variety to a commonly-used variety in public places in as little as two decades and has acquired a growing number of speakers among the younger generation (Xue, 2010).

The three suburban dialects selected in Shànghǎi were the Sōngjiāng dialect, the Bǎoshān dialect and the Nánhuì dialect. These three dialects were selected mainly because they have been documented extensively; dating back to Chao (1928).

Sōngjiāng was a county of Shànghǎi in 1958 and became Sōngjiāng District in 1998. It is located in the southwest of Shànghǎi with a land area of 605.64 km^2 and a population of 1,582,398 as of 2010. Its main town is "Sōngjiāng New City". Shànghǎi was administered by Sōngjiāng prefecture from Yuan Dynasty (1271-1368) in 1292 until officially becoming a city in 1927[18]. As it has been less influenced by urbanization than Sōngjiāng central area, Sōngjiāng preserves more historical and architectural treasures as well as the local dialect features. The survey was conducted in two subdistricts in "Sōngjiāng New City", Yǒngfēng Subdistrict (永丰街道) and Zhōngshān Subdistrict (中山街道).

Bǎoshān was a county of Sōngjiāng and became Bǎoshān District in 1988, the first one among seven newly-incorporated districts. It is located in the north of Sōngjiāng with a land area of 424.58 km² and a population of 1,905,000 as of 2010. Baosteel, the second-largest steel producer in the world in 2010, is located in Bǎoshān and brings many migrants from all parts of the country. The survey was conducted in two villages in Luódiàn Town (罗店镇) in Bǎoshān, Dōngnánnòng Village (东南弄村) and Luóxī Village (罗溪村).

Nánhuì, formerly Nánhuì County, became Nánhuì District in 2001 and was merged into Pǔdōng New District in August 2009. As shown in Figure 2.4, Nánhuì is located in the east of Shànghǎi, the south part of Pǔdōng New District, and along the coast of East China Sea with a land area of about 809.5km² and a population of 975,017 as of 2006. Although Nánhuì District has been merged with Pǔdōng New District when our survey was conducted in 2010, this study still uses the concept of Nánhuì community because (1) the concept of Nánhuì dialect is still being used in Chinese dialectology; (2) The identity as Nánhuì people had not changed when the survey was conducted. When asked how they identify themselves, most participants replied "Nánhuìnese". The survey was conducted in two villages of Huìnán Town of Nánhuì, Mínlè Village (民乐村) and Míngguāng Village (明光村).

2.4.2. Past tonal development in Shànghǎi

2.4.2.1. The development of citation tone

The earliest scientific descriptions of the Shànghǎi urban dialect were made by western missionaries. Two books were published in 1853; 10 years after Shànghǎi became a treaty port. They are *The Gospel of Saint John: In the Chinese Language, According to the Dialect of Shànghǎi* written by John Summer and *A Grammar of Colloquial Chinese* written by Joseph Edkins. Both books distinguished eight

---

[18] The official website of Governmental Office of Shànghǎi Chorography, http://www.shtong.gov.cn/node2/node2247/node4560/index.html, retrieved on May 20, 2014.

citation tones in SHUD and gave their corresponding text notes. Chinese dialectologists reconstructed the records of Joseph Edkins according to a 5-point scale and the results are listed in Table 2.2.

Table 2.2 Inventory of citation tones of SHUD in the 1850s

	T1	T3	T5	T7
Qián (2003)	52	44	35	5
Chén (2007)	53	44	35	55
	T2	**T4**	**T6**	**T8**
Qián (2003)	22	113	13	2
Chén (2007)	22^{3}	113	13	12

These two reconstructions are basically the same. The pitch values of T4 and T6 are very close to each other. Edkins (1853) found that some T4 syllables already merged into the T6 category. It demonstrates that SHUD was transferring from an 8-category tonal system towards a 7-category system.

The next milestone study of the Shànghǎi dialects was Chao's (1928) *Studies in the Modern Wú Dialects*. It is widely acknowledged as an authoritative account of 33 Wú dialects of the 1920s. These dialects covered SHUD and four suburban dialects: BSS, BSL, NHZ and SJC. Data of these five sites for transcribing tone was all from male adult subjects (Chao 1928: 77). It is a pity that their ages have not been recorded.

Table 2.3 Inventory of the citation tones in the Shànghǎi dialects by Chao (1928)

	SHUD	BSS and BSL	NHZ	SJC
T1	41	53	$4^{\#}1^{\#}$	$6^{b}2$
T2	13^{b}	$^{2}31$	14	13^{b7}
T3	$3^{b}2^{b2}$	$^{4}35$	$^{4}35$	432
T4 syllables with sonorant initial in literary readings	=T3; T6	=T3; =T6	N/A	=T3; =T6
T4 syllables with sonorant initial in colloquial readings	=T6	=T6	$^{3b}23$	$^{7b}67$
T4 obstruent syllables				=T6
T5	$^{2}2^{b}3$	$44^{\#}$	25	435
T6	13^{b}	$^{2}1^{\#}4$	$3^{b}5$	62
T7	$4^{\#}$	$44^{\#}$	4	$4^{\#}/5^{\#}$
T8	23	$3^{b}5$	23	71
n. of tonal categories in total	7(6)	7	8	8, 9?

Data was elicited by reading a word list. Before inventing the 5-point scale, Chao transcribed the tone values musically with the help of a sliding pitch pipe and noted them down in numerical music notations. The tone values of 33 survey points were listed in one large table and the Shànghǎi parts are extracted in Table 2.3.

Chao's records demonstrate that (1) SHUD had already completed the merger of T4 and T6 in the 1920s, while T2 was still merging with T4/6. (2) Nánhuì was the most stable area in keeping eight tonal categories, Sōngjiāng was the second most stable, and Bǎoshān showed the closest similarity with SHUD. "The methodology Chao used in his research for *Studies* are looked upon as embryonic precursors to the techniques that predominate today, methods embodied in the 'Fāngyán diàochá zìbiǎo' (方言调查字表, *Dialect Survey Word List,* 1956) and 'Fāngyán diàochá cíhuì biǎo' (方言调查词汇表, *Dialect Survey Lexicon List"* (Simmons 2006: 189). In the running speech, the text of the fable *the North Wind and the Sun* was used for reading. In short, data were elicited by word-list reading and paragraph reading. Modern Chinese dialectology (from Chao's time) was also greatly influenced by the European tradition. European dialectologists and sinologists, e.g., Georg Wenker (German, 1852-1911) and Bernhard Karlgren (Swedish, 1889-1978), introduced their techniques to investigate Chinese dialects in the late nineteenth and early twentieth century. For example, modern Chinese dialectologists select one or a few NORMs (Non-mobile Older Rural Males) as the best informants like their European colleagues. Except where special declaration is made, all the studies reviewed in the following Sections 2.4.2 and 2.4.4 used NORM's data or only published NORM's data. Some studies investigated young men as well as NORMs to compare the old and new varieties.

In the 1950s, a national survey of Chinese dialects was conducted in 1,849 survey points in China. The purpose of the survey was to assist the promotion of Pǔtōnghuà by comparing dialectal phonological systems with that of Pǔtōnghuà. Surveys in Jiāngsū Province and Shànghǎi area were carried out by a group of leading dialectologists and their results were published in *General situation of dialects in Jiāngsū Province and Shànghǎi* (1960, hereafter GSDJPS). These surveys found that SHUD had five tones (T3=T5, T2=T4=T6), Bǎoshān had six tones (T3=T5, T4=T6), Nánhuì had seven tones (T4=T6) and Sōngjiāng had eight tones (Dialect survey group of Jiāngsū Province and Shànghǎi (江苏省上海市方言调查指导组), 1960). Among them, only SHUD and Sōngjiāng were given transcriptions of tonal values. The comparison of the numbers of tonal categories in the 1920s and 1950s shows the tonal development in Shànghǎi. SHUD finished the merger of T2, T4 and T6 and further merged T3 and T5 in the Yīn register. The Bǎoshān dialects also merged T3 and T5 in addition to the merger of T4 and T6, but kept T2 separate. The Nánhuì dialect merged T4 and T6 while Sōngjiāng remained stable.

Another important study on Wú is *Studies in the Contemporary Wú Dialects* written by Qián Nǎiróng (钱乃荣). In the 1980s, Qián reinvestigated the same 33 survey points as Chao had. This makes it possible to trace the tonal development with a depth of 60 years. Meanwhile, he investigated three age groups: old, middle and young but only published data of two groups: elders (around age 70) and youth (around age 16). The old variety recorded by Qián is quite similar with Chao's 1920s new variety, regarding both the tonal categories kept for each dialect and their tonal values. One interesting finding of Qián's old variety in SHUD is the intermediate state of T4 before merging with T6; this new variant was transcribed as a mid level tone 44. Comparing the old and new varieties, the merging among T2, T4, & T6, as well as T3 & T5 continues. Changes in the NHZ were the most radical. T2 and T4 of NHZ were both merged into T6 according to Qián. However, the value of T6 also changed from 35 to 113, which is the value of T2 in the old variety. If the tone value is the determining factor in defining the tonal category, it is better to say that T4 and T6 merged into T2. This issue will be further examined in Chapter 6.

Table 2.4 Inventory of citation tones in the Shànghǎi dialects by Qián (1992: 40-47)

	Old variety				**New variety**			
	SHUD	BSS & BSL	NHZ	SJC	SHUD	BSS & BSL	NHZ	SJC
T1	52	52	52	52	52	52	52	52
T2	=T6	^{2}31	113	^{2}31	=T6	^{2}31	=T6	^{2}31
T3	=T5	^{4}35	44	44	=T5	^{4}35	44	44
T4	44/=T6	=T6	323	22	=T6	=T6	=T6	=T6
T5	334	435/44	335	335	334	=T3	335	335
T6	113	213	35	113	113	213	113	113
T7	55	55	55	55	55	55	55	55
T8	23	23	23	23	23	23	23	23
n. of tonal categories	5/6	7	8	8	5	6	6	7

As Qián followed the methodology of Chao, Qián's study can be considered as Chao's trend survey. Comparing Qián's new variety with that of Chao (data both from the young generation at that time), I found that SHUD and NHZ changed the fastest within the sixty years between 1920 and 1980, both lost 2 tones; BSL lost one tone and SJC was the most stable. Considering the distances of these suburban towns away from the city center (SHUD=0, NHZ=16km, BSL=29km, SJC=40km), it looks as if the closer the suburban town is to the city center, the more it was effected by the change. This will also be tested in our own data. Moreover, the mergers of T2, T4 and T6 of those four dialects have one similarity: the results of merges are always /113/ or /213/, a delayed rising tone. When those dialects complete the merger of T2, T4 and T6, those four dialects will have the same T2/4/6. In other words, the regional differences among lax tones in the lower register will be narrowed or even eliminated.

Almost at the same time, around 1990, dialectologists from Fudan University surveyed the dialects in the Shànghǎi urban area and its 10 counties thoroughly (Yóu, 2010). Then this survey is temporarily called as "the 1990s survey". As stated above, nine counties among them have been changed into suburban areas administratively, including Nánhuì, Bǎoshān and Sōngjiāng, which are relevant to the current study. According to Yóu (2010), the detailed and complete results of the 1990s survey have not been published in public. However, the county annals of these areas documented the main results briefly. Unlike Chao and Qián, the survey in Nánhuì was conducted in Huìnán (惠南, hereafter NHH), a town 20 kilometers further away from the city center than NHZ and the survey in Bǎoshān was conducted in the administrative center of Bǎoshān (BS center), which is 10 kilometers closer towards the city center. The results are summarized Table 2.5. The data of Bǎoshān and Nánhuì was not collected from BSL and NHZ as Chao (1928) and Qián (1992) did, so results of these two regions in the annals are not compared with them.

Table 2.5 Inventory of citation tones of SHD on the annals of the 1990s

	SHUD	BS center	NHH	SJC
T1	53	53	53	53
T2	=T6	31	=T6	31
T3	=T5	35	44	44
T4	=T6	=T6	113	22
T5	34	=T3	35	35
T6	23	113	13	13
T7	55	55	55	5
T8	12	23	23	3
n. of tonal categories	5	6	7	8

The subjects of the 1990s surveys were from the old generation, but around ten years younger than the old subjects that Qián had previously investigated. Comparing the results of the 1990s surveys and Qián (1992), it is found that the results of SHUD in annals are almost the same with the new variety of Qián (1992). The results of SJC are closer to the old variety of Qián (1992) as it had eight tonal categories. SJC was the slowest in change as it still kept eight tones among elders in the 1990s.

2.4.2.2. The development of tone sandhi

In this section, tone sandhi patterns initialed with T2, T4 and T6 (i.e., T2.X, T4.X and T6.X) which are used in SHUD, NHZ & NHH, BSL and SJC are reviewed one by one.

SHUD has been studied intensively. Regarding its sandhi tones, two representative studies are reviewed here. First, Xú, Tāng, & Qián (1981) is the most comprehensive study focusing on the sandhi patterns of Shànghǎi dialects. It documented the sandhi patterns used in the new variety around 1980. The new variety they refer to is SHUD, used by people who have merged T2, T4 and T6 in their citation system. Its results in T2.X, T4.X and T6.X are summarized at the top of Table 2.6. It can be seen that T2.X, T4.X and T6.X of SHUD have merged. It is a typical left-dominant system, extending the citation tone of its first syllable /23/ rightward, resulting in a /22.44/ sandhi pattern. Moreover, Xú, Tāng, & Qián (1981) observed a few words or phrases whose σ_2 goes with T1 to T6, can only be realized as /55.31/. In a second study, Qián (1992) also found that /22.44/ or /22.4/ is the dominant patterns used in SHUD. Qián (1992) did not mention /55.31/ but found a /22.52/ pattern used in T1.X. These differences will be studied in Chapter 6, using our data.

Table 2.6 Tone sandhi patterns initialed with T2, T4 and T6 in SHUD

<table>
<tr><th colspan="10">Xú, Tāng, & Qián (1981)</th></tr>
<tr><td colspan="2"></td><td>T1</td><td>T2</td><td>T3</td><td>T4</td><td>T5</td><td>T6</td><td>T7</td><td>T8</td></tr>
<tr><td>T2</td><td rowspan="3">23</td><td colspan="6" rowspan="3">22.44
55.31 (a few)</td><td colspan="2" rowspan="3">22.4</td></tr>
<tr><td>T4</td></tr>
<tr><td>T6</td></tr>
<tr><th colspan="10">Qián (1992: 641)</th></tr>
<tr><td colspan="2"></td><td>T1</td><td>T2</td><td>T3</td><td>T4</td><td>T5</td><td>T6</td><td>T7</td><td>T8</td></tr>
<tr><td>T2</td><td rowspan="3">113</td><td rowspan="3">22.44
22.52</td><td colspan="5" rowspan="3">22.44</td><td colspan="2" rowspan="3">22.4</td></tr>
<tr><td>T4</td></tr>
<tr><td>T6</td></tr>
</table>

To the knowledge of the author, only two studies on Nánhuì sandhi have been published: Qián (1992) and the one published in the county annals of Nánhuì (Xuē, 1992). Qián (1992) documented the sandhi patterns of Nánhuì Zhōupǔ County (NHZ) whereas the county annals documented the sandhi patterns of Huìnán County (NHH). The results of these two studies are summarized in Table 2.7.

Table 2.7 Tone sandhi patterns initialed with T2, T4 and T6 in NHZ and NHH

NHZ (Qián 1992: 640)									
		T1	T2	T3	T4	T5	T6	T7	T8
T2	$^{2}31$	22.33						22.5 22.3	
T4 T6	213	22.52		24.31	24.31 22.24	22.24		22.5	
Nánhuì county annals, NHH (Xuē, 1992: 762; 766-767)									
		T1	T2	T3	T4	T5	T6	T7	T8
T2	113	22.33/22.23				22.35		22.3	
T4	113/13	13.53		13.31		31.35		13.53	
T6	13			31.35		13.31		35.53	

If we dismiss the fine phonetic differences due to different criteria of transcribers, e.g., /22.52/ in Qián (1992) and /13.52/ in Xuē (1992), there are also four types of sandhi patterns in Table 2.7. The first pattern /22.33/ is found to be the dominant pattern used in T2.X. NHZ used /22.33/ or /22.3/ in T2.X, while NHH did not use it in T2.T5 and T2.T6. The second pattern is /22.52/ and /22.5/ in Qián (1992: 640) or /13.53/ and /13.53/ used in the county annals (Xuē, 1992: 766-767) as they are a low tone followed by a high falling tone. Two studies both found that this pattern was used by T4.T1-2, T4.T7-8 and T6.T1-2 but had different opinions for T6.T7−8. The third pattern is /24.31/ in Qián (1992: 640) or /13.31/ used in the county annals (Xuē, 1992: 766-767) as they are both a rising low followed by a falling tone. This pattern is the dominant pattern used in T4.T3-4 and T6.T3-4 in both two studies and less frequent pattern used in T4.T5-6 and T6.T5-6 found in the annals. The fourth pattern is /22.24/ in Qián (1992: 640) or /31.35/ used in the county annals (Xuē, 1992: 766-767) as they are both a low tone plus a rising tone. This pattern is found to be the dominant pattern used in T4.T5-6 and T6.T5-6 in both studies and the less frequent pattern used in T4.T3-4 and T6.T3-4.

Differences between these two studies are probably due to regional differences. But in general, dialects in NHZ and NHH have many similarities. NHH is more complicated than NHZ. Táo Huán (陶寰), a dialectologist working on SHD for more than 30 years, suggested to the author of this dissertation to investigate NHH rather than NHZ, as he found NHZ was already very similar to the urban dialect. Because

NHZ is so close to the city center (16km) and newly developed Pǔdōng (5km) geographically, its dialect has changed considerably since the development in Pǔdōng in the 1990s. So this study uses the data of NHH to represent the dialect in suburban Nánhuì.

To the knowledge of the author, only two studies on Bǎoshān sandhi patterns have been published: Qián (1992) and the one published in the county annals of Bǎoshān in 1992. Qián (1992) documented the sandhi patterns of BSL whereas the county annals documented the sandhi patterns near the administrative center of Bǎoshān. The results of these two studies are summarized in Table 2.8:

Table 2.8 Tone sandhi patterns initialed with T2, T4 and T6 in BSL

<table>
<tr><th colspan="10">Qián (1992: 639)</th></tr>
<tr><td colspan="2"></td><td>T1</td><td>T2</td><td>T3</td><td>T4</td><td>T5</td><td>T6</td><td>T7</td><td>T8</td></tr>
<tr><td>T2</td><td>113</td><td colspan="2">22.52</td><td colspan="2">24.31</td><td colspan="2">24.31
22.23</td><td colspan="2">24.3
23.4</td></tr>
<tr><td>T4</td><td>323</td><td colspan="2">22.52</td><td colspan="2">22.52
22.23
24.31</td><td>22.23
22.52</td><td>22.23
24.31
33.52</td><td colspan="2">24.3
22.4</td></tr>
<tr><td>T6</td><td>35</td><td>22.52</td><td>22.52
22.23</td><td colspan="4">24.31
22.23</td><td colspan="2">24.3</td></tr>
<tr><th colspan="10">Bǎoshān county annals (Zhū, 1992: 1005)</th></tr>
<tr><td colspan="2"></td><td>T1</td><td>T2</td><td>T3</td><td>T4</td><td>T5</td><td>T6</td><td>T7</td><td>T8</td></tr>
<tr><td>T2</td><td>31</td><td colspan="6">33.55
35.55</td><td colspan="2">35.<u>31</u></td></tr>
<tr><td>T4</td><td rowspan="2">113</td><td colspan="6" rowspan="2">35.31
33.53</td><td colspan="2" rowspan="2">35.<u>31</u></td></tr>
<tr><td>T6</td></tr>
</table>

As can be seen from Table 2.8, the results of these two studies differ widely. The results of Qián (1992) are quite complicated while the results of county annals are quite simple. Since those two studies documented sandhi patterns in two different areas of Bǎoshān, the differences are probably due to regional differences rather than age differences. The results of the Bǎoshān county annals show that the tonal system in Bǎoshān center is much simpler than BSL both in citation form and sandhi form. To trace the conservative forms in Bǎoshān, I decided to collect data from BSL. In BSL, three sandhi patterns /22.52/, /22.23/ and /24.31/ are found to be dominant. It seems that the application of these patterns depends on the MC tonal category of σ_2. Meanwhile, these patterns can also be used in the same category. For instance, all three patterns were found in T4.T3 and T4.T4. These variances shown in Qián (1992) are still open to interpretation.

Regarding SJD, two studies on Sōngjiāng sandhi patterns - Qián (1992) and the county annals published in 1991 - are compared. Those two studies both investigated Sōngjiāng center, known as “Sōngjiāng New City” (松江新城). The results of these two studies are summarized in Table 2.9:

Table 2.9 Tone sandhi patterns initialed with T2, T4 and T6 in SJC

<table>
<tr><th colspan="10">Qián (1992: 642)</th></tr>
<tr><td></td><td></td><td>T1</td><td>T2</td><td>T3</td><td>T4</td><td>T5</td><td>T6</td><td>T7</td><td>T8</td></tr>
<tr><td>T2</td><td>²31</td><td colspan="2">22.52</td><td>22.52
24.31(a few)</td><td>22.52
24.31</td><td colspan="2">24.31</td><td colspan="2">24.31</td></tr>
<tr><td>T4</td><td rowspan="2">113</td><td colspan="2">24.31</td><td>24.31
23.44</td><td>24.31
23.44
22.23</td><td colspan="2">22.23</td><td>24.31</td><td>24.31
23.4</td></tr>
<tr><td>T6</td><td>23.44
22.52</td><td colspan="2">23.44</td><td colspan="3">22.23
23.44</td><td colspan="2">24.31</td></tr>
<tr><th colspan="10">Sōngjiāng county annals (Hé, 1991: 967-968)</th></tr>
<tr><td></td><td></td><td>T1</td><td>T2</td><td>T3</td><td>T4</td><td>T5</td><td>T6</td><td>T7</td><td>T8</td></tr>
<tr><td>T2</td><td>31</td><td colspan="4">T6.T2=13.31</td><td colspan="2">T4.T6=22.13</td><td>T2.T7
=31.5</td><td>T2.T8=
31.3</td></tr>
<tr><td>T4</td><td>22</td><td colspan="6">T6.T4=13.22</td><td rowspan="2">T6.T7
=13.5</td><td rowspan="2">T6.T8=
13.3</td></tr>
<tr><td>T6</td><td>13</td><td colspan="4">T4.T4=22.22</td><td colspan="2">T4.T6=22.13</td></tr>
</table>

As can be seen from Table 2.9, the results of these two studies also differ enormously, even though they investigated the same area. The results of Qián (1992: 642) are quite complicated while the results of county annuals are relatively simple. Since those two studies documented sandhi patterns of two different areas of Bǎoshān, the differences are probably due to region rather than age. In the former, three sandhi patterns /22.52/, /22.23/ and /24.31/ are found to be the dominant patterns used in BSL. It seems that the application of these patterns depends on the MC tonal category of σ_2. Meanwhile, these patterns can also be used in the same category. For instance, all three patterns were found in T4.T3 and T4.T4. These variances shown in Qián (1992) are still open to interpretation.

Comparing the sandhi patterns in four dialects of Shànghǎi (Table 2.6 to Table 2.9), the sandhi patterns in SHUD are the simplest as only one dominant pattern /22.44/ is found; NHZ/NHH is moderately complicated as its T4.X and T6.X have merged; BSL and SJC are extremely complicated since their T2.X, T4.X and T6.X are kept distinct from each other. The similarities of the three suburbs can be summarized as follows: (1) they all have patterns /22.52/, /22.23/ (/22.24/ in NHD) and /24.31/, according to Qián (1992); (2) the σ_1 and σ_2 of the sandhi pattern both determine the sandhi pattern applied. However, based on the literature, it is difficult to formulate the rules of sandhi application in the Shànghǎi suburban dialects and therefore hamper our understanding of the sandhi variations there.

2.4.3. The speech community of Wúxī

As in the case of Shànghǎi, a few concepts about Wúxī should be clarified. The Wúxī area refers to the administrative areas of Wúxī at present and in the past, including both the urban areas and county regions within its jurisdiction. Taking Wúxī as a speech community, the Wúxī dialect (hereafter WXD) refers to all the dialect varieties spoken in the Wúxī area. Wúxī is a prefecture-level city in Jiāngsū Province and it administers nine county-level divisions, including seven districts and two county-level cities (Yíxīng 宜兴 and Jiāngyīn 江阴). The city proper refers to the four old districts formed in the 1950s: Běitáng District (北塘区), Chóng'ān District (崇安区), Náncháng District (南长区) and part of Bīnhú District (滨湖区)[19], they are collectively called urban Wúxī. Traditionally, the Wúxī urban dialect refers to the variety spoken in the core district - Chóng'ān district. However, because of the reconstruction of the core district, most local people were moved to relatively marginal urban areas, namely Běitáng, Náncháng and part of Bīnhú district. In this study Wúxī urban dialect refers to the variety spoken in the urban Wúxī. Suburban Wúxī includes the remaining three districts, with the exception of the two county-level cities, where the language varieties spoken are collectively called Wúxī suburban dialects.

Wúxī is located in the south of Jiāngsū Province, around 130 km northwest of Shànghǎi and 180 km southeast of Nánjīng in the YRDA (c.f. Figure 2.3). Wúxī has a land area of about 4788 km² and a population of 6.37 million in 2010.

Wúxī is one of the origins of Wú civilization. Dating back to the end of the Shāng dynasty (1600 BC – 1046 BC). Tàibó (泰伯), the eldest son of King Tài (古公亶父周太王) of the Clan Zhōu[20](周), left his family in present-day central China with his younger brother Zhòngyōng (仲庸) in order to give authority to his youngest brother, Jìlì (季历). They settled a thousand miles away in the wild southeast area and founded a state named "Gōuwú" (句吴) in Méilǐ (梅里, currently known as Méicūn 梅村) in present-day Wúxī, which was the origin of latter Wú Guó. Two Zhōu princes brought Wú area the advanced culture and technology of that time and helped develop Wú's agriculture and waterways; hence, the area soon flourished.

---

[19] Bīnhú District was renamed in 2000. Its old name was Jiāoqū (郊区), meaning "suburban".
[20] Clan Zhōu annexed Shāng Dynasty in 1046 BC and then became the Zhōu dynasty (1046 BC - 256 BC).

Figure 2.5 Map of Wúxī

Wúxī has the Yangtze River to its north and Tài Lake to its south. Since the 7th century, the Běijīng-Hángzhōu Grand Canal (京杭大运河), the longest canal in the world, has passed through the city. Thanks to Wúxī's climate and waterway transportation, Wúxī developed into a "land of agriculture and fisheries" (鱼米之乡). In the early 20th century, Wúxī also got a nickname "Little Shànghǎi" (小上海) because of its close geographic proximity to Shànghǎi and its rapid development and booming economy. At that time, though Wúxī was one of the cradles of national industry and commerce, it was nevertheless rather self-sufficient compared to the metropolis and treaty port Shànghǎi, having much less floating population and language contact. Therefore, WXD was quite stable during the first three-quarters of the 20th century (more details in Section 2.4.4.1).

Entering the 1980s, Wúxī experienced socio-economic redevelopment due to a change in national economic policies. It was well known for its manufacturing and textile industries, and was accordingly designated as one of the fifteen National Economic Centers in 1981[21]. Currently, Wúxī is developing its high-technology industry with the aim of becoming the "Oriental Silicon Valley"[22]. According to

---

[21] http://www.wuxi.gov.cn/wxdk/215116.shtml (in Chinese), retrieved on May 20, 2014.
[22] http://en.wuxi.gov.cn/sitePages/subPages/1300350001340157.html?sourceChannelId=23440&did=337639 (in English), retrieved on May 20, 2014.

Forbes ranking, Wúxī reached fifth place in the "Top 10 best cities for business" in Mainland China in 2012[23]. Two large industrial parks have been built in the southern suburbs and they are attracting huge investments and returning overseas talent, and recruiting workers from across the nation and the world. The total population of Wúxī was 6.37 million persons in 2010, an increase of 1.19 million from the 2000 census, including a floating population of 1.93 million (Zhuāng, 2011). The floating population increased by 1.05 million, which is 2 times more than 10 years ago and accounts for 88% of the population growth for the past ten years (ibid.).

Huàzhuāng (华庄) is a subdistrict administered by the Bīnhú District. It is adjacent to Tài Lake, 13km southwest of downtown Wúxī and 7.5km southwest of new industrial parks. It has a land area of about 59 km² and a population of 71,624 in 2009. Huàzhuāng was purely a rural area two decades ago, but nowadays has been affected dramatically by the process of urbanization.

2.4.4. Past tonal development in Wúxī

2.4.4.1. The development of citation tone

There were no previous records of the dialects spoken in the Wúxī suburbs, so this section will focus exclusively on tonal change in the urban area. Table 2.10 summarizes the inventories of the citation tones transcribed by some representative studies on WXD.

The reliable sources of the Wúxī data include Chao (1928), *General situation of dialects in Jiāngsū Province and Shànghǎi* (1960, hereafter GSDJPS), Qián (1992) and Wúxī Annals. In these surveys, the methods of collecting data were the same as those of the Shànghǎi surveys (see Section 2.4.2.1). Two more sources of Wúxī tonal data are Chan & Ren (1989) and Xú (2007). Chan & Ren (1989) wrote the first English-language article about Wúxī tones. They used acoustic approaches to transcribe the pitch values of monosyllabic and bisyllabic words. Recordings were made at the University of California, Los Angeles (UCLA) in 1982 and 1986, for two native speakers: one subject was a visiting scholar at UCLA and the other one was the second author. The subjects' ages were not reported. Ren was a graduate student at that time; hence it may be assumed that he represented the young or

[23] http://www.china.org.cn/top10/2012-12/13/content_27404677_6.htm (in English), retrieved on May 20, 2014.

middle-aged native speakers of WXD at that time. Chan & Ren (1989) found that the inventory of Wúxī citation tones had seven tones with T2 and T6 merged, matching Qián's new variety.

Another source was Xú (2007). The subjects in this study were six native speakers (average age 57.7): three males and three females. Data was also elicited by word-list reading. Xú first transcribed T4 as a low-rising tone with tonal values similar to T6, indicating that T4 was migrating towards T6. Chan & Ren (1989) had a quite conservative T4, close to Chao's transcription from the 1920s. So, it can be deduced that the peaking contour of T4 was the old variant, the mid level shape /33/ was intermediate, and the low rising variant /13/ recorded by Xú (2007) was the most advanced, or recent.

Table 2.10 Inventories of the citation tones in WXDs in chronological time

<table>
<tr><th></th><th>Chao (1928)</th><th>GSDJPS (1960)</th><th>Chan & Ren (1989)</th><th>Qián (1992) old variety</th><th>Qián (1992) new variety</th><th>Wúxī annals (1995)</th><th>Xú (2007)</th></tr>
<tr><td>T1</td><td>$\underline{53}^{b}$</td><td>55</td><td>52</td><td>$^{5}44$</td><td>$^{5}44$</td><td>53</td><td>44/53</td></tr>
<tr><td>T2</td><td>$1\underline{3}^{b}$</td><td>14</td><td>=T6</td><td>14</td><td>=T6</td><td>=T6</td><td>=T6</td></tr>
<tr><td>T3</td><td>$\underline{\underline{3}^{b}2^{b}\underline{3}^{b}}$</td><td>324</td><td>313</td><td>323</td><td>323</td><td>323</td><td>323/223</td></tr>
<tr><td>T4 syllables with sonorant initial in literary readings</td><td>=T3;
=T1</td><td rowspan="3">33</td><td rowspan="3">131</td><td rowspan="3">33</td><td rowspan="3">213</td><td rowspan="3">232</td><td rowspan="3">13</td></tr>
<tr><td>T4 syllables with sonorant initial in colloquial readings</td><td rowspan="2">$\underline{\underline{2}^{b}3^{b}\underline{2}}$</td></tr>
<tr><td>T4 obstruent syllables</td></tr>
<tr><td>T5</td><td>$\widehat{3^{b}\underline{4}}$</td><td>35</td><td>34</td><td>34</td><td>34</td><td>35</td><td>$^{3}34$</td></tr>
<tr><td>T6</td><td>$\underline{\underline{2}\underset{\cdot}{7}3}^{b}$</td><td>213</td><td>213</td><td>213</td><td>213</td><td>13</td><td>113/13</td></tr>
<tr><td>T7</td><td>$\underline{4}$</td><td>5</td><td>$\underline{53}$</td><td>$\underline{55}$</td><td>$\underline{55}$</td><td>5</td><td>5</td></tr>
<tr><td>T8</td><td>$\underline{\underline{1^{\#}4}}$</td><td>2</td><td>$\underline{13}$</td><td>$\underline{23}$</td><td>$\underline{23}$</td><td>$2^{3}$</td><td>$\underline{13}$</td></tr>
<tr><td>n. of tonal categories</td><td>8</td><td>8</td><td>7</td><td>8</td><td>7</td><td>7</td><td>7/6</td></tr>
</table>

From Table 2.10, the development of the Wúxī citation tones can be traced clearly. Like the other northern Wú dialects, tones T2, T4 and T6 in the Wúxī dialect were undergoing changes. T2 and T6 merged first while T4 showed a tonal pattern highly similar to T2/6 in Xú (2007). With respect to T4, Chao separated syllables with a sonorant initial in literary readings, syllables with sonorant initials in colloquial readings and obstruent syllables, and found that these three sets did not have the same tone values. However, this separation was not reported in the following studies and hence no special case of T4 was recorded in these studies.

The current study does not compare the literary and colloquial reading styles, but it separates sonorant and obstruent initials. For the transcription of T4, a considerable amount of variation occurs in earlier studies. T4 was recorded as a peaking tone /$\underline{2}^{b}3^{b}\underline{2}$/ in Chao (1928), Chan & Ren (1989) and Wúxī annals (1995), as a level tone /33/ in GSDJPS (1960) and the old variety in Qián (1992), as a low dipping tone /213/ in the new variety of Qián (1992), and as a low rising tone /13/ or /113/ in Xú (2007). T4 shares its transcription with T2 and T6 in the new variety of Qián (1992) and Xú (2007), indicating the merger of T4 with T2 and T6. Hence, /213/ and /113/ can be identified as the advanced variant of T4. There are two different transcriptions of the conservative T4: the peaking tones transcribed by Chao (1928), Chan & Ren (1989) and Wúxī annals (1995) and the level one transcribed by Qián (1992) and GSDJPS (1960). This discrepancy is investigated in our data and discussed in Chapter 5.

2.4.4.2. The development of tone sandhi

Chao (1928) first found that the sandhi pattern for T1 in WXD was quite distinct, sounding like the isolated T3. This feature was commonly considered to be "the Wúxī accent" (Wúxī Qiāng, 无锡腔). Chén (1989) further developed this notion and summarized Wúxī sandhi patterns as two circles:

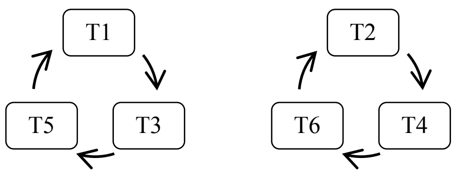


Figure 2.6 Tone sandhi circles in WXD (Chén 1989)

The left circle in Figure 2.6 illustrates that the T1 sandhi pattern sounds like the melody of isolated T3; the T3 sandhi pattern sounds like the melody of isolated T5; and the T5 sandhi pattern sounds like the melody of isolated T1. Similarly, the right circle shows that the T2 sandhi pattern sounds like the melody of isolated T4; the T4 sandhi pattern sounds like the melody of isolated T6; and the T6 sandhi pattern sounds like the melody of isolated T2. However, Chén (1989) did not give the tone value of each citation tone and sandhi pattern, while Chan & Ren (1989), Qián (1992), Cáo (2003) and Xú (2007) did address this issue. As this dissertation will focus on T2, T4 and T6 sandhi patterns, only the values of these three tone patterns as transcribed by Chan & Ren (1989), Qián (1992), Cáo (2003) and Xú (2007) are included in Table 2.11.

Table 2.11 summarizes data from four studies. For each study, the leftmost column lists the tone of σ_1 and its citation tone; the topmost row indicates the tone of σ_2. Chan & Ren (1989) used a different transcription system from the other three studies due to its research purpose. Chan & Ren (1989) sought to interpret the Wúxī sandhi in order to establish its abstract phonological rules while the other three studies had mostly descriptive purposes. On the basis of their data and work, Chen (2000) further distinguishes two types of constructions of Wúxī sandhi, which is listed following Table 2.11.

Type I	1.	reduplicated verbs
	2.	verb + resultative or directional complements
	3.	expressions with "number + classifier"
	4.	reduplicated nouns (in child-directed speech only)
Type II	1.	compounds (including regular reduplicated nouns)
	2.	phrases

(Chen, 2000: 321)

Type I sandhi is typically "left-dominant", like SHUD. It is used in reduplicated verbs, verbs followed by resultative or directional complements, expressions with "number + classifier" and reduplicated nouns. It undergoes tone deletion and spread: tone deletion eliminates all but the leftmost tone; then the citation form of the leftmost tone is extended into the entire sandhi domain. According to Chan & Ren (1989), T2 in citation form is "LHL", so T2.X in Type I also uses "LHL"; T4 in citation form is "HHL", so T4.X in Type I uses "HHL"; T6 in citation form is "LLH", so T6.X in Type I uses "LLH". However, it should be noted that, in Chan & Ren (1989), T2 and T6 had the same phonetic transcription /213/ but different phonological presentations: LHL and LLH.

Table 2.11 Tone sandhi patterns initialed by T2, T4 and T6

<table>
<tr><th colspan="10">Chan & Ren (1989)</th></tr>
<tr><td colspan="2">Type I</td><td>T1</td><td>T2</td><td>T3</td><td>T4</td><td>T5</td><td>T6</td><td>T7</td><td>T8</td></tr>
<tr><td>T2</td><td>213
LHL</td><td colspan="8">LHL</td></tr>
<tr><td>T4</td><td>131
HHL</td><td colspan="8">HHL</td></tr>
<tr><td>T6</td><td>213
LLH</td><td colspan="8">LLH</td></tr>
<tr><td colspan="2">Type II</td><td>T1</td><td>T2</td><td>T3</td><td>T4</td><td>T5</td><td>T6</td><td>T7</td><td>T8</td></tr>
<tr><td>T2</td><td>213
LHL</td><td colspan="8">HHL</td></tr>
<tr><td>T4</td><td>131
HHL</td><td colspan="2">LHL</td><td colspan="4">LLH</td><td colspan="2">LHL</td></tr>
<tr><td>T6</td><td>213
LLH</td><td colspan="8">LHL</td></tr>
<tr><th colspan="10">Qián 1992</th></tr>
<tr><td colspan="2"></td><td>T1</td><td>T2</td><td>T3</td><td>T4</td><td>T5</td><td>T6</td><td>T7</td><td>T8</td></tr>
<tr><td>T2</td><td>14</td><td colspan="6">24.31</td><td colspan="2">24.2</td></tr>
<tr><td>T4</td><td>33</td><td colspan="2">22.55</td><td>22.55
21.23</td><td>22.55
21.23
55.31</td><td>21.23
55.31</td><td>22.55
21.23</td><td>22.5
55.3</td><td>22.5</td></tr>
<tr><td>T6</td><td>213</td><td colspan="2">22.55
21.23 (a few)</td><td colspan="4">22.55
21.23</td><td colspan="2">22.5</td></tr>
<tr><th colspan="10">Cáo 2003</th></tr>
<tr><td colspan="2"></td><td>T1</td><td>T2</td><td>T3</td><td>T4</td><td>T5</td><td>T6</td><td>T7</td><td>T8</td></tr>
<tr><td rowspan="2">T2</td><td rowspan="2">223</td><td colspan="6">A.13.21 B.22.44</td><td colspan="2" rowspan="2">A.13.21
B.22.4</td></tr>
<tr><td colspan="5"></td><td>C.22.24</td></tr>
<tr><td>T4</td><td>13</td><td colspan="6">A.22.24 B.13.21 C.22.44</td><td colspan="2">A.22.4
B.13.21</td></tr>
<tr><td rowspan="2">T6</td><td rowspan="2">223</td><td colspan="6">A.22.44 B.22.24</td><td colspan="2" rowspan="2">A.22.4
B.22.24</td></tr>
<tr><td colspan="4"></td><td colspan="2">C.13.21</td></tr>
<tr><th colspan="10">Xú 2007</th></tr>
<tr><td colspan="2"></td><td>T1</td><td>T2</td><td>T3</td><td>T4</td><td>T5</td><td>T6</td><td>T7</td><td>T8</td></tr>
<tr><td rowspan="2">T2</td><td rowspan="2">113/ 13</td><td colspan="6">[a] 35.31 [c] 33.55</td><td colspan="2" rowspan="2">[a] 35.31
[c] 33.5</td></tr>
<tr><td colspan="3"></td><td colspan="3">[b] 33/43.34</td></tr>
<tr><td rowspan="3">T4</td><td rowspan="3">13</td><td colspan="2">[a] 33.55</td><td colspan="4">[a] 33.34</td><td colspan="2" rowspan="3">[a] 33.5
[c] 35.41</td></tr>
<tr><td colspan="2">[b] 33.34</td><td colspan="4">[b] 33.55</td></tr>
<tr><td colspan="6">[c] 35.41</td></tr>
<tr><td rowspan="2">T6</td><td rowspan="2">113/ 13</td><td colspan="6">[a] 33.55 [c] 33.34</td><td colspan="2">[a] 33.5
[c] 33.34</td></tr>
<tr><td colspan="4"></td><td colspan="2">[b] 35.31</td><td></td><td>[b] 35.31</td></tr>
</table>

Type II construction is used by common compounds and phrases. It undergoes an extra step of pattern substitution before tone deletion and spread. HHL and LHL are the substitution forms for T2.X and T6.X, separately. The substitution for T4.X is "right-dominant" as its substitution rule is determined by σ_2. If σ_2 is T3, T4, T5 or

T6, "LLH" replaces "HHL" and further spread to the entire sandhi domain; if σ_2 is T1, T2, T7 or T8, "LHL" substitutes "HHL". Chan & Ren (1989) argued that pattern substitution in Wúxī originally involved last syllable dominance, attributing the unusual tone sandhi behavior in modern Wúxī to a shift in dominance from the last to the first syllable. Chan & Ren's data partly supports Chén's (1989) conclusion that the T2 sandhi pattern sounds like the melody of isolated T4, and that T6 sandhi pattern sounds like the melody of isolated T2. However, in Chan & Ren's (1989) data, there are two sandhi patterns of T4.X, determined by σ_2. If σ_2 is T3, T4, T5 or T6, the sandhi pattern sounds like the melody of citation T6. If σ_2 is T1, T2, T7 or T8, the sandhi pattern sounds like the melody of citation T2. This pattern split in T4.X will be further investigated in our dataset to see whether it is involved in any ongoing changes.

Turning now to the other three studies in Table 2.11, Cáo (2003) identified three sandhi patterns in all of T2.X, T4.X and T6.X and marked these by "A", "B" and "C" according to their frequency of use, where "A" has the highest frequency. Cáo (2003) indicated that most type B words she tested were special grammatical structures like reduplicated verbs and expressions with "number + classifier". They overlapped with Chen's (2000) Type I. Type I words were also discussed by Cáo (2003) but not by Qián (1992). Xú (2007) also found different types of sandhi patterns in each tonal category. Type [a] and [b] are variants in phonological context, yet type [c] equals Chen's (2000) Type I, being conditioned by morphosyntactic factors. Xú's (2007) type [a] and [b] are variants of Chen's (2000) Type II words. Xú's (2007) type [a] corresponds to Cáo's (2003) type A; Xú's (2007) type [b] corresponds to Cáo's (2003) type C.

Table 2.12 Conversion table of transcriptions from different studies

Pattern		Transcriptions per study
1	Qián 1992	/24.31/ and /24.2/
	Cáo 2003	/13.21/ and /13.21/
	Xú 2007	/35.31/, /35.31/, /35.41/ and /35.41/
2	Qián 1992	/22.55/ and /22.5/
	Cáo 2003	/22.44/ and /22.4/
	Xú 2007	/33.55/ and /33.5/
3	Qián 1992	/21.23/
	Cáo 2003	/22.24/ and /22.24/
	Xú 2007	/33.34/, /43.34/ and /33.34/
4	Qián 1992	/55.31/ and /55.3/

The Wúxī tone sandhi situation looks quite disorderly based on these phonetic transcriptions. Yet dismissing the fine phonetic differences due to differences between transcribers, Table 2.12 only includes four types of sandhi patterns regardless of the citation value of $σ_1$. These are listed in Table 2.12 with their transcriptions in each study.

Based on Table 2.4 and Table 2.12, it can be easily established that only the patterns /55.31/ and /55.3/ were previously observed by Qián (1992) for T4.X. This might be caused by an occasional borrowing of T5.X. /55.31/ was primarily used by T5.X, which is the result of $σ_1$'s pattern extension. The first T5 syllable has a voiceless initial, which naturally induces a high tone. If a T4 morpheme, either in isolation or as the $σ_1$ of a sandhi form, is occasionally pronounced with a voiceless initial, its voiceless initial may also induce a high tone. The remaining three patterns /22.24/, /22.44/ and /24.31/ are found among T2.X, T4.X and T6.X. T2.X is predominantly /24.31/ while T6.X is predominantly /22.44/. The situation for T4.X is somewhat complex: Qián (1992) found that /22.24/ occurred when $σ_2$ is T3, T4, T5 and T6, whereas both Xú (2007) and Cáo (2003) found that a few words with tone combinations T4.T1-2 also used /22.24/, albeit less frequently than /22.44/. If Qián's (1992) records are trustable, the differences between Qián's study and the other two indicate a diffusion of the /22.24/ pattern among T4.X. There might be ongoing changes involved in the diffusion process of /22.24/. For this reason, in the design of the current study (Chapter 3), T4.T1-2 words and phrases that use /22.24/ according to Xú (2007) and Cáo (2003) will be of particular interest.

This study aims at investigating the tonal variation and changes in Wúxī and Shànghǎi against the background of the rapid urbanization process. The current chapter has reviewed tonal developments that have occurred in Wúxī and Shànghǎi based on earlier studies. Summarizing, it was found that T4 of WXD, T2 of SJC and BSL, and T6 of NHH have not completed the merger with the other lax tones of the Yáng register in the citation tones; these are more likely to be involved in ongoing changes. Regarding tone sandhi, it was found that besides SHUD, the other five regions all have relatively complicated sandhi patterns: WXD mainly uses /22.24/, /22.44/, and /24.31/ patterns for bisyllables initialed by T2, T4 and T6, whereas NHH, BSL and SJC mainly use /22.24/ (/22.23/), /22.44/ and /24.31/ patterns. However, it is still unknown whether the choice of those tone patterns is involved in any change in progress. Meanwhile, the acoustic realization of these tone patterns also might involve change in progress. Linking and comparing previous studies provide us some clues where the sociolinguistic variables of tone might exist in Wúxī and Shànghǎi. Based on these clues, the first goal of this study is to scan our own data to assess the realization of these sociolinguistic variables and fully describe them.

The descriptive results can further serve the second theoretical goal: to establish how internal, external and extra-linguistic factors interact in tonal variation and change. To fulfill this goal, this study investigates linguistic variables that are both linguistically and socially constrained. The variations and changes in the lax tone of the Yáng register meet this requirement because these are not only constrained by social factors due to urbanization but also are internally motivated due to their links with the Zhuó Shǎng Guī Qù, one of the major phonological changes during the MC era.

Chapter 3. Methodology

This chapter outlines the methods employed in fieldwork and data analysis. It first describes the research design with respect to data collection, including the criteria for selecting participants in Section 3.1, the procedures for recruiting participants in Section 3.2, recording procedures in Section 3.3 and speech material for data elicitation in Section 3.4. The primary task of variationist research is to correlate linguistic variation as the dependent variable with independent variables (Chambers 2003: 17). Given the design of data collection, Section 3.1 ends with summaries of the social predictors examined in this study, and Section 3.4 ends with summaries of the linguistic and stylistic predictors. Based on the different types of dependent and independent variables, Section 3.5 justifies the use of regression models used for identifying the features associated with the variation and presents an overview the modeling procedures.

3.1. Participants

Variationists firmly believe that language is an "orderly heterogeneity" (Weinreich, et al., 1968) rather than a homogeneous ideal, and variationists differ from theoretical linguists in their goal of understanding "the impact of the interaction of language, culture and society on the structures and processes of traditional linguistics." [24] Beside purely linguistic considerations, social factors are also essential to the investigation and explanation of language variation and change (LVC).

Within LVC research, there are three types of investigation with different research aims and methodological priorities, or "three waves" according to Eckert (2000; 2012). The first wave is the traditional Labovian study; namely, research that studies the relationship between linguistic variability and broad social factors such as age, sex, socioeconomic status, etc. The second wave of variation study employs ethnographic methods to examine the relation between linguistic variation and more narrowly-defined categories, which can be defined by participants' practices, e.g., gang membership (Fought, 1999), nerd girls (Bucholtz, 1999), and social networks

---

[24] Definition of LVC from the Journal *Language Variation and Change,* http://journals.cambridge.org/action/displayJournal?jid=LVC, retrieved on May 20, 2014.

(Milroy, 1981). The third wave focuses on the social meaning of variables. Its research targets not only linguistic variables (e.g. changes in progress) but any linguistic material that serves a social/stylistic purpose (Eckert, 2008). For instance, Podesva (2007), a typical third wave study, found that a medical student used falsetto phonation across situations to convey an expressive meaning, which can be used to construct a "diva" persona and perhaps a gay identity.

This study, as an initial attempt to understand tonal variation and change in the Wú dialects, aims to offer a big-picture analysis of this change, and its social spread in the context of rapid urbanization. According to this aim, the methodology of first-wave studies, i.e., using major demographic categories as the basic social units, is more efficient for collecting data and providing an overview of sound change. Using major demographic categories also enables the data collection over multiple localities (Wúxī vs Shànghǎi, urban vs suburban, different suburbs in Shànghǎi). The results of this study can serve as a basis for further research, like investigating one specific social network or local practice for the correlation of one tonal variable (second-wave study) or how one individual deploys the tonal variants to convey fine-grained social meaning and identities (third-wave study). In sum, this study will follow the work of first-wave variationists. The remainder of this section will discuss the factors considered in participant selection: age, sex, space, socioeconomic status and ethnicity.

3.1.1. Age

Unlike historical linguistics, variationist sociolinguistics not only deals with systematic and inherent variation in language in the past (diachronic), but also in the present (synchronic) (Tagliamonte, 2011). Over the past half century, most variationist studies have been synchronic, "making use of the present to explain the past" (Labov, 1975). This field has accumulated a large array of results using the "synchronic approach". The two most commonly used methods are: (1) *real-time construct*: revisiting the same community at different periods, that is, conducting a longitudinal study; (2) *apparent-time construct*: indicating the linguistic changes among different generation of a population at one point in time (Bailey, 2004). Real-time evidence is difficult to acquire: existing data is difficult to compare over time and re-surveys are notoriously time-consuming and complex. Hence, analyzing linguistic features in apparent-time is the most widely used analytical tool for variationist research. In the apparent-time construct, generational differences are used to reflect the diachronic changes that happened in the (recent) past of a language.

In this study, generational change (linguistic age) will be compared between an old generation (born before 1950, aged 60 and above at the time of data collection in 2010) and a young generation (born between 1987 and 1992, aged 18 to 23). The old generation represents people who were born and raised in the non-urbanized China; while the young generation is the first generation in China to grow up during fast-paced urbanization, starting in the 1980s (C. Lǐ 2005). In this study, age group, or more precisely, age cohort (old vs young), is used to trace the tonal variation and change under the impact of urbanization.

People aged 60 or older were retired at the moment of data collection. In comparison with people in their 50s they have more time to participate in this type of research. Young adults between the ages of 18 and 23 were also easier to recruit than those under 18, as they had graduated from high school and did not have the pressure of securing a place in college.

There is another important difference in the level of education of the young and old generation as discussed in Chapter 1. On the whole, young people have received a higher education than the old group for historical reasons. In the analyses of tonal variation in Chapters 5 and 6, the level of education of the participants is taken into account.

3.1.2. Space

Section 1.2 described Goal I of this study: to examine the ongoing tonal changes in the Wú dialects in the context of urbanization. Having such a goal determines this study's nature as a geographically informed variation analysis. Then space is an indispensable variable in the research design. It is the core external variable in traditional dialectology. However, in variationist research, space has received less attention as a social category than socioeconomic status, gender or style (Britain, 2004). The existing variationist research mainly discusses space as a stage in the diffusion of linguistic innovations (Bailey, et al., 1993). Therefore, space can act as a surrogate for apparent time in defining and analyzing linguistic variables (Horvath & Horvath, 2001). This study will therefore take account of the role of space in reflecting and shaping tonal variation. As mentioned in Chapter 2, two cities, Shànghǎi and Wúxī, are selected. Within each city, both the urban and suburban sites will be further separated.

The split of urban and suburban in this study not only represents a geographic division, but also a social division. As discussed in Chapter 2, the urban and rural communities have distinct identities due to the *hukou* system. In this study, all the speakers recruited from the suburban areas have *rural hukou*, while all the urban speakers hold *urban hukou*. Therefore, the identity contrast between urban and suburban people is, to some extent, the same as the urban-rural contrast. For the old generation, it is almost impossible to change their *hukou* identity. They can migrate but cannot integrate into the city. Nowadays, however, the young generation in the rural/suburban areas has a much higher chance of getting an *urban hukou* due to recent policy changes. Obtaining an *urban hukou* is like obtaining a new social status for them. Therefore, among the young generation, the urban-suburban contrast is more or less like the stratification of socioeconomic status. One of the issues this study will examine is whether there are some linguistic innovations marking this group of "social climbers".

3.1.3. Sex

Sex, as a common social factor, has revealed its constraints on linguistic change by hundreds of variationist studies through many years. Labov formulated the gender paradox that "women conform more closely than men to sociolinguistic norms that are overtly prescribed, but conform less than men when they are not" (Labov, 2001). It seems that the constraining effects of sex are generally omnipresent. However, a number of investigations into Chinese speech communities have yielded different results (Rau 2001; Huang 2000; Chen 2004; Xu 1992; H. Chan 1984; among others; cited in Xu to appear). In those studies, although some other social constraints were uncovered, the sex effect did not emerge. However, sex differences were found in a recent variationist study of the vernacular at Lián Island (连岛, the Liányúngǎng dialect, Mandarin) (Lǐ, 2012). Moreover, two generations had two different linguistic patterns constrained by sex. Young women on Lián Island were found to use a higher proportion of standard/prestigious variants than young men, while women in the older generation favored the nonstandard forms more than men. As the author explained, the change of sex patterns between two generations was because Lián Island was changing from a closed rural community into an urbanizing community. Therefore, along with the change of Chinese society, the constraining effects of sex may change, so this study will also explore whether sex is a factor embedded in the social context of language, both in the urbanized area (Shànghǎi urban and Wúxī urban) and suburban areas which are undergoing urbanization. For this reason, a balanced selection of male and female participants was made in each locality and within each age group.

3.1.4. Socioeconomic status

Socioeconomic status is a concept determined by economic factors. China, before the 1980s, had a command economy or planned economy; almost all employees worked for the state-owned enterprises during that period. The social stratification in China was quite distinct from that in the countries with a market-based economy. The basic social division of China's urban population at that time was "worker" (gōngrén, 工人) and "cadre" (gànbù, 干部) (D. Xu, 1992), which is comparable to "blue collar" and "white collar". With the transformation from a planned economy to market-based (or, more accurately, mixed) economy, the social stratification has changed dramatically. Q. Zhang (2010) found an influence of this new social stratification on linguistic change. The common social stratification of upper, middle, and working classes are emerging in contemporary China (Li, 2009). However, this dissertation only investigates working class participants. The reasons are: (1) two thirds of our participants, whether old or young, came from the suburbs, which were rural areas two decades ago (Zhāng & Dù, 2001). Nowadays, these suburbs are communities predominantly composed of working class people. (2) Among older people in the urban areas, the social stratification is not obvious. As for age, the non-working class seems to become younger over time (C. Lǐ, 2005). The class stratification of elders after retirement has not changed much in the past decades and the working class predominance has been maintained. (3) Most of the young urban speakers were college students. There is class stratification in this environment, determined by their parents' socioeconomic status, so, in order to ensure the comparability with the rest of the participants, young urban participants were all recruited from working class families.

3.1.5. Ethnicity

Han Chinese is the dominant ethnic group of the Yangtze River Delta Area (YRDA). According to the 2010 National Population Census, 98.8% of Shànghǎi's residents and 99.2% of Wúxī's are of the Han Chinese ethnicity[25]. Therefore, this study will only investigate Han Chinese participants.

3.1.6. Social factors

Considering the five aspects above, 120 participants were selected, 80 from Shànghǎi and 40 from Wúxī, 20 per site. They were all Han Chinese from the working class. The participants were further designated by age (old vs young) and sex (male vs female). In addition, all participants met the following criteria: (1) Participants and their parents were born and raised locally. (2) They are native speakers of the local dialect. (3) They use the local dialect as their only home language. (4) They are literate and received at least a primary education.

Table 3.1 presents the design of the speaker selection. Numbers in the columns "female" and "male" are number of speakers for each group. The abbreviations before the numbers are the group code, using the combination of bold letters in Table 3.1 to indicate the city, area, age group and sex of each cell from left to right, in which the geographic information is in uppercase. For example, "**SUof**" means **S**hànghǎi **U**rban **o**ld **f**emales. These abbreviations will be used in the following chapters to indicate the speaker group.

Table 3.1 Design of speaker selection, split by city, area, age and sex

city	area	age group	age		**f**emale		**m**ale	
			M	SD	abbr.	n	abbr.	n
Shànghǎi	**U**rban	**o**ld	67.10	7.98	SUof	5	SUom	5
		young	18.80	1.48	SUyf	5	SUym	5
	Sōngjiāng	**o**ld	64.00	4.19	SSof	5	SSom	5
		young	18.70	1.49	SSyf	5	SSym	5
	Bǎoshān	**o**ld	67.90	7.03	SBof	5	SBom	5
		young	19.60	1.58	SByf	5	SBym	5
	Nánhuì	**o**ld	68.90	6.57	SNof	5	SNom	5
		young	20.40	0.84	SNyf	5	SNym	5
Wúxī	**U**rban	**o**ld	66.80	4.39	WUof	5	WUom	5
		young	18.40	1.90	WUyf	5	WUym	5
	Huàzhuāng	**o**ld	65.00	4.06	WHof	5	WHom	5
		young	19.60	1.65	WHyf	5	WHym	5

For each group the mean age (M) and standard deviation (SD) is given. Univariate

[25] Shànghǎi data is available from http://jfdaily.eastday.com/j/20110503/u1a878723.html; Wúxī data is available from http://tjj.zhenjiang.gov.cn/tjzl/tjfx/201212/t20121226_842388.htm, both retrieved on May 20, 2014.

ANOVAs were conducted with age as dependent variable and REGION and SEX as independent variables, for the two age groups separately. In the old group, there were no significant effects of SEX ($F(1, 48)=3.66$, $p=0.062$), REGION ($F(5, 48)=0.96$, $p=0.451$) and SEX*REGION ($F(5, 48)=0.60$, $p=0.697$). In the young group, there were no significant effects either: SEX ($F(1, 48)=0.01$, $p=0.934$), REGION ($F(5, 48)=2.35$, $p=0.055$) and SEX*REGION ($F(5, 48)=0.95$, $p=0.458$). It can be concluded that the old groups and young groups across areas and sex do not differ and that the areal and sex differences are not biased by age differences.

Based on the above aspects of the sociological information of these speakers, five independent variables were set up for the testing of social conditioning effects on the tonal variation. In what follows, predictor names are presented in uppercase, and levels of categorical predictors are in italics. In our analyses, the social predictors include:

(1) **AGE** (*old* and *young)*
(2) **SEX** (*male* and *female*)
(3) **REGION** (*Shànghǎi urban, Sōngjiāng, Bǎoshān, Nánhuì, Wúxī urban* and *Huàzhuāng*)
(4) **EDU** (*primary school*, *middle school*, and *college*)
(5) **PTH** (*can speak PTH* and *cannot speak PTH*)

It is noteworthy here that above five predictors are all classified as 'social factor' because they are based on the speaker's demographic information but REGION and PTH among them are external linguistic factors as explained in Chapter 1. AGE, SEX and EDU are extra-lingusitic factors. Moreover, those five predictors are all nominal variables. The detailed description of the first five predictors can be found in Section 3.1.1 to 3.1.5. The last one "PTH" is the information regarding each individual's PTH proficiency. There was a question in the interview asking the participant whether they can speak PTH or not. Their self-reported answers "*can*" or "*cannot*" are considered to be two levels of this variable.

3.2. Participant recruitment

Participants from Wúxī were recruited by the author's personal connections, such as neighbors, friends of family, former teachers and their social connections. Those Wúxī urban speakers were all from Běitáng District, Chóng'ān District and Náncháng District, three districts located in the center of Wúxī City. Those from Huàzhuāng were recruited by a family friend of the author, who was born in 1947 in

Huàzhuāng and lived all his life there.

Participants in urban Shànghǎi were recruited with the help of the author's friends and linguists of Fudan University. Participants from the Shànghǎi suburbs were found through the local government. The author tried to look for appropriate participants in Shànghǎi suburbs herself, as she did not have similar connection as in the urban area. After several attempts, she found this approach ineffective, as suburban residents did not trust a stranger and were not prepared to participate in an interview for more than one hour, although they were to be remunerated. Generally speaking, it is not easy to do such a corpus-based community investigation in China. The latest data shows that the overall level of social trust has declined. Less than 30% people would trust strangers, based on the result of an investigation of over 1900 participants in seven cities[26]. The top-down method of gaining assistance through local government was proven in this study to be the most effective way of recruiting qualified participants in an unfamiliar community.

Specifically, the top-down recruitment was executed through the following steps. First, the author selected two representative villages of each area, and then went to the superior government of those villages to find the officers in charge of language and culture. The author explained the research goals and operation steps to them and asked for their assistance. Some of them were kind and glad to support. If they refused the requests, the author then tried another department of the government or even went directly to find the chief executive of the town or district until finding someone willing to help. The officers of the town government or district government then discussed with the head of target villages and asked for their assistance in searching appropriate participants. As the head of the village can access the basic information (original family *hukou*, age, sex, address, home phone number, etc.) of villagers and also knows the villagers quite well, it is easy for them to find qualified speakers and make appointments. Meanwhile, the selected villagers thought the interview was reliable because it was organized by their local government, so they were willing to collaborate.

All participants were paid 50 RMB, slightly more than half day's income of a Shànghǎi urban participant[27].

---

[26] News "Chinese distrust strangers, lack shared values", released on February 18, 2013. The link is english.peopledaily.com.cn/90882/8132101.html, retrieved on May 20, 2014.

[27] In 2010, the municipality registered 31836 RMB in per capita net income of urban households in

3.3. Recording

The participants were interviewed in their neighborhood and their speech was recorded in a quiet room, and the doors and windows were closed during the recording to reduce background noise. The survey was conducted in the summer, June to August of 2010. Due to the hot weather, rooms equipped with noiseless air conditioning were selected as recording rooms, creating a comfortable interview situation without disturbing background noise.

The stimuli were presented one by one in simplified Chinese characters on a 12.1-inch laptop screen. The font size was 175 point, easier for the elders to read when they sit 0.8 - 1.5 meter away from the screen. The order of presentation was randomized. Five sets of stimuli with different randomized orders were presented to the five participants from the same cell in Table 3.1. The display laptop was IBM X200, chosen for its portability. Its weight was less than 2kg with battery. It was found that, for presenting stimuli, a laptop with battery was more reliable than a desktop because the battery can avoid any interruptions caused by a power cut. The power-cuts happened several times when doing fieldwork in the Shànghǎi suburbs. The hot weather (over 40 degrees Celsius) caused a spike in electricity consumption which resulted in the power outage of that area.

The reading passage was printed on one side of an A4 size paper, avoiding turning the page in the course of reading. Two versions with different font sizes were prepared for the elders and the youth, 22 point print for the elders and 14 point for the youth. The papers were covered with transparent plastic sheets to reduce the noise of handling the paper.

The participants were asked to read the stimuli in their dialect. They were able to control the speed of the stimulus presentation by clicking the computer mouse. For elders who had difficulties with controlling the mouse, the interviewer clicked the mouse for them according to pre-designated hints, like moving fingers or touching the table. Speech was recorded with a portable TASCAM DR-100 solid state recorder and two AKG C420 headset microphones. The recordings were sampled at 48kHz (24 bits).

Shànghǎi, that equals 87 RMB a day (http://china.org.cn/top10/2012-03/16/content_24912437_9.htm, retrieved on 5 June, 2014).

The author of the present study was the interviewer in Wúxī, both the urban and suburban area. Three fieldwork assistants who were all native speakers of Shànghǎi dialect were recruited from Fudan University and Shànghǎi University. They were all graduate students of linguistics and had experience in doing dialectological fieldwork.

3.4. Speech material

3.4.1. Research design

The design of data elicitation combined the methods of Chinese dialectology, sociolinguistics and phonetics. With respect to data collection methods, Chinese dialectologists aim for a representative and comprehensive description of dialects, thus, like western dialectologists, they usually use "direct probes" to elicit dialect forms (Wolfram & Schilling-Estes, 1998: 21), such as word lists. Phoneticians aim for experimental control and replicability, thus they also often elicit read speech. Considering the goal of obtaining the full spectrum of phonological contexts for tonal variation in the Wú dialects and making comparisons among different sites, the best choice for this study was to use mainly read speech. Sociolinguists aim for natural, vernacular speech to avoid "observer's paradox"; thus, they usually try to obtain spontaneous speech. To take the concerns of sociolinguistic tradition, three types of reading style (a reading passage, a word list, and a minimal pair list), which is a manifestation of intraspeaker variation, is also given considerations.

The classic sociolinguistic interview consists of four parts: (1) reading a list of minimal pairs, (2) reading a list of words in isolation, (3) reading a short narrative, and (4) talking with the interviewer. The first three yield highly monitored, careful speech, the fourth one was designed to elicit spontaneous speech. The order of tasks was as follows: (1) reading a list of monosyllable (free or bound) morphemes, (2) reading a list of bisyllabic words, (3) reading a passage, (4) reading a list of minimal pairs, and (5) having a structured interview to elicit spontaneous speech. The recording session of each participant was approximately 70 mins in length, 10 mins for becoming familiar with the interviewer and 60 mins for the actual recording.

The elicitation of monosyllabic morphemes and bisyllabic words was designed according to the methodological framework of modern Chinese dialectology. In modern Chinese dialectology, the standard method of investigating the phonological system of a given dialect is to use *Fāngyán Diàochá Zìbiǎo* (*Dialect Survey Word*

List, 方言调查字表, 1956, hereafter referred to as *Zìbiǎo*) to elicit data. *Zìbiǎo* includes just over 3700 monosyllabic morphemes, which are arranged according to the MC phonological categories. In other words, the arrangement of these morphemes in *Zìbiǎo* indicates their historical categories in the MC rime table. This scheme enables researchers to observe the development of a given dialect with respect to MC and also to other Chinese dialect varieties. Given the idea of *Zìbiǎo*, the speech materials of the present research were also arranged according to the MC phonological categories; to be specific, Bādiào (T1-T8) of MC. Because one third of the 33 Wú dialects Chao (1928) investigated in the 1920s had eight tonal categories, which coincided with Bādiào, rather than the four categories of Sìshēng (*Píng*, *Shǎng*, *Qù* and *Rù*).

Zìbiǎo only contains monosyllable morphemes, thus it can just elicit forms of citation tone. For the examination of tone sandhi, this study followed Chao's (1928) *Studies in the Modern Wú Dialects* (1928: 147-148), filling bisyllabic words into 64 tonal combinations (eight tonal categories of the first syllable and eight tonal categories of the second syllable). This method is also presented in the textbook of Chinese dialectology (Zhān, 2001: 171) .

Like the phonetic experiments, the selection of monosyllabic morphemes, bisyllabic words and minimal pairs follow the principle of "maximal repetition" to give full play to the control power of a reading task. "Maximal repetition" means to use the same segmental structures repeatedly in different tones and tasks as much as possible in order to control the segmental effects, even lexical effects if using the same character. Sections 3.4.2 to 3.4.5 will illustrate this principle in detail.

3.4.2. Monosyllabic morphemes

Monosyllabic morphemes were selected on the basis of Middle Chinese tonal categories. More precisely, they are all characters from *Zìbiǎo*. Appendix I lists the final selection, 136 characters from eight MC tonal categories. It is noteworthy here that not only in the task of monosyllabic morphemes, but also in all the other tasks, data for the eight MC tonal categories was collected. An exploratory analysis of tonal variables showed that only the variation for T2, T4 and T6 was found in all six regions. Given the limited time and the priority of regional comparison, only T2, T4 and T6 are analyzed in this dissertation. Materials covering eight MC tones are presented in Appendixes I to III for reference. The materials for selecting T2, T4 and

T6, applying the same criteria, are elaborated upon in this section. The criteria for selecting monosyllabic morphemes are discussed below:

(1) All are real morphemes in the Shànghǎi and Wúxī dialects.

(2) All are CV syllables in order to simplify the research design.

(3) Most of the selected characters are from the list of the top 2500 high frequency characters in modern Chinese (State Language Commission, 1988), except the three words marked by a frame in Appendix I.

(4) Monosyllabic words, i.e., free monosyllabic morphemes, were chosen as much as possible but bound monosyllabic morphemes were also included. As defined in Norman (1988: 7), a morpheme is considered free if it can be uttered independently in any context and bound only if it is bound in all contexts. In Modern Chinese, many monosyllabic words from Classical Chinese were developed into polysyllabic words (ibid.). For example, 渠 (qú, 'channel') was a free morpheme in Classical Chinese but has to be used in bisyllabic word 沟渠 (gōu qú, 'channel') to express the meaning of 'channel' in Modern Chinese. The majority of words in Modern Chinese are bisyllabic, a combination of two monosyllabic morphemes. So it is difficult to exclusively use free monosyllabic morphemes while fulfilling other criteria. When analyzing the variation of citation tone, the bound–free status of monosyllabic morphemes in the Wú dialects is considered as a factor.

(5) As discussed in Section 2.3, it is not sufficient to use the eight tonal categories to trace the development of tone. Initial consonants also play a significant role in differentiating the paths of tonal development in some Wú dialects (Chang, 1975; Ho, 1988; Norman, 1988). So, within the frame of Bādiào, further division of initials should be considered.

Chapter 2 introduced the MC tone, which has registers of Yīn (upper) and Yáng (lower). The split of Yīn and Yáng registers are defined by the initial consonant. If the initial is voiceless, i.e., qīng (clear, 清) in terms of traditional Chinese phonology, the syllable belongs to the Yīn register. Syllables with voiced initial, i.e., zhuó (muddy, 浊), are all from the Yáng register.

In MC, the unvoiced group (qīng) can be further split into two subgroups (1) unaspirated unvoiced obstruents, i.e., quánqīng (fully clear, 全清) and (2) unvoiced aspirated stops and affricates, i.e., cìqīng (secondary clear, 次清); the voiced group (zhuó) can be further split into (1) voiced obstruents, i.e., quánzhuó (fully muddy, 全浊) and (2) sonorants, i.e., cìzhuó (secondary muddy, 次浊). So, according to the classification of MC tones and consonants, two independent variables are embedded in the design: (1) eight MC tonal categories and (2) the two-way split of initials within each MC tonal category. Therefore, 16 possible combinations are made in Table 3.2 and then monosyllabic morphemes were selected to fill in these 16 cells.

Table 3.2 Sixteen categories for eliciting monosyllabic morphemes

register	initial classification	tonal Category			
		Píng	***Shǎng***	***Qù***	***Rù***
Yīn	A: unaspirated	T1A	T3A	T5A	T7A
	B: aspirated	T1B	T3B	T5B	T7B
Yáng	A: obstruent	T2A	T4A	T6A	T8A
	B: sonorant	T2B	T4B	T6B	T8B

For each cell of Table 3.2, consonants with different places of articulation and different manners of articulation were selected as initials to make them as diversified as possible as the interaction of consonants and tones has not been comprehensively investigated. The first three rows of Table 3.2 contain obstruents while the last row contains sonorants. With respect to the obstruents, alveolar plosive /t/, alveolar affricate /ts/, labio-dental fricative /f/, alveolo-palatal /tʂ/, bilabial plosive /p/ and velar plosive /k/ were used for the unaspirated voiceless morphemes, i.e., morphemes of T1A, T3A, T5A and T7A in the first row. The aspirated versions of the above five stops, i.e., /tʰ/, /tsʰ/, /tʂʰ/, /pʰ/ and /kʰ/, were applied to the morphemes of T1B, T3B, T5B, T7B in the second row, paired with those unaspirated voiceless morphemes. For the obstruents in the Yáng register, the same six unaspirated obstruents were applied to the third row but followed by a voiced glottal fricative /ɦ/, in other words, /tɦ/, /tsɦ/, /fɦ/, /tʂɦ/, /pɦ/ and /kɦ/ were used for the morphemes of T2A, T4A, T6A and T8A. Because, in the Wú dialects, the phonetic feature of morphemes in the Yáng register is a voiceless obstruent followed by a voiced glottal fricative [ɦ] and breathy-voiced vowels (Chen 2010). With respect to sonorant, bilabial nasal /m/, alveolar nasal /n/, glottal approximant /ɦ/ and alveolar lateral approximant /l/ were applied to the morphemes of T2B, T4B, T6B and T8B in the fourth row.

(6) Vowels. High frequency monophthongs in the Shànghǎi and Wúxī dialects were chosen for testing monosyllables. Vowels with different pitch height are expected to have an intrinsic influence on the F_0 height: high vowels or close vowels (low F_1) have a higher F_0 than low vowels or open vowels (Hombert, et al., 1979; Lehiste, 1970). Diphthongs were not used in order to avoid pitch variation caused by differences in vowel heights within the syllable.

The monophthongs of *Píng*, *Shǎng*, and *Qù* syllables are /ɑ/, /ɛ/, /ʌ/, /u/, /y/, /i/, and /ʅ/ in the Wúxī urban dialect and /ᴀ/, /ɛ/, /ɔ/, /u/, /y/, /i/, and /ʅ/ in the Shànghǎi urban dialect. In Wú dialects, monophthongs of *Rù* end in a glottal stop. Monosyllables of *Rù* are /ʌʔ/, /ɔʔ/, /ɪʔ/ and /əʔ/ in the Wúxī urban dialect and /ɐʔ/, /ɔʔ/, /iɪʔ/ and /əʔ/ in the Shànghǎi urban dialect. These monophthongs are not only frequently used, but were chosen to maintain an acoustic balance of vowel position. They cover close and open, front and back vowels. Since vowel height is intrinsically related to F_0 (Hombert, et al., 1979; Lehiste, 1970), vowel height will be considered as a linguistic predictor in the analyses of pitch height.

This study focuses on the variations in T2, T4 and T6 morphemes, so the monosyllabic morphemes in these three tonal categories in MC are listed in Table 3.3 to Table 3.5. The column "Wú Pron" in each table provides the pronunciation in "Common Northern Wu system" proposed by Simmons (1999: 50-77). As reviewed in Section 2.2, dialects spoken in Wúxī and Shànghǎi all belong to the "Tàihú Sub-Group" and therefore share common phonology. The "common Northern Wu system" can represent the common phonological categories of the dialects of the Tàihú area (Simmons 1999: 50). The phonetic realizations in the Wú dialects are not provided here. The dialects from the six selected regions show variation in the initials and finals and it is reported that they are undergoing changes in those dialects (Cáo, 2003; Qian, 2010). As their exact phonetic nature is outside the scope of this study, we stick to a phonemic representation.

Table 3.3 Monosyllabic morphemes of MC T2 in word list reading

character		Wú Pron.	PTH Pron.	morpheme	gloss
I.	排	pɦa	[pʰai^{35}]	F	to be in a row
II.	抬	tɦé	[tʰai^{35}]	F	to lift up
III.	曹	sɦau	[tsʰau^{35}]	F	a surname
IV.	湖	ɦu	[xu^{35}]	F	lake
V.	渠	kɦiu	[tɕʰy^{35}]	B	channel
VI.	题	tɦi	[tʰi^{35}]	B	title
VII.	词	sɦy	[tsʰʅ^{55}]	F	word
VIII.	咸	ɦén	[xian^{35}]	F	salty
i.	埋	mɦa	[mai^{35}]	F	to bury
ii.	篮	lɦén	[lan^{35}]	F	basket
iii.	牢	lɦau	[lau^{35}]	F	sturdy and strong
iv.	吴	ng	[u^{35}]	F	a surname
v.	余	ɦiu	[y^{35}]	B	extra
vi.	迷	mɦi	[mi^{35}]	B	be crazy about
vii.	元	djy	[uan^{35}]	B	currency unit Yuan
viii.	毛	mɦau	[mau^{35}]	F	hair, feather

Table 3.4 Monosyllabic morphemes of MC T4 in word list reading

character		Wú Pron.	PTH Pron.	morpheme	gloss
I.	罢	pɦa	[pa^{51}]	B	to quit
II.	诞	tɦén	[tan^{51}]	B	birth
III.	造	sɦau	[tsau^{51}]	F	to make
IV.	户	ɦu	[xu^{214}]	B	household
V.	聚	sɦiu	[tɕy^{51}]	F	to get together
VI.	弟	tɦi	[ti^{51}]	B	young brother
VII.	似	zy	[sɿ^{51}]	B	similar
VIII.	腐	fɦu	[fu^{214}]	B	rotten
IX.	上	djang	[ʂɑŋ^{51}]	F	to go up
X.	道	tɦau	[tau^{51}]	B	road
i.	买	mɦa	[mai^{214}]	F	to buy
ii.	懒	lɦén	[lan^{214}]	F	lazy
iii.	老	lɦau	[lau^{214}]	F	old
iv.	午	ng	[u^{214}]	B	noon
v.	雨	ɦiu	[y^{214}]	F	rain
vi.	米	mɦi	[mi^{214}]	F	rice
vii.	满	mɦu	[man^{214}]	F	full
viii.	冷	lɦang	[ləŋ^{214}]	F	cold
ix.	礼	lɦi	[li^{214}]	F	gift

Table 3.5 Monosyllabic morphemes of MC T6 in word list reading

character		Wú Pron.	PTH Pron.	morpheme	gloss
I.	败	pɦa	[pai^{51}]	B	failure
II.	蛋	tɦén	[tan^{51}]	F	egg
III.	召	sɦau	[tsau^{51}]	B	call together
IV.	护	ɦu	[xu^{51}]	B	to protect
V.	具	kɦiu	[tɕy^{51}]	B	tool
VI	第	tɦi	[ti^{51}]	B	order, degree, class
VII.	寺	zy	[sɿ^{51}]	B	temple
VIII.	字	zy	[tsɿ^{51}]	F	word, character
IX.	大	tɦū	[ta^{51}]	F	big, great
X.	附	fɦu	[fu^{51}]	B	to attach
i.	卖	mɦa	[mai^{51}]	F	to sell
ii.	烂	lɦén	[lan^{51}]	F	rotten
iii.	涝	lɦau	[lau^{51}]	B	water logging
iv.	误	ng	[wu^{51}]	B	error
v.	预	ɦiu	[ju^{51}]	B	to predict
vi.	谜	mɦi	[mi^{35}]	F	puzzle
vii.	面	mɦien	[mian^{51}]	F	flour
viii.	闹	nau	[nau^{51}]	F	noisy

3.4.3. Bisyllabic words

Data for bisyllabic words was collected by means of the word lists in Appendix II. The 128 bisyllabic words in Table 1 were chosen according to the following criteria: (1) the table is designed in the frame of MC tonal categories, dividing the first syllable (σ_1) into 16 groups. The leftmost three columns indicate σ_1's tonal category and the property of its initial consonant. The top row indicates the second syllable's

(σ_2) tonal category. (2) Bisyllabic words' σ_1 uses the morphemes tested in the monosyllabic list whenever possible, in order to get comparable data with the monosyllabic morphemes. Those comparable syllables are marked in bold. (3) All bisyllabic words are well-formed in PTH. The criteria for well-formedness in PTH are those accepted in *The Contemporary Chinese Dictionary* (Institute of Linguistics, 2002) , but not indicated as dialect vocabularies. These were used in place of the typical dialect to ensure that the young speakers can read them naturally and comfortably, because it is important to be able to compare the bisyllabic with the monosyllabic data. The dialect speakers of the young generation have a somewhat reduced vocabulary in comparison with the older generation. It was observed by the author that some young speakers in Wúxī and Shànghǎi do not know many words that are unique to the dialect and thus cannot read them.

Table 2 in Appendix II contains the dialect words of the Wúxī and Shànghǎi dialects. The criteria for choosing these words were (1) Words accepted by the *Dialects in Shànghǎi Urban area* (Xǔ & Tāng, 1988) and *The Dictionary of Modern Chinese Dialects (Shànghǎi dialect Volume)* (Xǔ & Táo, 1997; Xǔ & Táo, 1997) and also frequently used both in Wúxī and Shànghǎi areas according to the judgments of two Shànghǎi native speakers and two Wúxī native speakers; (2) Words that are not accepted by the *Contemporary Chinese Dictionary* or accepted but classified as dialect words. As discussed above, young speakers have difficulties in reading some of these stimuli. The unreadable tokens for young speakers will be discussed separately in Chapters 5 and 6.

Xú (2007) did a thorough investigation on the Wúxī sandhi patterns and found some exceptions to the tone sandhi rules summarized by Chan & Ren (1989) and Chen (2000). These exceptions include both dialect and PTH words, which are listed in Table 6 of Appendix II and tested both in the Shànghǎi and Wúxī dialects.

Table 3.6 to Table 3.8, below, present the bisyllabic testing words initialed with T2, T4 and T6 morphemes separately. These words are phrases of T2.X, T4.X and T6.X in Table 1, Table 2 and Table 6 of Appendix II. The column "PTH Pron." in these tables indicates not only those words' PTH pronunciation, but also whether they are well-formed in PTH. The exclusively dialect words are marked with "N/A".

Table 3.6 T2.X phrases elicited in word list reading

tone	phrase	Wú Pron.	PTH Pron.	gloss
T2A.T1	曹操	sɦau tsʿau	[tsʰau^{35} tsʰau^{55}]	Cao Cao, a historical figure in China
T2A.T1	调羹	tɦiau káng	[tʰiau^{35} kən^{55}]	spoon
T2A.T1	停当	tɦing tɦang	[tʰiŋ^{35} tɑŋ^{55}]	be completed
T2A.T1	辰光	ɕɦeng kuang	N/A	time
T2A.T1	台灯	tʿé teng	[tʰai^{35} tən^{55}]	desk lamp
T2A.T1	层单	sɦeng tén	N/A	sheet
T2A.T1	提高	tɦi kau	[tʰi^{35} kau^{55}]	to raise or improve
T2A.T1	馋胚	sɦén pʿé	N/A	glutton
T2B.T1	迷宫	mi kong	[mi^{35} kuŋ^{55}]	maze
T2B.T1	洋机	ɦiang ki	[iɑŋ^{35} tɕi^{55}]	sewing machine
T2B.T1	阳台	ɦiang tʿé	[iɑŋ^{35} tʰai^{55}]	balcony
T2B.T1	莲心	lɦi sing	N/A	lotus seed
T2A.T2	排球	pɦa kɦiou	[pʰai^{35} tɕʰiou^{35}]	volleyball
T2B.T2	梧桐	ɦu tɦong	[xu^{35} tuŋ^{35}]	Chinese parasol tree
T2B.T2	名堂	ming tɦang	N/A	variety
T2B.T2	胡咙	ɦu long	N/A	throat
T2B.T2	洋盘	ɦiang pu	N/A	one not smart enough
T2B.T2	浑堂	ɦueng tɦang	N/A	bathhouse
T2A.T3	词组	sɦy tsu	[tsʰɿ^{35} tsou^{214}]	phrase
T2A.T3	团长	tɦun tɕáng	[tʰuɛn^{35} tʂɑŋ^{214}]	team leader
T2B.T3	篮板	lɦén pén	[lan^{35} pan^{35}]	backboard
T2B.T3	难板	nén pén	N/A	seldom
T2B.T3	门坎	meng kʿén	[mən^{35} kʰan^{214}]	door sill
T2B.T3	元宝	djy pau	[yɛn^{35} pau^{35}]	(gold or silver) ingot
T2A.T4	渠道	kɦiu tɦau	[tɕʰy^{35} tau^{51}]	channel
T2B.T4	厘米	lɦi mɦi	[li^{35} mi^{214}]	centimeter
T2A.T5	咸菜	ɦén tsʿé	[ɕiɛn^{35} tsʰai^{51}]	brined vegetables
T2A.T5	残废	sɦén fi	[tsʰan^{35} fei^{51}]	disabled
T2B.T5	牢靠	lɦau kʿau	[lau^{35} ku^{51}]	secure
T2B.T5	文旦	fɦeng tén	N/A	pomelo
T2B.T6	余地	ɦiu tɦi	[y^{35} ti^{51}]	leeway
T2B.T6	闲话	ɦén ɦuo	N/A	gossip
T2B.T6	原旧	ngiuon kɦiou	[yɛn^{35} tɕiou^{51}]	as usual
T2B.T6	和调	ɦū tɦiau	N/A	to play a joke
T2B.T6	迷露	mɦi lɦu	N/A	fog
T2A.T7	湖北	ɦu poq	[xu^{35} pei^{214}]	a province in China
T2B.T7	毛笔	mɦau pieq	[mau^{35} pi^{214}]	writing brush
T2B.T7	难得	nén teq	[nan^{35} tɤ^{35}]	rare
T2A.T8	题目	tɦi mɦoq	[ti^{35} mu^{51}]	topic
T2B.T8	埋没	mɦa mo	[mai^{35} mo^{51}]	to bury
T2B.T8	寒热	ɦun ngieq	N/A	fever

Table 3.7 T4.X phrases elicited in word list reading

tone	phrase	Wú Pron.	PTH Pron.	gloss
T4A.T1	罢工	pɦa kong	[pa^{51} kuŋ^{55}]	strike
T4A.T1	像腔	ziang k‘iang	N/A	to behave properly
T4A.T1	上腔	djáng k‘iang	N/A	to retort sarcastically
T4B.T1	买方	mɦa fang	[mai^{214} fɑŋ^{55}]	buyer
T4B.T1	午休	ng hiou	[u^{214} ɕiou^{55}]	noon break
T4B.T1	雨披	ɦiu p‘i	[y^{214} pʰi ^{55}]	poncho
T4B.T1	老官	lɦau kuon	N/A	husband
T4B.T1	懒胚	lɦén p‘é	N/A	lazy bone
T4B.T1	眼泡	ngén p‘au	N/A	eyelid
T4B.T2	冷盆	lɦang pɦeng	[ləŋ^{214} pʰən^{35}]	cold dishes
T4B.T2	理由	lɦi ɦiou	[li^{214} iou^{35}]	reason
T4B.T2	羽毛	iu mɦau	[y^{214} mɑu^{35}]	feather
T4B.T2	领头	lɦing tɦou	N/A	collar
T4A.T3	户口	ɦu k‘ou	[xu^{51} kʰou^{214}]	registered residence
T4B.T3	老虎	lɦau hu	[lau^{214} xu^{214}]	tiger
T4B.T3	老酒	lɦau tsiou	N/A	liquor
T4B.T3	耳朵	ngi tū	[ɚ^{214} tuo]	ear
T4A.T4	腐乳	fɦu djy	[fu^{214} ʐu^{214}]	preserved beancurd
T4A.T4	上下	djáng ɦia	[ʂɑŋ^{51} ɕia^{51}]	up and down
T4A.T4	丈姆	ɕɦáng m	N/A	mother-in-law
T4B.T4	蚂蚁	mo ngi	[ma^{214} i^{214}]	ant
T4A.T5	上进	djáng tsing	[ʂɑŋ^{51} tɕin^{51}]	enterprising
T4A.T5	造句	sɦau kiu	[tsau^{51} tɕy^{51}]	to make sentences
T4A.T5	负数	fu su	[fu^{51} ʂu^{51}]	negative number
T4A.T5	鳝片	ɕɦūn p‘ien	[ʂan^{51} pʰian^{51}]	eel slice
T4A.T5	上照	djáng tɕau	N/A	photogenic
T4A.T5	上劲	djáng king	N/A	enthusiastic about doing something
T4B.T5	礼拜	lɦi pa	[li^{214} pai^{51}]	week
T4B.T5	惹气	dja k‘i	N/A	angry
T4B.T5	有劲	ɦiou king	N/A	interesting
T4A.T6	近视	kɦing sɦy	[tɕin^{51} ʂi^{51}]	myopic
T4A.T6	部队	pɦu tɦé	[pu^{51} tuei^{51}]	army
T4A.T6	上路	djáng lɦu	N/A	reliable
T4A.T6	道地	tɦau tɦi	N/A	considerable
T4B.T6	懒惰	lɦén tū	[lan^{214} tuo^{51}]	lazy
T4A.T7	道德	tɦau teq	[tau^{51} tɤ^{35}]	morality
T4A.T7	挺刮	t‘ing kuaq	[tʰiŋ^{214} kuɑ^{55}]	(clothes) neat
T4B.T7	米色	mɦi seq	[mi^{214} sɤ^{51}]	cream color
T4A.T8	聚集	sɦiu sɦieq	[tɕy^{51} tɕi^{35}]	to gather
T4A.T8	序幕	sɦiu mɦu	[ɕy^{51} mu^{51}]	prologue
T4B.T8	满月	mɦu ngiueq	[man^{214} yɛ^{51}]	the achievement of the first month of life
T4B.T8	暖热	nun ngieq	N/A	warm

Table 3.8 T6.X phrases elicited in word list reading

tone	phrase	Wú Pron.	PTH Pron.	gloss
T6A.T1	召开	sɦau k'é	[tʂau^{51} kʰai^{55}]	to convoke
T6A.T1	自家	zy ka	N/A	oneself
T6A.T1	便当	pɦien tɦang	[piɛn^{51} tɑŋ^{55}]	convenient
T6B.T1	闹钟	nau tɕong	[nau^{51} tʂuŋ^{55}]	alarm clock
T6A.T2	蛋黄	tɦén ɦuang	[tan^{51} xuɑŋ^{35}]	yolk
T6A.T2	寿头	djou tɦou	N/A	a fool
T6A.T2	旧年	kɦiou ngien	N/A	last year
T6A.T2	调查	tɦiau sɦo	[tiau^{51} tʂʰa^{35}]	to investigate
T6B.T2	烂糊	lɦén ɦu	[lan^{51} xu^{35}]	pulpy
T6B.T2	弄堂	lɦong tɦang	N/A	alley
T6B.T2	用场	ɦiong tɕ'áng	[yŋ^{51} tʂʰɑŋ^{214}]	function
T6A.T3	字典	zy tien	[tsɿ^{51} tian^{214}]	dictionary
T6A.T3	事体	sɦy t'i	N/A	affair; matter
T6A.T3	具体	kɦy t'i	[tɕy^{51} tʰi^{214}]	specific
T6B.T3	效果	ɦiau kū	[ɕiau^{51} kuo^{214}]	effect
T6B.T3	面孔	mɦien k'ong	[miɛn^{51} kuŋ^{214}]	face
T6A.T4	护士	ɦu sɦy	[xu^{51} ʂʅ^{51}]	nurse
T6A.T4	运道	ɦiuing tɦau	[yŋ^{51} tau^{51}]	fortune
T6A.T4	导演	tɦau ien	[tau^{214} iɛn^{214}]	directors of film, drama, etc.
T6B.T4	料重	lɦiau ɕɦong	N/A	(burden) heavy
T6A.T4	万岁	mɦén sué	[øuɑŋ^{51} suei^{51}]	long live
T6B.T4	号码	ɦau mo	[xau^{51} ma^{214}]	number
T6B.T4	闹猛	nau mɦeng	N/A	bustling with noise and excitement
T6B.T4	谜语	mɦi ɦiu	[mi^{35} y^{214}]	riddle
T6A.T5	大蒜	tɦū sun	[ta^{51} suan^{51}]	garlic
T6A.T5	硬劲	ngáng king	N/A	manage to do sth with difficulty
T6B.T5	卖相	mɦa siang	[mai^{51} ɕiɑŋ^{51}]	appearance
T6B.T5	夜快	ɦia k'ua	N/A	at night fall
T6B.T5	浪费	lɦang fi	[lɑŋ^{51} fei^{51}]	waste
T6A.T6	寺庙	zy mɦiau	[sɿ^{51} miau^{51}]	temple
T6A.T6	自愿	zy ngiuon	[tsɿ^{51} yɛn^{51}]	voluntarily
T6B.T6	夜饭	ɦia fɦén	N/A	supper
T6B.T6	烂饭	lɦén fɦén	N/A	rice cooked with too much water
T6B.T6	预料	ɦiu lɦiau	[y^{51} liau^{51}]	to predict
T6B.T6	议论	ngi lɦeng	[i^{51} lun^{51}]	to discuss
T6A.T7	硬扎	ngáng tsāq	N/A	firm, strong
T6A.T7	败笔	pɦa pieq	[pai^{51} pi^{214}]	a faulty expression in writing
T6B.T7	预测	ɦiu ts'e	[y^{51} tsʰɤ^{51}]	to forecast
T6A.T8	附属	fɦu djoq	[fu^{51} ʂu^{214}]	affiliated
T6B.T8	赖学	lɦa ɦōq	[lai^{51} ɕyɛ^{35}]	to play truant
T6B.T8	卖力	mɦa lɦieq	[mai^{51} li^{51}]	to work hard

In comparison with the Shànghǎi dialects, the tone sandhi rules in Wúxī are more complicated, and conditioned by grammatical structures. The complexity of tone sandhi patterns in the Wú dialect has been documented extensively (Chan & Ren, 1989; M. Y. Chen, 2000; Xú, 2007). According to Chan & Ren's (1989) instrumental study of tone sandhi phenomena in Wúxī, Chen (2000: 320-325) distinguished two types of tone sandhi in the Wúxī dialect (see Section 2.4.4.2).

The words in Table 3.6 to Table 3.8 only cover words of Type II. The sandhi patterns of Type I (a), (b) and (c) are presented separately in Tables 3 to 5 in Appendix II. Words initialed with T2, T4 and T6 are listed in Table 3.9 to Table 3.11 below. Table 3.9 lists the reduplicated forms of six frequently used verbs using T2A, T2B, T4A, T4B, T6A and T6B.

Table 3.9 Phrases for eliciting reduplicated verbs in T2, T4 and T6

tone	phrase	Wú Pron.	PTH Pron.	gloss
T2A.T2A	陪陪	pɦé pɦé	[pʰei^{35} pʰei]	to accompany
T2B.T2B	量量	liang liang	[liaŋ^{35} liaŋ]	to measure
T4A.T4A	动动	tɦong tɦong	[tuŋ^{51} tuŋ]	to move
T4B.T4B	咬咬	ngau ngau	[ŋʌ^{13} ŋʌ]	to bite
T6A.T6A	垫垫	tɦie tɦie	[tian^{51} tian]	to pad with
T6B.T6B	练练	lie lie	[lian^{51} lian]	to practice

In Table 3.10, Part 1 lists three verbs from MC T2, T4 and T6. Part 2 lists eight morphemes from T1-T8, expressing behavioral outcomes or trends. The combinations of Part 1 and Part 2 morphemes result in 24 verb phrases for testing the tone sandhi of Type I (2): "verb + resultative or directional complements". In Table 3.11, Part 1 lists two numerical words in T4 and T6[28]. T2 is an accidental gap. Part 2 of Table 3.11 lists eight classifiers from eight tonal categories. Those eight classifiers are followed by numerical words "两" (two) and "廿" (twenty) to result in 16 verb phrases for testing the tone sandhi of Type I (3): "expressions with 'number + classifier'".

[28] It is noteworthy here that Niàn (廿, twenty) is the only choice for eliciting a T6 numerical word. It might be controversial to use Niàn (廿) because it is not listed in *Zìbiǎo*, though it is indicated as T8 in another authoritative source: *Gǔ-Jīn Zìyīn Duìzhào Shǒucè* (古今字音对照手册, *Handbook of Comparison of Modern and Ancient Sounds of Characters*, Dīng & Lǐ, 1966). However, the pronunciation of Niàn (廿) has changed to the homophone of T6 word Niàn (念, to read loudly) in the Wú dialects, which can be dated back to the 11th century (Hè, 2012). So, it is better to test this controversial Niàn (廿) rather than leaving another gap.

Table 3.10 Words in T2, T4 and T6 for eliciting expressions with “verb + resultative or directional complements”

tone	char.	Wú Pron.	PTH Pron.	gloss
Part 1: σ₁				
T2	磨	mū	[mo³⁵]	to grind
T4	斩	tsén	[tʂan²¹⁴]	to chop
T6	弄	long	[noŋ⁵¹]	to make, to do
Part 2: σ₂				
T1	开	kʰé	[kai⁵⁵]	open
T2	平	pɦing	[piŋ³⁵]	flat
T3	好	hau	[hau²¹⁴]	the action has been done (satisfyingly)
T4	断	tɦun	[tuan⁵¹]	broken into two parts
T5	破	pʰū	[po⁵¹]	broken into pieces
T6	坏	ɦua	[huai⁵¹]	broken
T7	出	tɦʰueq	[tɕʰu⁵⁵]	out
T8	落	lōq	[luo⁵¹]	the action has been done

Table 3.11 Morphemes in T2, T4 and T6 for eliciting expressions with “number + classifier”

tone	char.	Wú Pron.	PTH Pron.	gloss
Part 1: σ₁				
T2	/	/	/	/
T4	两	liang	[liaŋ²¹⁴]	two
T6	廿	ngien	[niɛn³⁵]	twenty
Part 2: σ₂				
T1	斤	king	[tɕin⁵⁵]	a weight used in China, currently equal to 0.5 kg
T2	年	ngien	[niɛn³⁵]	year
T3	把	po	[pa²¹⁴]	a handful of
T4	米	mi	[mi²¹⁴]	meter
T5	块	kɦua	[kʰuai⁵¹]	a measure for pieces and monetary units
T6	段	tɦun	[tuan⁵¹]	paragraph
T7	只	tseq	[tʂʅ⁵⁵]	a measure for animals like hens, birds, tigers, etc.
T8	粒	lieq	[li⁵¹]	a measure for small round pieces

3.4.4. Minimal pairs

The variations among T2, T4 and T6 are the focus of this study. A minimal pair test is designed to check whether the participants can distinguish tonal differences among T2, T4 and T6. There are two parts in the list of minimal pairs. Table 3.12 consists of three sets of two-tone pairs, T2 & T4, T4 & T6 and T2 & T6. There are six pairs in each set, three with obstruent initials and three with sonorants. In addition, 12 three-tone pairs, six pairs with obstruents and six with sonorants were also tested. Morphemes in three-tone pairs are all the morphemes tested in the monosyllabic list, marked by Roman numerals (I-VI & i-vi) in Table 3.3 to Table 3.5. Morphemes marked with the same Roman numeral constitute one three-tone pair.

Table 3.12 Word lists of two-tone minimal pairs

tone	char	Wú pron.	PTH pron.	morpheme	gloss
T2	糖	tɦang	[tʰɑŋ^{35}]	F	sugar
T4	荡		[tɑŋ^{51}]	F	to swing
T2	狼	lɦang	[lɑŋ^{51}]	F	wolf
T4	朗		[lɑŋ^{214}]	B	loud and clear
T2	盘	pu	[pʰan^{35}]	F	plate
T4	伴		[pan^{51}]	F	partner
T2	瞒	mɦu	[man^{35}]	F	hide the truth from
T4	满		[man^{214}]	F	full
T2	球	kɦiou	[tɕʰiou^{35}]	F	ball
T4	舅		[tɕiou^{51}]	B	mother's brother
T2	邮	ɦiou	[iou^{35}]	B	to post
T4	友		[iou^{214}]	B	friend
T4	肚	tou	[tu^{51}]	B	belly, abdomen
T6	度		[tu^{51}]	B	extent
T4	鲁	lɦu	[lu^{214}]	F	a surname
T6	路		[lu^{51}]	F	road
T4	键	kɦien	[tɕian^{51}]	F	key, bond
T6	健		[tɕian^{51}]	F	healthy
T4	免	mɦien	[mian^{214}]	B	to avoid
T6	面		[mian^{51}]	F	flour
T4	奉	fɦong	[fəŋ^{51}]	F	to accept
T6	凤		[fəŋ^{51}]	F	phoenix
T4	涌	ɦiong	[ioŋ^{214}]	B	to gush out
T6	用		[ioŋ^{51}]	F	to use
T2	头	tɦou	[tʰou^{35}]	F	head
T6	豆		[tou^{51}]	F	bean
T2	楼	lɦou	[lou^{35}]	F	building
T6	漏		[lou^{51}]	F	to leak
T2	逃	tɦau	[tʰau^{35}]	F	to escape
T6	稻		[tau^{51}]	B	paddy, rice
T2	毛	mɦau	[mau^{35}]	F	hair, feather
T6	帽		[mau^{51}]	F	hat
T2	坟	fɦeng	[fən^{35}]	F	grave, bomb
T6	份		[fən^{51}]	B	portion
T2	云	ɕɦən	[jən^{35}]	F	cloud
T6	运		[jən^{51}]	F	to transport

3.4.5. Paragraph reading

Chao (1928) first used the fable *The North Wind and the Sun* to elicit the spoken language in Chinese dialects. Since then it has also been adopted widely for phonetic demonstrations in Chinese dialectology (e.g. Qián, 1992; Hóu, 1999). But *The North Wind and the Sun* does not offer appropriate materials for examining tone contrasts, mainly because it contains only 168 words in Wú version (Chao 1928: 157-158), which is not enough to elicit all types of citation and sandhi forms. To elicit as many citation forms as possible, this study used an essay introducing Chinese names (see

Appendix III). This essay has two paragraphs and 736 words in total. The text is widely known by the Chinese people and easily understood. It was also designed with three linguistic aims: (1) it consists of monosyllabic morphemes in all eight isolated tonal categories; (2) these morphemes are mostly free morphemes that are less influenced by intonation when reading. When reading those words, punctuation before and after can make participants insert pauses, resulting in stressed words and reducing the impact of intonation on tone. (3) The text consists of as many types of sandhi patterns as possible. It is not feasible to have every type of 64 bisyllabic combinations in two short paragraphs. The text is designed to include combinations initialed with eight MC tones. Materials relevant to T2, T4 and T6 are present in Table 3.13 (monosyllabic morphemes) and Table 3.14 (bisyllabic words).

Table 3.13 T2, T4 and T6 monosyllabic morphemes in paragraph reading

tone	character	Wú Pron.	PTH Pron.	morpheme	gloss
T2A	陈	ɕɦeng	[tʂʰən^{35}]	F	a surname
T2A	晨	ɕɦeng	[tʂʰən^{35}]	B	morning
T2B	刘	lɦiou	[liou^{35}]	F	a surname
T2B	仁	djeng	[ʐən^{35}]	F	benevolence
T4A	赵	ɕɦau	[tsau^{51}]	F	a surname
T4B	李	lɦi	[li^{214}]	F	a surname
T4B	礼	lɦi	[li^{214}]	F	etiquette
T4B	勇	ɕɦiong	[ioŋ^{214}]	F	courage
T6A	寿	djou	[ʂou^{51}]	F	longevity
T6A	健	kɦien	[tɕian^{51}]	F	health
T6B	廖	lɦiau	[liau^{51}]	F	a surname
T6B	义	ngi	[ji^{51}]	F	righteousness

Table 3.14 T2.X, T4.X and T6.X phrases elicited in paragraph reading

tone	word	Wú Pron.	PTH Pron.	gloss
T2A.T1	其中	kɦi tɕong	[tɕʰi^{35} tʂuŋ^{55}]	among
T2A.T1	时间	ɕɦy kien	[ʂi^{35} tɕiɛn^{55}]	time
T2A.T2	排名	pɦa mɦing	[pʰai^{35} miŋ^{35}]	rank
T2A.T6	含义	ɦūn ngi	[xan^{35} i^{51}]	meaning
T2A.T6	长寿	ɕɦáng djou	[tʂʰɑŋ^{35} ʂou^{51}]	longevity
T2A.T6	常用	ɕɦáng ɦiong	[tʂʰɑŋ^{35} yŋ^{51}]	often used
T2A.T7	全国	sɦiuon kū	[tɕʰuɛn^{35} kuo^{35}]	the entire country
T2B.T2	男人	nūn nging	[nan^{35} ʐən ^{35}]	man
T2B.T2	威武	ué fɦu	[uei^{55} u^{214}]	powerful
T2B.T3	人口	nging k‘ou	[ʐən^{35} kʰou^{214}]	population
T2B.T6	名字	mɦing zy	[miŋ^{35} tsɿ ^{51}]	name
T4B.T2	女人	ngiu nging	[ny ^{214} ʐən^{35}]	woman
T4B.T3	某种	mɦou tɕong	[mou^{214} tʂuŋ^{214}]	some kind
T4B.T5	两个	lɦiang ke	[liɑŋ^{214} ke^{51}]	two
T6A.T1	健康	kɦien k‘ang	[tɕiɛn^{51} kʰan^{55}]	healthy
T6A.T4	地点	tɦi tien	[ti^{51} tiɛn ^{214}]	location
T6A.T4	具有	kɦiu ɦiou	[tɕy^{51} iou^{214}]	have, process
T6B.T6	另外	lɦing nga	[liŋ^{51} uai^{51}]	moreover
T6B.T6	愿望	ngiuon fɦuang	[yɛn^{51} uɑŋ^{51}]	desire, wish

3.4.6. Structured interview

The final part of elicitation was a one-on-one interview. The list of interview questions is given in Appendix IV. The interview was roughly divided into two broad categories: (1) language: language aptitude, language attitude, the experience of language study and language use; (2) personal information: participant's name, age, schooling, occupation and several questions on the topic of food. The linguistic questions concerned not only the dialect but also the PTH used by the participants. There were two versions of the questionnaire, if the participants were not able to speak PTH, version 2 was used. Version 2 does not include any questions on PTH.

The interview was structured to acquire background information on each speaker and some spontaneous speech on the topic of food, a neutral topic that can engage speakers in interaction. But difficulties were encountered in the acoustic analysis of tonal data. First, there were few isolated words spoken in the speech, not enough tokens for investigating the variation of citation tone. Second, the sandhi forms of bisyllabic words strongly interact with intonation. For example, the slope of a falling tone is drastically reduced in a high rising intonation marking a question or uncertainty, and the slope of a falling tone increases when it is stressed. To the author's knowledge, the interaction between lexical tones and intonation on tonal variation has not been systematically examined. Since the analysis of spontaneous speech data is complicated, this dissertation will not perform acoustic analyses on them but focus on (part of) the structured data.

3.4.7. Linguistic factors

Based on the design of data elicitation, there is a stylistic conditioning effect on tonal variation for both citation tone and tone sandhi. The contextual style has three levels for the citation tone (*word list*, *paragraph* and *the list of minimal pairs*) and two levels for the tone sandhi (*word list* and *paragraph*.) Additionally, 6 independent variables were analyzed to test the linguistic effects on the variation of citation tone whereas 12 independent variables were analyzed for the variation of tone sandhi. The linguistic predictors for citation tones include:

(1) **MC_TONE** : MC tonal category (*T2*, *T4* and *T6*)
(2) **MC_AB**: the split of *Quánzhuó* (fully muddy) and *Cìzhuó* (secondary muddy) in MC (*Quánzhuó* is denoted as *A* after MC_TONE, e.g., T2A*,* and *Cìzhuó* is denoted as *B*, e.g., T2B)

(3) **MORPHEME**: the classification of each given morpheme in the monosyllable stimuli (*free morpheme* and *bound morpheme*)
(4) **HEIGHT**: the height of nuclear vowels (*close, near close, close mid, mid, open mid, near open* and *open*)
(5) **PTH_TONE**: the corresponding tonal category in PTH (*PT1, PT2, PT* 3 and *PT4*)
(6) **DUR**: the duration of the domain of tone in each given syllable (duration in milliseconds of each token)

Among those linguistic predictors for citation tone, MC_TONE, MC_AB, HEIGHT and DUR are internal linguistic factors while MORPHEME and PTH_TONE are external linguistic factors. The correlations of MC_TONE, MC_AB, MORPHEME and HEIGHT are discussed in the beginning of Section 3.4.2. These four predictors are also embedded in the design of collecting monosyllabic morphemes. The explanations of PTH_TONE and DUR are given here. PTH_TONE is a predictor to test whether a process of tonal variation is motivated by contact with PTH or not. Meanwhile, PTH_TONE is derived from MC_TONE (see Table 3.15 below). The significant conditioning effect of PTH_TONE is possibly caused by MC_TONE (see Table 3.15). So when both PTH_TONE and MC_TONE show the significant conditioning effects on one variable, further justifications are needed to decide whether this variable is induced by contact with PTH or constrained by the MC tonal categories.

Table 3.15 The relationship between MC_TONE and PTH_TONE

MC_TONE	T1	**T2**	T3	**T4A**	**T4B**	T5	**T6**	T7	T8A	T8B
PTH_TONE	PT1	PT2		PT3	PT4			PT1-4	PT2	PT4

Regarding DUR, there is indeed an intrinsic durational variation among different tonal shapes. Howie (1976), Rose (1980) and Zhu (1999) pointed out that a rising tone is usually longer than a level tone, and both are longer than a falling tone. DUR is not the focus of this study and is therefore not examined as a dependent variable. DUR is used as a covariant for dependent variables relevant to pitch height and pitch shape to prevent omitted variable bias. As DUR is not the focus of our study, its significant effect is not discussed if its weight goes in the expected direction, e.g., longer duration opting for a rising tone in comparison to a falling tone. If its weight goes in the unexpected direction, the interpretation or discussion is provided additionally.

The linguistic predictors for tone sandhi are more complicated in comparison to the citation tone. Since according to previous research the sandhi patterns investigated in the present study are not purely left-dominant, such as the Wúxī dialect, the segmental and tonal attributes of not only σ_1 but also those of σ_2 have possible conditioning effects. The linguistic predictors investigated in the present study include:

(1) and (2)	σ_1_MC_TONE and σ_2_MC_TONE: MC tonal category of σ_1 (*T2*, *T4* and *T6*) and σ_2 (*T1*, *T2, T3, T4, T5, T6, T7* and *T8*)
(3)	MC_AB: MC split of *Quánzhuó* (fully muddy) and *Cìzhuó* (secondary muddy) for T2, T4, T6 morphemes (*Quánqīng* and *Quánzhuó* are denoted as A, *Cìqīng* and *Cìzhuó* are denoted as *B*)
(4)	LEXICAL ATTESTEDNESS: a phrase occurs only in the Wú dialects or both the Wú dialects and PTH (*Wú lexical item* and *PTH & Wú lexical item*)
(5) and (6)	σ_1_HEIGHT and σ_2_HEIGHT: the height of the nuclear vowel in σ_1 and σ_2 (*close, near close, close mid, mid, open mid, near open* and *open*)
(7) and (8)	σ_1_PTH_TONE and σ_2_PTH_TONE: the corresponding tonal category in PTH for σ_1 and σ_2 (*PT1, PT2, PT3* and *PT4*)
(9) and (10)	σ_1_DUR and σ_2_DUR: duration of σ_1 and σ_2 (duration in millisecond of each token)
(11)	σ_2_CODA: σ_2 is ended in a glottal stop /ʔ/ or not (*checked syllable* and *smooth syllable*). The glottal stop /ʔ/ is the only stop consonant found in the Wúxī and Shànghǎi dialects to terminate a syllable. Such a stop-closed syllable is called "checked syllable" since it also carries a checked tone. While an open or sonorant-closed syllable is called smooth syllable since it carries a lax tone. In addition, σ_1 discussed in this study is always a smooth syllable as it is in T2, T4 or T6 so only σ_2 has the distinction of "checked syllable" and "smooth syllable".

Among those linguistic predictors for tone sandhi, σ_1_MC_TONE & σ_2_MC_TONE, MC_AB, σ_1_HEIGHT & σ_2_HEIGHT, σ_1_DUR & σ_2_DUR, σ_2_CODA are internal linguistic factors while LEXICAL ATTESTEDNESS and σ_1_PTH_TONE & σ_2_PTH_TONE are external linguistic factors. σ_2_CODA is intrinsically correlated with σ_2_MC_TONE. In the Wú dialects, σ_2_CODA is determined by MC_TONE. If a morpheme has MC_T7 or MC_T8, it ends with a glottal stop in the Wú dialects and thus becomes a checked syllable, the duration of which is much shorter than that of the smooth syllable. Consequently, σ_2_CODA and σ_2_MC_TONE cannot be entered into the regression model jointly. In Chapters 5 and 6, the choice between σ_2_CODA and σ_2_MC_TONE is made based on further discussions.

3.5. Statistical modeling

3.5.1. Drawbacks to VARBRUL

This study aims to investigate tonal variation within the Labovian framework of LVC. The data analyzed within this framework traditionally follows the "principle of multiple causes" and "principle of quantitative modeling" (Young & Bayley, 1996). The "principle of multiple causes" means that linguistic variability is attributed to more than one contextual factor, including both linguistic and social factors. The "principle of quantitative modeling" means that the effects of the contextual factors on linguistic variability can be examined quantitatively with statistical models. Correspondingly, a computer program **vari**a**b**le **rul**e (abbreviated as VARBRUL) was developed by variationists to identify which contextual factors can significantly affect the variability, in what direction, and to what degree.

The mathematical background of VARBRUL was developed by Cedergren and Sankoff (1974) on the basis of the "variable rules" (Labov, 1969) and the theory of probability. The procedure at the heart of VARBRUL is binary logistic regression. The independent variable has to be binary and the dependent variable has to be categorical. VARBRUL has several versions: Goldvarb 2.0 (Rand & Sankoff, 1990), Goldvarb 2001 (Robinson, et al., 2001), Goldvarb X (Sankoff, et al., 2005) and Goldvarb Lion (Sankoff, et al., 2012). Recently, VARBRUL has developed versions implemented in the statistical language R, e.g. R-Varb (Paolillo, 2002) and Rbrul (Johnson, 2009).

VARBRUL has made an important contribution to the development of the quantitative paradigm in sociolinguistics over the past 30 years and has become the key and standard methodological tool of variationist sociolinguistics (Tagliamonte, 2006). However, its popularity has also harmed the field to a certain extent, as (young) scholars have neglected new developments in statistics, and the output is not understood by scholars in other fields of linguistics and social sciences, hampering the dissemination of research results (Johnson, 2009). Furthermore, due to the developments in statistics and increases in computational power, the same underlying statistical techniques (e.g. crosstab, binary logistic regression) are available in well-known statistical toolboxes (SPSS, SAS, R, etc.). In addition, more developed statistical techniques, such as (multinomial) logistic mixed-effects and linear mixed-effects models, are available for analyzing linguistic variation. New models may fit the data better. For example, Tagliamonte & Baayen (2012) introduce mixed-effects models, random forests and conditional inference trees to examine the variation between *was* and *were* in English plural existential constructions. After comparing VARBRUL with the above three models, they found that the mixed-effects model and the random forest provide the closest fits to the data. Their study also concludes that the random forest is more adept at handling unbalanced datasets, while the mixed-effects model prefers balanced datasets.

Turning to this study, which is a sociophonetic study of tonal variation, what is the optimal statistical tool? Variationists have tended to code the independent phonetic variable as a categorical variable on the basis of an impressionistic transcription. Since the development of sociophonetics, the categorical transcription has been criticized from the perspective of phonetics, which has asserted that most current variationist work lacks the analysis of fine phonetic detail when the independent variable is gradient (Jannedy & Hay, 2006). As a sociophonetic study of tone, this work has continuous dependent and independent variables based on instrumental analysis. So this study opts for statistical models that are able to handle continuous variables.

Except Rbrul, the other VARBRUL programs cannot handle either continuous dependent variables or continuous predictors. Rbrul can handle both continuous numeric responses of the independent and dependent variables, but cannot analyze multinomial responses of independent variables (Johnson, 2009). Since the use of VARBRUL is restricted to research in sociolinguistics and second language acquisition, this study will use R (R Core Team, 2013) to construct linear and binary mixed models in order to promote exchanges with tonal studies in other linguistic fields.

Random forest is a recent innovation in statistical classification and regression and thus is less well-known in the field of linguistics. As discussed in Sections 3.2 and 3.4, this study has a relatively balanced dataset. The reading materials for word lists and paragraphs were systematically constructed. So the current balanced dataset allows for using mixed-effects model rather than random forest. Section 3.5.2 focuses on the modeling procedures for mixed effect models.

3.5.2. Mixed modeling procedures

Before delving into the specifics of model building for each variable in Chapters 5 and 6, it is useful to overview the general thoughts on model building here. Two types of mixed models are built. For categorical dependent variables, generalized linear mixed models are built by using the glmer function of the package “lme4” (Bates, et al., 2013). For continuous dependent variables, linear mixed models are built by using the lmer function of the package “lme4” (ibid.). The random factors are SPEAKER (i.e., participants) and WORD (i.e., monosyllabic morphemes or bisyllablic phrases). These two random factors are crossed; in other words, each WORD is presented to each SPEAKER only once. In most cases a word only occurs once, except for a few monosyllabic items in the word lists and minimal pairs.

The summary of mixed model fit by the glmer or lmer function in package “lme4” only provides estimates of the fixed-effects parameters, standard errors for these parameters and a t-ratio, but does not give a degree of freedom and p-value[29]. However, I found *p*-values are useful for understanding findings. To facilitate the discussion of output, the “lmerTest” package is used to calculate p-values for mixed models in this study.

A good regression model is one that is rich enough to describe the data, but does not overfit the data (Agresti 2002: 211). As stated above, five social factors and six linguistic factors are the possible fixed predictors for the tonal variation in citation form; the same five social factors and 11 different linguistic factors are the possible fixed predictors for the tonal variation in sandhi form. To avoid the overfitting problem, not all the factors are put into the final mixed model simultaneously. After the basic exploratory data analyses such as scatterplots for continuous factors and crosstabs for categorical factors, all the fixed factors are investigated by means of likelihood ratio tests to compare two mixed models as a preliminary one by one (Quené, 2012: 15). The first mixed model contains only random factors, while the second model contains one fixed factor and the random factors. These two models are then compared with the ANOVA function. If including the fixed factor yields a better model with significantly smaller residuals than the random-effect-only model, it means this fixed factor significantly improves

[29] The author of package “lme4” provides the reasons for not giving degree of freedom and p values in r-help mailing list: https://stat.ethz.ch/pipermail/r-help/2006-May/094765.html.

the model fitting. Thus this fixed factor is defined as a significant main effect and is included in the final mixed model. After finding the significant fixed factors, their interactions with the other fixed factors are examined. The same procedures are done for all the possible two-way interactions related to the significant main effects selected beforehand. If including a given interaction yields a better model than the model with only significant main effects and random effects, this interaction is included in the final mixed model as well.

In sum, there are two sources of predictors for building the final mixed model: (1) the fixed effect which yields a significantly better model than the model involving only random effects, and (2) the interactions which yield a significantly better model than the model involving only significant main effects and random effects.

Chapter 4. Acoustic measurement of tone in the Wú Dialects

Methodological issues as the definition of the domain of tone and the choice of a tonal normalization procedure had to be researched as a prerequisite for an analysis of the patterns of variation and change. Section 4.1 investigates the domain of tone in the Wú dialects in order to explore from which part of the syllable the F_0 should be extracted for tonal analysis. The section ends with the procedures for tonal segmentation, incorporating some (semi-) utomatic methods for large datasets. Section 4.2 deals first with the F_0 measurement procedure, then a continuous description of tonal contour based on these F_0 points is presented. Section 4.3 discusses the issue of tone normalization and selects the method that serves best our research aims. Section 4.4 investigates the voice quality of the Wú dialects to see if any valid parameters can be found to distinguish breathy voice and modal voice. Finally, a brief summary is provided in the end of this chapter.

4.1. Tonal segmentation: the domain of tone in the Wú dialects

4.1.1. Introduction

In tone languages, the tonal features are restricted to the voiced part of a syllable. But not all voiced parts of the syllable are tone-bearing. Therefore, the concept "domain of tone" is needed to define which entity of the voiced part is associated with the recognition of tone. A similar concept is "tone bearing unit" (TBU), which serves for understanding the autosegmental nature of tone (Bao 1990; Clements & Ford 1979b; Duanmu, 1994; Yip 2002). TBU is an abstract concept used in phonological representation. This study will not discuss this issue at this highly abstract level, but will rather use the term "domain of tone".

Why is it necessary to define the domain of tone? The importance of this issue is illustrated with the example of T2 in Mandarin. It is well-known that Mandarin T2 is a mid-rising tone /35/. According to Zhū (2010: 277), the key of differentiating rising (e.g., /13/) and dipping tones (e.g., /213/) is the position of P_{min}. If P_{min} of a contour, i.e., the turning point, is at around the first 20% of the whole contour, this contour is perceived as a rising tone. If P_{min} is at around 40% of the whole contour, it is perceived as a dipping tone.

Figure 4.1 presents the mean F_0 contour of the four Mandarin tones in the monosyllable /ma/, produced in isolation (reproduced from Xu 1997: 67). The four F_0 contours all consist of /m/ and /a/. The cross symbol "+" on the contour indicates the boundary between /m/ and /a/. The arrow symbol above the Tone 2 contour points to P_{min} of Tone 2.

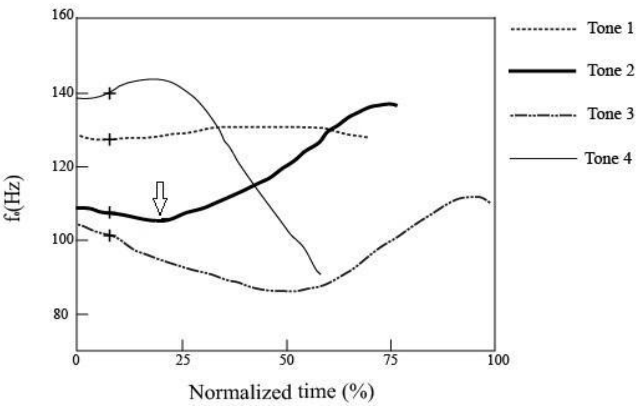


Figure 4.1 The mean F_0 contour of four Mandarin tones in the monosyllable /ma/, produced in isolation (reproduced from Xu 1997: 67)

In Figure 4.1, the time is normalized, with all tones plotted with their average duration proportional to the average duration of T3. "+" marks the boundary between /m/ and /a/. As illustrated in Figure 4.1, the T2 contour ends at 75% normalized time and its P_{min} appears at around 20% normalized time. So the portion before P_{min} accounts for around 27% (20%/75%) of the entire contour. Meanwhile, the portion before P_{min} showed a slightly falling trend. So, its contour shape looks like a dipping curve, though T2 can be perceived as a rising tone according to Zhū's criteria. Whereas, neglecting the initial /m/ (around 8% normalized time), the falling portion before P_{min} is around 18% (18%=(20%-8%)/(75%-8%)), reducing to 9% out of the entire duration, i.e., one third of the falling part. Then the entire contour (neglecting /m/) can reflect the rising trend better as the turning point moved forward. In Standard Mandarin, the 9% difference (from 27% to 18%) might not have any significant consequence. But if it is in a dialect which has a dipping tone changing to a rising one, or conversely, this 9% difference might have more consequences. It is essential for any research involving tonal description to discover, before doing acoustic measurements, which part(s) of the syllable is the domain of tone. A few studies have focused on this issue in Mandarin, but very few attempts have been made on the Wú dialects. This dissertation, as a tonal variationist study of the Wú dialects, has to define the domain of tone in the Wú dialects before any tonal analysis.

To answer this question, the remainder of this section will first review relevant literature of Mandarin on the domain of tone in order to assist seeking answers in the Wú dialects. Before reviewing, it is useful to refresh the Chinese syllable structure. As discussed in Chapter 1, in most Chinese dialects, the maximal syllable consists of CGVX, i.e., consonantal onset (C), a prenuclear onglide (G), the nucleus (V), and a coda (X) (Chen 2000: 4).

Based on impressionistic accounts, Wang (1967), Chao (1968: 19, 25) and Duanmu (2000) claimed that the entire voiced portion of a Mandarin syllable is the domain of tone. The entire voiced portion in Mandarin includes voiced onset (C; /m/, /n/, /l/), prenuclear onglide (G; /i/, /u/, /y/), nucleus (V, monophthong or diphthong) and final nasal consonant (X; /n/, /ŋ/). Conflicting views were presented by Kratochvil (1970), Dow (1972), Howie (1974) and Lín (1995). Kratochvil (1970) argued that the voiced initial consonants should be excluded from the domain of tone while Dow (1972) even proposed to exclude the final nasal consonant. Howie (1974) further examined this issue based on an acoustical investigation of F_0 contours in monosyllables with citation tone. He compared the pitch contours of nine different types of syllables which are all the possible combinations of the four voiced segments listed above. These nine types of syllables are:

1.	V	e.g., yī [i^{55}]	'clothes'
2.	GV	e.g., yā [ia^{55}]	'duck'
3.	GVX	e.g., yāng [iaŋ^{55}]	'center'
4.	C(voiceless fricative)V	e.g., fā [fa^{55}]	'to send out'
5.	C(voiced unaspirated continuant)V	e.g., mā [ma^{55}]	'mother'
6.	C(voiceless aspirated stop)V	e.g., tā [t^{h}a^{55}]	'he/she/it'
7.	C(voiceless unaspirated stop)V	e.g., bī [pi^{55}]	'to compel'
8.	C(voiceless aspirated affricate)V	e.g., cān [ts^{h}an^{55}]	'a meal'
9.	C(voiceless unaspirated affricate)V	e.g., zān [tsan^{55}]	'hairpin'

The high level tone /55/ was used in the examples above. The other three tones of Mandarin were also tested in Howie (1974). Nine types of words were further divided into three groups: (1) Types 1, 6, 7 (initialed with voiceless stops or none); (2) Types 2, 3, 5 (initialed with onglide or nasal consonants); (3) Types 4, 8, 9 (initialed with voiceless affricate). The averaged curves of each type of word for each tone were plotted, resulting in 12 graphs (3 groups × 4 tones). Comparing the 12 graphs, Howie (1974) concluded that the domain of tone in Mandarin is not the entire voiced part of the syllable, but is confined to the nucleus and final nasal consonant. The initial voiced consonant and prenuclear onglide are not included in the domain of tone in Mandarin. The perceptual results of Lín (1995) also support the exclusion of initial voiced consonants or prenuclear onglides in monosyllables. Furthermore, Lín (1995) found that subjects could identify the four tonal categories of Beijing Mandarin correctly on the basis of stimuli containing a nucleus and its adjacent transition. In other words, the nucleus of a given syllable provides the key and sufficient information in the judgment of tonal category whereas other parts of the syllable are less crucial.

Since the Wú dialects have a similar segmental structure to Mandarin, tests will be done to see whether (a) the initial voiced consonant (Section 4.1.2), (b) the prenuclear onglide (Section 4.1.3) and (c) the final nasal consonant (Section 4.1.3) are included in the domain of tone. These tests are restricted to the production data, which is part of the data collected during this project. Inspired by Howie (1974), the shapes of the tonal contours are compared and illustrated quantitatively. Moreover, the duration of different types of syllables is compared as well.

4.1.2. Test on the initial voiced consonants

This section aims to test whether the initial voiced consonant is part of the domain of tone in the Wú dialects. It contains (i) the data and quantitative methods used for comparing tonal shapes are first introduced and (ii) results and conclusions.

4.1.2.1. Data and methods

Syllables with voiced initials in the Wú dialects are all associated with the tonal categories of the Yáng register, i.e., T2, T4, T6 and T8 words. As discussed in Chapters 2 and 3, there are two types of initials in the lower register. One is sonorant (/m/, /n/, /l/, etc.), which is intrinsically voiced and is denoted as T2B, T4B, T6B and T8B; the other is obstruent, which has been proven to be voiceless (Cao & Maddieson 1992; Z. Chén 2010; Ren 1992) and is denoted as T2A, T4A, T6A and T8A. For the CV syllables initialed with voiceless obstruent, the domain of tone is undoubtedly the nucleus (V). The questionable portion is the voiced consonants in the T2B, T4B, T6B and T8B words. Though syllables of the lower register can differ in initials (voiced sonorant vs voiceless obstruent), they carry the same citation tone if they are from the same MC tonal category. Given this condition, the problem is simple because the tonal contours of syllables with voiceless obstruent, i.e., T2A, T4A, T6A and T8A words, can be regarded as a baseline. Tonal contours of T2B, T4B, T6B and T8B words can be compared with this baseline to see whether including or excluding the initial sonorants has more in common with the baseline.

In the Shànghǎi dialect, three lax tones of the Yáng register, i.e., T2, T4 and T6, have merged and become a rising tone of /14/ (Zee & Maddieson 1979); further discussed in Chapter 6). They will be labeled as **SH-T2/4/6** hereafter. Shànghǎi T8, a checked tone with the value /14/ (ibid.) will be labeled as **SH-T8**. In the Wúxī dialect, T2 and T6 have merged together into the dipping shape /213/ (**WX-T2/6**). Wúxī T4 is subject to an ongoing merger with Wúxī T2 and T6 (Xú, 2007). Due to this variation, Wúxī T4 morphemes will not be used to examine the issue of the domain of tone. Wúxī T8 (**WX-T8)** has the value of /13/ (ibid.). To sum up, Section 4.1.2 deals with the four tones listed in Table 4.1.

Table 4.1 Stable tones of the Yáng register in the Shànghǎi and Wúxī urban dialect

tone name	tone value	n	
		Base Set	Set C̬V
SH-T2/4/6	/14/	28	17
SH-T8	/14/	8	7
WX-T2/6	/213/	19	10
WX-T8	/13/	8	7

(Base Set: words with voiceless obstruent onsets; Set C̬V: words initialed with voiced sonorant consonants)

The speech materials analyzed in this section are from 10 native speakers in the Shànghǎi urban area and 10 native speakers in the Wúxī urban area, five males and five females in each. For each speaker, 68 citation CV syllables of the Shànghǎi dialect or 52 citation CV syllables of the Wúxī dialect are analyzed. Rhyming parts of these syllables are all monophthongs. The word list consists of two sets: words with voiceless obstruent initials (Base Set) and words with voiced sonorant initials (m, l; Set C̬V). As was discussed above, the Base Set is the baseline for examining the domain of tone. The exact numbers of tokens for each tone are listed in the column "n" of Table 4.1, subdivided by Base Set and Set C̬V. Syllables of Set C̬V were segmented into two parts: the voiced consonant portion (C̬) and the vowel portion (V_{CV}), i.e., **Set C̬V= Set C̬+ Set** $\mathbf{V_{CV}}$. The F_0 values of Base Set, C̬ and V_{CV} were all measured at 10 equidistant points in semitone, resulting in a time-normalized F_0. Contours of each Base Set were averaged across tones and speakers. In addition, absolute duration was measured in milliseconds in all four sets, providing additional evidence bearing on the domain of tone.

The principal idea of examining the domain of tone of the voiced initial morphemes is to make comparisons between F_0 contours from syllables containing voiced onsets (Set C̥V) and portions from which voiced onsets are removed (Set $V_{C̥V}$) in order to see which set fits the Base Set better. Figure 4.2 illustrates the definition of Base Set, Set C̥V(Set C̥ and Set $V_{C̥V}$).

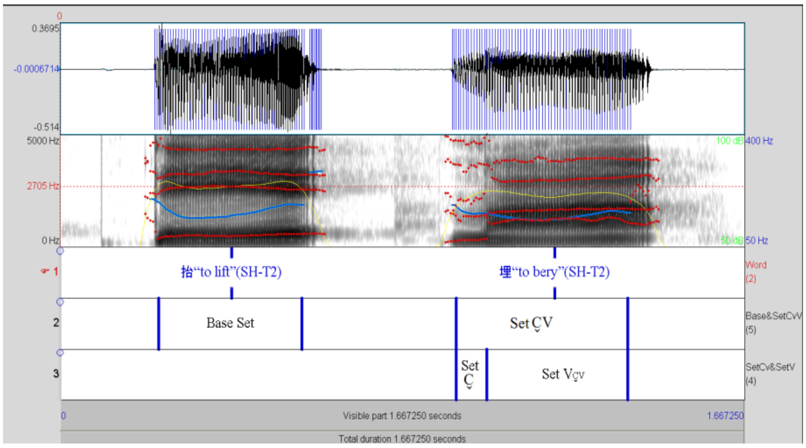


Figure 4.2 Illustrations of Base Set, Set C̥V, Set C̥ and Set $V_{C̥V}$ in Shànghǎi T2 morphemes "[tɦɛ̤] (抬, to lift)" and "[mɦa̤] (埋, to bury)" (recordings from SUof3)

The Root Mean Square Deviation (RMSD) is used to calculate the distance between two contours by summarizing the sizes of residuals. Residuals are defined here as the F_0 point differences between the reference group and test group. The equation of RMSD is:

$$(1) \qquad RMSD(pA, pC̥V/V) = \sqrt{\frac{\sum_{i=1}^{n}(x_{A,i} - x_{C/V,i})^2}{n}}$$

In equation (1), "p_A" denotes points of the Base Set, "$p_{C/V}$" are points of test group Set C̥V or Set $V_{C̥V}$, "$x_{A,i} - x_{C/V,i}$" means residuals and "n" equals 9 in this case.

4.1.2.2. Results

To obtain a general impression of the durational differences between oblique tones (SH-T8 and WX-T8) and lax tones (SH-T2/4/6 and WX-T2/6), the mean values of duration across sets and tones are presented in Table 4.2.

Table 4.2 Mean duration (msec) of the Base Set, Set C̬V, Set C̬ and Set V_{CV}, split by tonal category

tone	Base Set		Set C̬V		Set C̬		Set V_{CV}	
	M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
SH-T2/4/6	277	49	353	51	65	11	288	46
SH-T8	101	30	181	36	80	19	102	27
WX-T2/6	260	39	345	51	73	16	272	47
WX-T8	114	36	185	37	77	23	108	34

A repeated measures ANOVA was conducted to test the effect of excluding voiced initials on the duration of Set C̬V, i.e., Set V_{CV}. The results show that there was no significant difference for duration between the Base Set and Set V_{CV}, $F(1, 16)=3.382$, $p=0.085$. These results suggest that the duration of the voiced onset mainly accounts for the durational differences between words with voiced onsets and words with voiceless onsets.

Besides duration, the pitch contours need to be compared. Duration is a perceptual cue for tone but is not the most crucial one. For example, Mandarin native speakers are able to recognize the same tone when the stimuli are produced with different durations (Hallé, Chang, & Best, 2004; Shen & Lin, 1991). In other words, the F_0 related information is more important than the durational information in tonal perception. So I need to further examine the shape of pitch contours. Table 4.3 compares the RMSD results of Base Set & Set C̬V and Base Set & Set V_{CV}. The smaller RMSD indicates the better least-squares fit.

Table 4.3 Mean RMSD of the Base Set & Set C̬V, Base Set & Set V_{CV} in SH-T2/4/6, SH-T8, WX-T2/6 and WX-T8

tone	Base Set & Set C̬V		Base Set & Set V_{CV}	
	M	**SD**	**M**	**SD**
SH-T2/4/6	6.88	2.54	4.75	1.89
SH-T8	12.12	13.48	10.42	10.49
WX-T2/6	7.54	3.28	6.68	3.77
WX-T8	12.68	7.52	7.90	5.81

A repeated measures ANOVA was conducted to test the effect of excluding voiced initial on the tonal shape of Set V_{CV}. The results show that there were significant differences between "RMSD Base Set & Set C̬V" and "RMSD Base Set & Set V_{CV}", $F(1, 16)=9.153$, $p=0.008$. It shows that portions omitting the voiced onset (Set V_{CV}) in these four tones fit in with the reference group (the Base Set) better. Tone also plays a significant role. In SH-T8 and WX-T8, the RMSD differences between Set C̬V and Set V_{CV} is significantly higher than those of SH-T2/4/6 and WX-T2/6, $F(1, 16)=255.924$, $p=0.000$, which suggests that the voiced initial in checked syllables have more influence on the tonal shape. In sum, these results support the view that the voiced onset should be omitted from the domain of tone.

In sum, in light of evidence for tonal duration and tonal shape, the initial voiced consonants do not carry too much crucial tone information, but interfere with comparisons of duration and tone. Therefore, this dissertation will not treat the voiced sonorants as part of the domain of tone.

4.1.3. Tests on the prenuclear onglide and the final nasal consonant

This section aims to test whether the prenuclear onglide and the final nasal consonant are included in the domain of tone in the Wú dialects. Likewise, the data and methods will be introduced first, followed by the results and conclusion.

4.1.3.1. Data and methods

The tests on the prenuclear onglide and the final nasal consonant were conducted on the bisyllabic words, because all the monosyllables collected in this project are CV syllables, including neither a prenuclear onglide nor a final nasal consonant. The testing tokens are bisyllabic words' σ_1, which is initialed with three lax tones of the Yáng register, i.e., T2, T4 and T6. The speech materials are from 10 native speakers in the Shànghǎi Sōngjiāng and 10 native speakers in the Wúxī urban area, five old men and five old women for each area. Data of Shànghǎi urban area was not used because T2, T4 and T6 morphemes there were all realized as level tones when they were σ_1 in sandhi form. According to the observations on the voiced sonorant (Section 4.1.2), the sonorant portion tends to be level although the following nucleus is rising. This suggests that the voiced portion not carrying tonal information tends to be level; thus level tones are probably not able to reveal RMSD differences when comparing tonal shapes.

For each speaker, three types of syllables were analyzed: (1) syllables that do not have either a prenuclear onglide or a final nasal consonant, but only a voiceless initial and a nucleus. These syllables, 99 tokens for each speaker, constitute the "Base Set". (2) Syllables consisting of voiceless onset, prenuclear onglide and the nucleus. These syllables, 10 tokens for each speaker, constitute the "Set GV", in which "G" refers to the prenuclear onglide, and "V" refers to the nucleus, or "V_{GV}" when appearing alone, hence **Set GV = Set G + Set V_{GV}.** (3) Syllables consisting of voiceless onset, nucleus and final nasal consonant. These syllables, 11 tokens for each speaker, constitute the "Set VX", in which "V" refers to the nucleus, or "V_{VX}"

when appearing alone, "X" refers to the final nasal consonant, so **Set VX = Set V_{VX} + Set X**. Table 4.4 presents the tones and tokens of these three types of syllables.

To read Table 4.4, in the Sōngjiāng dialect, three lax tones of the Yáng register, i.e., T2, T4 and T6, in sandhi form, are labeled as **SJ-σ_1-T2/4/6** hereafter, in which "**SJ**" is Sōngjiāng for short, "σ_1" refers to the first syllable and "**T2/4/6**" indicates the tonal category. SJ-σ_1-T2/4 is realized as rising tone of /13/ while SJ-σ_1-T6 is a level tone, /11/. In the Wúxī dialect, three lax tones of the Yáng register, i.e., T2, T4 and T6, in sandhi form are labeled as **WX-σ_1-T2/4/6**. WX-σ_1-T2 is mostly rising tone /13/ while WX-σ_1-T4/6 are mostly level tones /11/. To sum up, section 4.1.3 will deal with the tokens listed in Table 4.4.

Table 4.4 Stimuli for testing prenuclear onglide and final nasal consonant

tone	**tone value**	**token n**		
		Base Set	**Set GV**	**Set VX**
SJ-σ_1-T2	/13/	28	4	7
SJ-σ_1-T4	/13/ or /11/	37	2	4
SJ-σ_1-T6	/11/	34	5	3
WX-σ_1-T2	/13/	28	4	7
WX-σ_1-T4	/11/	37	2	4
WX-σ_1-T6	/11/	34	5	3

(Base Set: words with only syllabic vowels; Set GV: words with prenuclear onglide and nucleus; Set VX: words with syllabic vowel and final nasal consonant)

Using the same methods as section 4.1.2, both duration and tonal shapes were compared between Base Set and Set GV, Base Set and Set VX. The results are presented in section 4.1.3.2.

4.1.3.2. Results

To gain a general impression of the durational differences, Table 4.5 and Table 4.6 present the mean values of duration across different sets and tones. Table 4.7 compares Base Set and Set GV, Table 4.8 compares Base Set and Set VX.

Table 4.5 Mean duration (msec) of Base Set, Set GV, Set G and Set VGV

tone	**Base Set**		**Set GV**		**Set G**		**Set V_{GV}**	
	M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
SJ-T2/4/6	177	57	227	64	56	22	175	34
WX-T2/4/6	166	47	210	57	60	23	165	21

Table 4.6 Mean duration (msec) of Base Set, Set VX, Set VVX and Set X

tone	**Base Set**		**Set VX**		**Set V_{VX}**		**Set X**	
	M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
SJ-T2/4/6	177	57	208	55	73	27	135	42
WX-T2/4/6	166	47	200	53	75	25	125	40

As can be seen from Table 4.5 and Table 4.6, whether in Sōngjiāng or Wúxī, both Set GV and Set VX had longer duration than the Base Set. First, to examine Set GV (Table 4.5), the duration of prenuclear onglide in Set GV, i.e., Set G accounted for less than one third of the duration of Set GV while the duration of Set V_{GV} was more than two thirds. In duration, Set V_{GV} was the main part of Set GV. A repeated measures ANOVA was conducted to test the effect of excluding prenuclear onglide on the duration of Set GV, i.e., Set V_{GV}. The results show that there is no significant difference between the duration of the Base Set and the duration of Set V_{GV}, $F(1, 16)=3.549$, $p=0.078$. It suggests that the onglide did not affect the duration of the following vowel, in contrast to the duration of the vowel in the Base Set. The duration of the onglide alone can account for all durational differences between the Base Set and Set GV. This supports the view that the domain of tone excludes the onglide.

Turning to Set VX (Table 4.6), the duration of the vowel in Set VX, i.e., Set V_{VX} was around 35% of the duration of Set VX while the duration of the final nasal consonant accounted for around 65%, the main proportion of Set VX. Both in Sōngjiāng and Wúxī, the duration of the vowel in Set VX was significantly shorter than the duration of syllabic vowel in the Base Set, according to the results of repeated measures ANOVA, $F(1,16)=260.054$, $p=0.000$. This means that the nasal consonant is not only the main proportion of Set VX but also can significantly reduce the duration of the preceding vowel. The influence of a nasal ending on the nucleus is different from the role of the onglide. This implies that the final nasal consonant's durational aspects are crucial to the domain of tone.

We will now test whether the prenuclear onglide and the final nasal consonant belong to the domain of tone. Table 4.7 compares the RMSD results of Base Set & Set GV and Base Set & Set V_{GV} in the rising SJ-σ_1-T2 and WX-σ_1-T2.

Table 4.7 Mean RMSD of Base Set & Set GV, Base Set & Set V in SJ-σ_1-T2 and WX-σ_1-T2

tone	**Base Set & Set GV**		**Base Set & Set V_{GV}**	
	M	**SD**	**M**	**SD**
SJ-σ_1-T2	23.71	11.43	22.71	11.43
WX-σ_1-T2	16.40	12.70	20.59	12.42

In Table 4.7, the RMSD values are the calculated means of 10 speakers, listed in the corresponding columns. It is obvious that the RMSDs of "Base Set & Set GV" and "Base Set & Set V_{GV}" in the Sōngjiāng area are very close. In the Wúxī data, RMSD in "Base Set & Set GV" is smaller than that in "Base Set & Set V_{GV}". A repeated measures ANOVA was conducted to test the effect of including or excluding onglide on the RMSD. REGION and SEX are the between-subject factors. The results show that the overall mean RMSD of "Base Set & Set GV" did not differ significantly from mean RMSD of "Base Set & Set V_{GV}", $F(1,16)=.914$, $p=0.065$. But the interaction of onglide and REGION is significant, $F(1,16)=6.661$, $p=0.020$. These analyses indicate that, from the perspective of tonal shapes, including or excluding the onglide in Sōngjiāng does not have an influence on contour shapes. However, in Wúxī, Set GV fits the Base Set better than Set V_{GV}, which means the rising trend is reflected not only in the nucleus but also in the onglide. In other words, the Wúxī data supports the inclusion of the onglide in the domain of tone.

To sum up, although the two dialects point to differing conclusions based on tonal shape data, the present evidence at least does not support the exclusion of the prenuclear onglide. However, the durational evidence does support the exclusion of prenuclear onglides. It appears to be difficult to find a solution that matches both tonal shape and duration. This study suggests keeping the prenuclear onglide in the domain of tone. First, among lax tones, the tonal shape plays a more crucial role in the tonal categorization than duration, so the prenuclear onglide should be included in the domain of tone when it carries the information of tonal shape. Second, to make the decision more practical, the difficulties and efforts in marking the boundary between onglide and nucleus must be considered. To save manpower and time, the same criterion for setting a boundary at the beginning of syllables without onglide was applied to the syllables with onglide.

Turning to the issue of the final nasal consonant, Table 4.8 compares the RMSD results of Base Set & Set V_{VX}.

Table 4.8 Mean RMSD of Base Set & Set VX, Base Set & Set V in SJ-σ_1-T2 and WX-σ_1-T2

tone	Base Set & Set VX		Base Set & Set V_{VX}	
	M	**SD**	**M**	**SD**
SJ-σ_1-T2	23.47	11.48	35.02	12.19
WX-σ_1-T2	17.55	14.65	30.15	12.27

From the data of Table 4.8, it is obvious that Set VX has a smaller RMSD, a better least-squares fit with the Base Set, than Set V_{VX}. A repeated measures ANOVA indicates that mean RMSD of Base Set & Set VX differed significantly from mean RMSD of Base Set & Set V_{VX}, $F(1, 16)=55.047$, $p=0.000$. In comparison with Set V_{VX}, using the tonal points of Set VX resulted in a significant reduction of RMSD. This means that portions containing the final nasal consonant have a better fit with the Base Set. Juxtaposing the findings in RMSD and the durational evidence that a nasal coda can significantly reduce the duration of the preceding vowel, it is concluded that the final nasal consonant should be included in the domain of tone.

In sum, in light of evidence from sandhi form, both the prenuclear onglide and final nasal consonant in the Wú dialects will be measured as part of the domain of tone. Adding the conclusion of Section 4.1.2, among the four voiced parts Ç̬ (voiced sonorant), G (onglide), V (nucleus) and X (final nasal consonants) in the Wú syllables, G, V and X are parts of the domain of tone while Ç̬ is not.

4.1.4. Procedures for tonal segmentation

After deciding on the omission of the voiced sonorant from the domain of tone, this section will discuss the operating procedures and criteria for the segmentation. Before segmentation, some preparatory work is necessary. The first step is data-cleaning. The interactions between interviewer and interviewee in reading tasks, any misread parts before self-correction, and background noises were deleted manually. This step is used to ensure all the voiced portions in the long sound file contain only target tokens for F_0 measurement. The second step is applying the script "marked pause" (Lennes, 2002b) to set boundaries at the beginning and the end of pauses. Those boundaries were manually checked to ensure they contained all the target tokens. Then another script (Lennes, 2002a) was applied to add text (Chinese characters) to the non-pause intervals. The results were also checked manually for the matching between sound tokens and labels. Afterwards the long sound files, labeled with text and token boundaries, were segmented into small sound files bearing one token in each file by the script "chop long sound file" (Welby, 2006). At this point, sound files were ready for segmentation. To speed up the segmentation, the voiced portion of each file was first automatically extracted and labeled as "v" by the script "extract voiced portion" (Tøndering, 2002). For female speakers, the pitch floor is 100 Hz and the pitch ceiling is 350 Hz while for male speakers, the pitch floor is 70 Hz and the pitch ceiling is 220 Hz. Then the boundaries of each voiced portion were manually adjusted based on Zhū's (2010) suggestions. Zhū (2010) pointed out that the intensity of the first glottal pulse is not strong enough for human hearing, so the starting point of a tonal contour is at the second glottal pulse of the nuclear. Meanwhile, the end point should be at the area where amplitudes decline sharply and the formants, especially F2, are vague (Zhū, 2010: 282). Additionally, the end point of a rising/dipping contour is at the maximum F_0 in the spectrogram, while the end point of a falling/peaking contour is where the intervals of glottal pulses become irregular. To facilitate the tonal segmentation, a script was used to open both a sound file and its corresponding TextGrid file from the specified directory at the same time, and then generate a window showing the waveform, a spectrogram of the current sound together with the TextGrid tier. This script is a simplified version of Xu & Prom-on (2012). In the TextGrid tier, the position of boundaries can be adjusted based on the information provided by the vocal pulse marks in the waveform panel and the pitch track plus formant tracks in the spectrogram panel. Then the F_0 values at the defined intervals can be automatically extracted and saved. Section 4.2 will illustrate the F_0 measurement in more detail.

4.2. F_0 measurement and the transcription of tone

After extracting F_0 values for the tonal contour of each token, this section will present the subsequent management of these F_0 values. First I report which F_0 points will be used in the analysis and how to deal with inaccurate F_0 values auto-extracted by Praat (Boersma, 2002). Then a continuous description of tonal contours will be presented, which can serve as input for the continuous dependent variables analyzed in the following chapters and also result in the semi-automatic classification of the tonal shapes.

4.2.1. F_0 measurement of tone

The F_0 values of the tone contour of each word were measured at 11 equidistant points (P_0, P_1… P_{10}), resulting in a set of time-normalized F_0 values. These values were all saved together with their original time scale. Apart from F_0 data, the script also enabled us to extract the duration, the P_{max} and P_{min} from P_1 to P_{10}. Differences in syllable onset voicing are known to affect the beginning of tonal contours (Hombert 1978; Hombert, Ohala, & Ewan 1979a). Though the decision for domain of tone has been made cautiously (c.f. Section 4.1), it did not take into consideration "F_0 perturbations" (Rose 1993) due to coarticulation. Taking this effect into account, the first 10% of the tone contour was neglected (Rose, 1987; Stanford, 2008), so P_0 is excluded and only the F_0 values of P_1 to P_{10} were used for further analysis.

Pitch can be difficult to measure correctly in non-modal voice quality (e.g., breathy, creaky, tense). Hand checks and hand measurements were used to ensure the accuracy of the results. If the difference between the F_0 values of two adjacent windows is exactly one or multiple octaves, it is probably due to an octave error.

Assisted by the author's perception judgments, and the observation of waveform and spectrogram, the error can be fixed by raising the value of the "octave-jump cost" and "octave cost" in the Advanced Pitch Settings of Praat. If the sound has irregular vibratory cycles in non-modal voice, the F_0 points were manually calculated following three steps: (1) calculating the original time of each point based on duration; (2) moving the cursor to that time point and selecting one complete cycle accurately by zooming in on the waveform, thus giving Praat the period in seconds (T); (3) calculating the F_0 value of that period by the equation: $F_0=1/T$.

4.2.2. The transcription of tone

As discussed before, the transcription for both the citation tone and tone sandhi in Chao's tone numerals are the author's impressionistic transcriptions. Since F_0 values have been found, it might be asked why not converting acoustic data directly to Chao's tone numerals, defining each tone level with a precise pitch range in Hz or semitones. The answer is that the previous research results are insufficient to support such conversion. There were several attempts to use acoustic evidence to quantify the transcriptions of five tone levels, developing into at least three approaches. (1) Vance (1977) and Du (1988) map the tone levels directly to the F_0; each tone level is spaced equal F_0 intervals for the entire tone range. (2) Chao's earlier description of the tone levels (1956) uses equi-exponential intervals, in which five tone levels are spaced two semitones apart. Another version of the equi-exponential approach (Shi 1986; Zhu 1999; Xu, 2007; etc.) maps five tone levels into the entire tone range; each tone level takes up one fifth of the range. (3) A later description of Chao (1968) provides non-equi-exponential intervals for tone levels. There is only one semitone between levels 1 and 2 and three semitones between levels 2 and 3. This modulated version may be tailor-made for Mandarin (Fon & Chiang, 1999). Recently, Lǐng & Lí (2009) investigated by means of perception experiments how the tone level of Cantonese is distributed. It finds that Cantonese uses non-equi-exponential intervals for tone level as well, in which level 2 takes up more than 40% of the entire space. In sum, all the studies above show that there has not been consensus in quantifying the transcriptions of five tone levels. Since the quantification might be language specific, the preliminary work of defining the pitch range of each tone level in a given language is to conduct several carefully designed perception experiments. The current study was not designed to solve these problems.

However, as a sociophonetic study, the transcription of its acoustic data has to meet the precision requirement of phonetics. It has been proven that polynomial equations are effective for synthesizing pitch contours, as in Zhu (1999), Sūn & Shèng (2001) and Andruski & Costello (2004). On the basis of previous research, there is good reason to believe that polynomial equations are an appropriate tool for describing tones quantitatively. On the one hand, the polynomial equation can provide an accurate and unique representation of each F_0 contour; on the other hand, it can further predict the tonal shapes. Tonal shapes are important to investigate because they are the basic category by which to differentiate tones, at least the Wú tones. Moreover, tones in different shapes have different crucial F_0 points for comparing pitch height.

The degree of polynomial equation depends on the complexity of tonal contours. Liú (2005) explicitly reviewed the tonal types of different Chinese dialects and found the Wú dialects have the following seven shapes: null (e.g., /3/, especially for abrupt tone), level (e.g., /33/), rising (e.g., /13/), falling (e.g., /31/), horn-shaped (e.g., /311/ or /113/), peaking (e.g., /131/) and dipping (e.g., /313/) tones. So, the tonal contour in Wú has maximally one turning point (for the horn-shaped, peaking or dipping curves). So the parabola $y=a+bx+cx^2$, the 2nd degree of polynomial equation, is the appropriate fitting curve for the Wú tones. F_0 or normalized values of P_1 to P_{10} are plotted on the Y-axis, the time point on the X-axis.

The quadratic equation has two functions. First, a-, b- and c-coefficients obtained from the equation can serve as continuous independent variables when examining tonal variations. For instance, the absolute value of the c-coefficient indicates the width of the parabola (mouth) stretch or compression, i.e., the degree of convexity or concavity in a peaking or dipping tone; the b-coefficient, the linear coefficient, indicates the slope of the tangent to the pitch contour at the y-intercept, i.e., the degree of obliqueness in linear tones; while the a-coefficient indicates the y-intercept.

Second, a-, b- and c-coefficients can indicate the tonal shapes approximately. When $y=a+bx+cx^2$ (c=0), the line is straight. In the actual quadratic fit of tone curves, it is almost impossible to have c=0. However, for the completeness of the description, these possibilities are also discussed. When c=0, the b-coefficient indicates its slope. If b=0, the line shape is level. If b is greater than 0, the line shape is rising. If b is negative, the line shape is falling.

For a quadratic equation $y=a+bx+cx^2$ ($c\neq0$), the position of the vertex is crucial for determining the curve shape. The vertex is the highest or lowest point of the parabola, depending on whether the parabola opens upward ($c>0$) or downward ($c<0$). The x-coordinate of the vertex can be obtained from $-b/2c$. In the pitch modeling, the horizontal X-axis represents time, thus the range of x in the equation $f(x)=a+bx+cx^2$ needs to be greater than or equal to the time point of P_1 (T_{P1}) but less than or equal to the time point of P_{10} (T_{P10}), i.e., $T_{P1}\leq x \leq T_{P10}$. If the x-coordinate of the vertex is within this range, it indicates that the pitch curve contains a vertex, thus its shape is dipping when the parabola opens upward or peaking when the parabola opens downward. If the vertex is less than T_{P1} or greater than T_{P10}, the pitch curve is either rising or falling, determined together by the a- and b-coefficients. In addition to coefficients, the value of R^2 is also calculated and saved, in order to evaluate the goodness of fit. Table 4.9 summarizes the relationships between the pitch curves and the different combinations of a-, b- and c-coefficients. A MATLAB program was written to do the curve-fitting and tonal shape judgment, based on Table 4.9.

Table 4.9 The relationship between the values of the a-, b- and c-coefficients and resulting curve shapes

coefficients of $y=a+bx+cx^2$	**curve shape**
$c=0 \cap b>0$	rising
$c=0 \cap b=0$	level
$c=0 \cap b<0$	falling
$c>0 \cap X_{vertex} < T_{P1}$	rising
$c>0 \cap T_{P1}\leq X_{vertex} \leq T_{P10}$	falling then rising (dipping)
$c>0 \cap X_{vertex} > T_{P10}$	falling
$c<0 \cap X_{vertex} < T_{P1}$	falling
$c<0 \cap T_{P1}\leq X_{vertex} \leq T_{P10}$	rising then falling (peaking)
$c<0 \cap X_{vertex} > T_{P10}$	rising

X_{vertex} is the x-coordinate of the vertex on the parabolas, which is in real time scale and can be obtained from $-b/2c$. Similarly, T_{P1} is the time value of P_1 and T_{P10} is the time value of P_{10}.

The curve shapes indicated by the quadratic equation cannot exactly model human perception because, on the one hand, fitting curves can represent the actual data only when the goodness of fit is high; on the other hand, slight skewedness can be distinguished by curve fitting but cannot be distinguished by human perception. For instance, the minimal perceptible interval for various level tones is one semitone (Clements & Ford, 1979a; Shi, Shi, & Liao, 1987). So the author, as a native speaker of the Wúxī dialect with near-native speaker fluency of the Shànghǎi dialect, transcribed the tonal shapes based on auditory perception. Then the author's transcription of tonal shape was compared with its automatic recognition. Discrepancies between auditory results and automatic recognition were further checked with the visual assistance of F_0 contours presented on a spectrogram. The manual check shows the classification accuracy of tonal shapes is nearly 80% in the citation tone. Corrections were made mainly for the level and dipping tones. The reason for the issues with level tone was indicated earlier in this paragraph. In regard to the rising and dipping tones, as stated in the beginning of this chapter, a contour having its turning point at around 20% into the entire duration can still be perceived as a rising tone. But such a contour is certainly classified as a dipping tone according to the algorithms in Table 4.9. Many manual corrections were needed for these curves that were automatically recognized as dipping shapes.

In conclusion, using the quadratic fitting curves as the continuous transcription of tonal contours in the Wú dialects is multifunctional. First, the automatic classification results can be a reference for the auditory transcription, thus enhancing the accuracy and speed of auditory transcription. Second, a-, b- and c-coefficients obtained from the quadratic equation can be used as continuous dependent variables for studying the tone normalization procedure (details in Section 4.3.2).

Now I have defined the domain of tone, and shown how to obtain F_0 data and transcribed the tonal contour continuously. The next step in the acoustic analysis concerns F_0 normalization, which is necessary for comparing tones of speakers who differ in the size of their vocal folds.

4.3. F_0 normalization

4.3.1. Introduction

In comparison with the large amount of work done on vowel normalization (see van der Harst 2011: 90-107 for an overview), tone normalization, or more specifically, the normalization of the fundamental frequency associated with linguistic tone, has received relatively little attention (Rose, 1987). In variationist studies, an effective technique for the normalization of vowels is one that (1) preserves phonemic variation, (2) minimizes anatomical variation and (3) preserves sociolinguistic variation (Adank, 2003; Adank, et al., 2004; van der Harst, 2011). These criteria can be applied to tone normalization as well. Tone normalization aims at eliminating the anatomical variation between speakers and at allowing between-speaker comparison. The anatomical variation resides in the frequency variation mainly caused by different sizes of the vocal folds. Laryngological studies, e.g., Chatterjee, Halder, Bari, Kumar, & Roychoushury (2011) revealed that longer vocal folds (i.e., typically found in men) result in slower vibrations, thus causing lower formant frequencies (including lower F_0). Such anatomy-based F_0 variation can also be found in elderly speakers, approximately 60 years old and above. Male voices go up in pitch with advanced age due to shrinking vocal folds, while female voices can either go down due to hormonal changes or go up due to atrophy of vocal folds after age 60 (Chatterjee, et al., 2011). Although the variation of vocal fold sizes is the source of the variation of fundamental frequencies, listeners are nevertheless able to neutralize such anatomy-based acoustic differences while retaining phonemic differences (Adank, et al., 2004; van der Harst, 2011; Rose, 1987). Hence, in order to accurately model the perception of tones by the human auditory system, these anatomy-based differences need to be removed by a proper normalization procedure.

Previous vowel normalization studies suggested that different normalization techniques serve different goals (Disner, 1980; Thomas, 2002; Thomas, 2011). So before going into details of tone normalization techniques, the above-mentioned criteria for evaluating normalization techniques are first reviewed in Section 4.3.2. A method of comparison used successfully in vowel normalization is adapted for tone normalization and introduced in Section 4.3.2 as well. In Section 4.3.3, 16 normalization procedures, ten existing procedures and six variants, are briefly described and compared. Section 4.3.4 presents our results and concludes with which normalization procedure is the best for a sociophonetic study of tone variation. That procedure is used to normalize all F_0 values in this study.

4.3.2. Evaluation of tone normalization methods

4.3.2.1. Previous evaluations on tone normalization methods

To the author's knowledge, Rose (1987) is the first study that compared tone normalization methods quantitatively. Rose (1987) evaluated two methods, Z-score and Fraction of Range (FOR) normalization by the method Dispersion Coefficient (DC, equation 2) on data of the Zhènhǎi dialect, a variety of Wú. The DC is the ratio of mean between-speaker variance to overall sample variance, and is a measure of the degree to which speakers' values cluster. Rose (1987) found that Z-score is preferable since it has a smaller DC, which indicates a better convergence result.

(2) $$DC = \frac{mean\ between{-}speaker\ and\ within{-}normalization{-}point\ variance}{sample\ variance}$$

Zhu (1999) followed Rose's study and further developed four normalization methods on the basis of Z-score and FOR, i.e., proportion of range (POR), ratio of log semitone distances (LD), logarithmic Z-score (LZ) and logarithmic proportion of range (LPOR). Zhu (1999) compared six methods on the basis of a Normalization Index (NI) on data of the Shànghǎi dialect; the LZ method comes out as the best method. Before describing these six methods, the method of comparison will be reviewed first as it is the key for the selection of methods.

NI is defined as a ratio between DC of F_0 values before normalization and DC of normalized values (Rose, 1987; Zhu 1999: 49, equation 3).

(3) $$NI = \frac{DC\ of\ F0\ values\ before\ normalization}{DC\ of\ normalized\ values}$$

A large NI indicates that normalized F_0 contours cluster tightly. If tonal studies, e.g., Rose 1987 and Zhu 1999, prefer normalization methods with larger NIs, it implies that their primary aim of normalization procedure is to categorize tones. High NI priority implies that a good normalization procedure can eliminate phonetic differences if they are not phonemic contrasts. In sum, the NI index serves for studies with the aim of drawing up the tone inventory of a given language; the method with a larger NI is preferred.

However, tonal variationist study not only aims at categorizing tones but also aims at recording the intermediate status between points on some scale between phonological and phonetic. If a tone is involved in a phonetically gradual change in progress, a proper normalization method for variationist study should keep the phonetic differences that are eliminated by high-NI normalization methods. Meanwhile, the F_0 of a tone carries not only phonemic and anatomical information, but also a great deal of sociolinguistic information, like speaker's regional background, socioeconomic class and ethnicity, etc. In such a case, NI cannot evaluate the success of normalization methods in preserving sociolinguistic information, which is its fatal flaw. Therefore, NI is not enough qualified to evaluate and then select the normalization methods for variationist studies.

4.3.2.2. Using discriminant analysis to evaluate tone normalization methods

Recent studies of vowel normalization (Adank, 2003; Adank, et al., 2004; van der Harst, 2011) used linear discriminant analysis (LDA) to compare different vowel normalization procedures to select the most successful one for fulfilling three criteria: (1) preserving phonemic variation, (2) minimizing anatomical variation and (3) preserving sociolinguistic variation. LDA is a standard technique for data classification. It maximizes the ratio of between-class variance to the within-class variance in order to guarantee the maximal separability. LDA generates models on the basis of the group variable and the independent continuous predictor(s). It can predict the group membership of each token from the model and then calculate the model's accuracy of prediction for the actual classification. For example, in a model testing how well the vowel category can be predicted on the basis of values resulting from the different normalization methods, the dependent variable is the actual phonemic category of vowel while the predictors are the values of F_0, F_1, F_2 and F_3 at 25%, 50% or 75% (van der Harst 2011: 108-109). Other details of the comparison method are also based on Adank (2003), Adank, Smits, & Van Hout (2004) and van der Harst (2011), which will be illustrated when reporting the results in Section 4.3.4.

The methodological differences of the present study from those vowel studies are the predictor(s) of LDA. Vowel studies use values of F_0, F_1, F_2 and F_3 at a specific point during the vowel, since those formant values are crucial to the classification of vowel categories and they are independent indexes. The main goal of tonal LDA is to predict tonal categories. To differentiate tones, information like pitch height, direction of pitch movement (contour shape) and duration are crucial. It is easy to imagine using pitch height values of P_1 to $P_{10,}$ i.e., ten equidistant points on the tonal contour, to represent not only the pitch height, but also the pitch shape. However, one of the LDA assumptions is that its predictors have to be independent. Pitch height values on the same tonal contour are strongly correlated with each other so they cannot be predictors at the same time. Thus I have to transform the information conveyed by P_1 to P_{10} into independent predictors, including their pitch heights, contour shape and the duration of tone. As discussed in Section 4.2, a-, b- and c-coefficients obtained from the quadratic equation can be used as continuant independent variables for studying the tonal variation. The quadratic coefficients are independent of each other as each of them represents one characteristic of the parabola (see Section 4.2.2). They can provide the information of pitch height as well as the contour shapes, so they all can be used as predictors for LDA. Another predictor is the duration of tone. A stepwise LDA will be performed to select the best predictor(s) among a-, b- and c-coefficients and duration.

Another difference between this tonal study and the vowel studies of Adank (2003), Adank et al. (2004) and van der Harst (2011) is that this study uses discrete indexes to quantify anatomical variation and sociolinguistic variation. When using LDA to evaluate how the normalization fulfills three aims of (1) preserving phonemic variation, (2) minimizing anatomical variation and (3) preserving sociolinguistic variation, the first step is to find proper indexes of phonemic, anatomical and sociolinguistic variation in the testing data. For the phonemic variation, the differences between vowel studies and tone studies are evident. Vowel studies use vowel categories as the dependent variable while the tone studies use tonal categories. It is noteworthy here that the level tones cannot be correctly predicted by the coefficients of the quadratic equation because, as discussed above, all the level tones are transcribed by the author as level tone but their predicted pitch contours, which is based on the classification of Table 4.9 can be of rising, falling, dipping or peaking contours, or even irregular curves. Therefore, the level tones will be excluded in the testing data when doing LDA analysis.

Vowel studies have used sex-related variation to represent anatomical variation and regional variation to represent sociolinguistic variation. If a normalization procedure can successfully remove the anatomical variation, the success rate for that procedure would be at chance level (i.e. 50%) to predict speaker's sex; if a normalization procedure can successfully preserve the sociolinguistic variation, the success rate for that procedure would be close to 100% to predict speaker's region. The fundamental premise of these evaluations is in the speech data of a corpus of 160 speakers of Dutch covering eight regional varieties of Standard Dutch, the sex-related variation is mainly anatomical rather than sociolinguistic, on the basis of the agreement of transcription (Adank 2003: 138), while there is obvious regional variation among eight data points (van der Harst 2011: 255-266). In the present research, can sex- and age- related variation represent anatomical variation and regional variation represent sociolinguistic variation as well? These premises have to be tested before LDA analysis.

The regional variation is the variation of phonemes across regions. Although data from six different regions and a comparable phoneme (i.e., the low rising tone in T2, T4 or T6) are available in this study, it was still unclear whether the regional variation exists before I made my own comparisons (in Chapters 5 & 6). Unlike the Dutch vowel studies, no systematic comparison is available on these low rising tones in the literature to date. The low rising tones were either transcribed as /13/ or /113/ in the dialectological studies, not showing obvious regional variation. So it is not appropriate to use regional variation to index sociolinguistic variation and another variable is needed. In the present study, the age cohort effect (old vs young) is used to trace the tonal variation. If ongoing changes exist, the old generation normally uses the conservative variant while the young generation uses the advanced variant; so age-related variation can index sociolinguistic variation. Though age-variation contains possible anatomical variation, as mentioned in Section 4.3.1, the anatomical F_0 changes caused by aging are mainly changes of pitch heights, so the variation of contour shapes constrained by AGE, rather than the variation of pitch height, is a reliable index of sociolinguistic variation.

When treating sex-related variation as anatomical variation, it needs to be sure that there is only anatomy-based variation between males and females in the testing data but no significant sociolinguistic variation. Like the anatomically age-related variation, the anatomically sex-related variation on F_0 is also the variation of pitch height, not associated with contour shape. So if there is sex-related variation of contour shapes, this type of variation cannot be purely anatomical and cannot be used as an index of anatomical variation.

In sum, an ideal dataset to test the ability of preserving sociolinguistic variation and minimizing anatomical variation in the tonal study should at least meet the following two requirements: (1) showing no variation of pitch shapes on the basis of transcription between males and females and (2) having variation of pitch shapes between the old and young generations. Such a variable is sought in the dataset of six regions.

4.3.2.3. Selecting and evaluating the dataset for the LDA analysis

A series of mixed models were conducted on the slope of rising tones of T2, T4 and T6 in each region. The independent variables were AGE and SEX. In the rising tones, the slope is an indicator of pitch shape. If I can find one or more regions in which there is a significant slope difference between old and young people, but no significant slope difference between males and females, it means that such kind of data meets the two requirements above and will be further tested in a perception test. The slope m_{rising} is calculated by the equation (4):

(4) $$m_{rising} = \frac{ST_{Pmax} - ST_{Pmin}}{NT_{Pmax} - NNT_{Pmin}}$$

In equation (4), ST_{Pmax} is the semitone value of maximum F_0 on the pitch contour, while ST_{Pmin} is the semitone value of minimum F_0 on the pitch contour. NT_{Pmax} is the time point of maximum F_0 while NT_{Pmin} is the time point of minimum F_0. m_{rising} is calculated for each token transcribed as rising tone in T2, T4 and T6 for all six regions. Table 4.10 lists the mean m_{rising} split by AGE and SEX, as well as the results of mixed-effected regression. The dependent variable is the m_{rising} of each region, which is numerical. The independent variables are AGE (0=young, 1=old) and SEX (0=female, 1=male), with SPEAKER (20 levels for each region) and WORD (33 levels in Wúxī urban and Huàzhuāng, 52 levels in Shànghǎi urban, 36 levels in Sōngjiāng, Bǎoshān and Nánhuì) as random effects.

Table 4.10 Mean m_{rising} and mixed models on m_{rising} in six regions

region	mean slope (ST/s) old male	old female	young male	young female	mixed-effects regression
T2 & T6 Wúxī Urban	23.87	26.12	11.33	13.91	AGE: t=5.41, p=0.031* SEX: t=−0.71, p=0.489 AGE*SEX: t=0.10, p=0.926
T2 & T6 Wúxī Huàzhuāng	29.04	17.37	13.32	13.38	AGE: t=0.733, p=0.474 SEX: t=−0.395, p=0.698 AGE*SEX: t=2.42, p=0.028*
T2, T4, & T6 Shànghǎi Urban	30.64	18.75	31.41	21.34	AGE: t=−0.576, p=0.573 SEX: t=2.158, p=0.046* AGE*SEX: t=0.352, p=0.730
T4 &T6 Shànghǎi Sōngjiāng	24.79	28.59	15.41	17.35	AGE: t=3.62, p=0.002** SEX: t=0.122, p=0.904 AGE*SEX: t=−0.465, p=0.647
T4 & T6 Shànghǎi Bǎoshān	33.93	22.83	6.13	16.15	AGE: t=1.57, p=0.135 SEX: t=−2.606, p=0.019* AGE*SEX: t=3.667, p=0.002**
T2 & T4 Shànghǎi Nánhuì	22.67	22.11	15.27	11.91	AGE: t=2.993, p=0.007** SEX: t=1.031, p=0.315 AGE*SEX: t=−0.621, p=0.542

Table 4.10 shows that, in Wúxī Urban, Shànghǎi Sōngjiāng and Shànghǎi Nánhuì, there is not any significant effect of SEX on the slope variation, either as a main effect or in the interaction. It means that these three regions have purely anatomical variation and thus meet the first requirement. There is a significant effect of AGE on the slope variation, Wúxī Urban, Wúxī Huàzhuāng, Shànghǎi Sōngjiāng and Shànghǎi Bǎoshān. In other words, the data of these four regions meets the second requirement that age-related variation can represent sociolinguistic variation. Only two regions, Wúxī urban and Shànghǎi Sōngjiāng, meet both requirements. The data of them was chosen to test the tone normalization methods.

The analyses of Table 4.10 used the semitone scale centered to 100 Hz to calculate slope. However, the semitone scale centered to 100 Hz is also a normalization procedure which causes the results of selecting best normalization procedure depend on one of the normalization procedure itself. So further perception tests are needed in order to avoid the problem of circularity. The slope variation of rising tone is a within-category variation, so five trained phoneticians who are specialized in tone or intonation research were asked to judge the steepness of rising tokens used in Wúxī urban and Shànghǎi Sōngjiāng.

In the word list reading of monosyllables, five rising tokens were randomly selected for each speaker, making 100 tokens in total for each region. Because the 20 speakers of each region were stratified for AGE (old vs young) and SEX (male vs female), the 100 rising tokens were stratified for AGE and SEX, resulting in 25 rising tokens in each cell. The phoneticians listened to the randomly ordered recordings one by one and made judgments on their rising scales from one (level) to four (steepest). Recordings of two morphemes in scale 1 and scale 4 were given as examples respectively before the test. Examples in scale one sounded like level tones, with m_{rising} around 7 ST/s. Examples in scale four were typical rising tones, with the two greatest m_{rising} values in each dataset, whose m_{rising} were around 60 ST/s.

Then the homogeneity or internal consistency of the judgments was further checked by Cronbach's alpha (Rietveld & Van Hout, 1993: 187-228). The Cronbach's alpha of the five raters was 0.853, which is high. On the basis of Cronbach's alpha, it was decided to keep all five transcribers. Then the correlations between each pair of transcribers were tested to see if the mean score of five transcribers is representative. As our scale for steepness is ordinal, both Pearson Correlation (range 0.37−0.73) and nonparametric correlation Kendall's tau_b (range 0.33−0.62) were tested on each pair of phoneticians. The results of the correlation coefficients were not as high as expected, but were acceptable. Given the results of Cronbach's alpha and correlations, I decided to calculate the mean score of the five transcribers for each stimulus and use it in the GLM analysis.

Table 4.11 presents the transcribed mean score of steepness split by REGION, AGE and SEX. A one-way within subjects ANOVA was further conducted to compare the effect of REGION (*Wúxī Urban* vs *Shànghǎi Sōngjiāng*), AGE (*old* vs *young*) and SEX (*male* vs *female*), and their interactions on the transcribed mean score of steepness.

Table 4.11 Mean scores of steepness split by REGION, AGE and SEX

REGION	**AGE**	**SEX**	**M**	**SD**	**n**
Wúxī Urban	old	male	3.42	0.59	25
		female	3.50	0.46	25
	young	male	2.77	0.65	25
		female	2.65	0.62	25
Shànghǎi Sōngjiāng	old	male	3.61	0.32	25
		female	3.83	0.22	25
	young	male	2.62	0.69	25
		female	2.45	0.77	25

The results of the one-way ANOVA showed that only AGE has a significant effect on the mean score (Wúxī: $F(1,192)=144.647$, $p=0.000$; Shànghǎi: $F(1,192)=7.471$, $p=0.007$). This means that the phoneticians perceived significant age-related variations, in which the old people use steeper rising tones and the young people use flatter rising tones. The effect of SEX was not significant, neither as a main effect, nor in the interaction with age ($p>0.05$). This indicates that, in both Wúxī urban and Shànghǎi Sōngjiāng, there was no significant sex-related difference in the slope steepness based on the judgments of the five phoneticians. In sum, the results of the perception task support the hypothesis that there is perceptible age-related variation in the realization of the steepness of rising tones, indicating sociolinguistic variation. Meanwhile, there is not much sex-related variation in the realization of the steepness of rising tone, indicating that the sex-related variation in the data is anatomically-based.

To conclude Section 4.3.2, an LDA analysis is conducted to evaluate different normalization methods; a-, b- and c-coefficients obtained from the quadratic equation, and tonal duration are used as continuant independent variables in the LDA analyses; dependent variables are chosen on the basis of different normalization aims: (1) tonal categories are used as the dependent variable to evaluate how well the normalization methods preserving phonemic variation; (2) according to the perception results, age-related variation is used as the dependent variable to evaluate how well the normalization methods preserve sociolinguistic variation and (3) sex-related variation is used as the dependent variable to evaluate how well the normalization methods minimize anatomically-based variation. Section 4.3.3 will give a brief description of the normalization methods to be compared. Section 4.3.4 will present the results of the comparison.

4.3.3. Description of tone normalization methods

As mentioned in Section 4.3.1, six tone normalizations were systematically described and compared in Zhu (1999: 45-48). To facilitate the understanding of these six methods, their equitation are cited with a brief description in this study as well. Those six methods are numbered as Methods 1, 2, 3, 4, 6 and 7 in this study.

In the following, equation (4) to equation (10), raw values of F_0 in hertz are represented by x_i, where i can take the value 1 to 10 for the measuring points P_1 to P_{10}. $F_0^{Method\ X}$ of each equation stands for normalized value via "Method X"; using the superscript identifies the name of normalization method.

Method 1: Z-score

Z-score values, $F_0^{Z-score}$, express "an observed F_0 value as a multiple of a measure of dispersion away from a mean F_0 value" (Rose 1987: 347). Z-score is calculated as follows:

(5) $$F_0^{Z-score} = \frac{x_i - m}{s}$$

In equation (5), m is the mean value of x_i, and s is the standard deviation, both calculated per speaker.

Method 2: Fraction of range (FOR)

F_0^{FOR} expresses "an observed F_0 value as a fraction of the difference between two range-defining F_0 values" (Rose 1987: 347). The normalized F_0 value is calculated by equation (6)

(6) $$F_0^{FOR} = \frac{x_i - x_L}{x_H - x_L}$$

Zhu (1999: 49) defined, in the Shànghǎi dialect, x_H as the highest F_0 value of T1 and x_L as the lowest F_0 value of T2/T4/T6 for each speaker. Because the highest F_0 value of T1 is usually the speaker's highest F_0 value, while the lowest point of T2/T4/T6 is usually the lowest point of all tones (Zhu 1999: 49).

Method 3: Proportion of range (POR)

F_0^{POR} also expresses F_0 value as a proportion of a range, being expressed by mean (m) and standard deviation (s). The equation is:

(7) $$F_0^{POR} = \frac{x_i - (m - cs)}{(m + cs) - (m - cs)} \ (c\ is\ a\ consonant)$$

Zhu (1999) used c=1 and c=2 in calculating POR. Same with Z-score, m is the mean value of x_i, both m and s are calculated by speaker.

Method 4: Ratio of log semitone distances (LD) & Method 5: T value

F_0^{LD} is the logarithmic version of Method 2: FOR. LD is calculated with equation (8). The x_H and x_L of LD is the same as in FOR.

(8) $$F_0^{LD} = \frac{log_{10}^{x_i} - log_{10}^{x_L}}{log_{10}^{x_H} - log_{10}^{x_L}}$$

Shi & Wang (2006) developed an adapted version of LD, which is called T value (F_0^T). The T value has been widely used in the Chinese literature on tone, more than any other normalization techniques, as T values range from 1 to 5 and match Chao's 5-point scale. The T value is calculated as:

(9) $$F_0^T = 5 \times \frac{log_{10}^{x_i} - log_{10}^{x_{min}}}{log_{10}^{x_{max}} - log_{10}^{x_{min}}}$$

In equation (8), x_{max} is the highest F_0 value of one speaker rather than the constant P_{max}; x_{min} is the lowest F_0 value, also different from the constant P_{min}. T value uses each speaker's two extreme F_0 values in order to make sure the result is between 0 and 5. T values between 0 and 0.99 were converted to Chao's tone letter 1, values in the range of 1 to 1.99 were Chao's tone letter 2, and so on, and thus T values between 0 and 5 can correspond to 5-point scale.

Method 6: Logarithmic Z-score (LZ)

F_0^{LZ} transform is the logarithmic version of the Z-score. The equation for calculating F_0^{LZ} is:

(10) $$F_0^{LZ} = \frac{y_i - m_y}{s_y} = \frac{log_{10}^{x_i} - \frac{1}{n} \times \sum_{i=1}^{n} log_{10}^{x_i}}{\sqrt{\frac{1}{n-1} \times \sum_{i=1}^{n} \left(log_{10}^{x_i} - \frac{1}{n} \times \sum_{i=1}^{n} log_{10}^{x_i}\right)^2}}$$

As shown in equation (10), $y_i = log_{10}^{xi}$, m_y and s_y are the mean and SD of y_i (i= 1, 2, 3, … 10) respectively.

Method 7: Logarithmic proportion of range (LPOR)

F_0^{LPOR} is the logarithmic version of POR. Therefore, it is calculated using equation (11):

(11) $$F_0^{LPOR} = \frac{y_i - (m_y - cs_y)}{(m_y + cs_y) - (m_y - cs_y)}$$

y_i, m_y and s_y in equation (11) are the same as in equation (10), $y_i = log_{10}^{xi}$, m_y and s_y are the mean and standard deviation of y_i. Like in POR, c is a consonant in LPOR.

F_0^{FOR}, F_0^{LD} and F_0^T value transform the original F_0 to a value relative to the range between x_H and x_L or between x_{max} and x_{min}, two range-defining (R-D) points. It is worth noting that R-D points should avoid tonal variation; otherwise the range could vary due to the tonal variation in addition to the anatomical differences. In addition, Rose (1987) pointed out that using this method should avoid the circularity of forcing congruence, as R-D points are assumed to be equivalent among the speakers. But their equivalence will only be clear after normalization if no external criterion is applied for evaluation beforehand. One possible external criterion is that R-D points are speaker-constants by virtue of their low within-speaker variance. In sum, the definition of R-D points is crucial. Their selection should meet the following requirements: (1) they must have consistent linguistic meanings across speakers to avoid between-speaker variations and (2) they must be speaker-constant values to avoid within-speaker variations.

In considering this issue in the Wú dialects, Zhu (1999: 49) proposed using the lowest F_0 value of T2/T4/T6 in the Shànghǎi urban dialect as the speaker-constant P_{min}, and the highest F_0 value of T1 in the Shànghǎi urban dialect as the speaker-constant P_{max}, which has relatively small standard deviations within each speaker. The other five data points (Shànghǎi Bǎoshān, Sōngjiāng and Nánhuì, Wúxī urban area and Huàzhuāng), like the Shànghǎi urban dialect, have at least one tone among T2, T4 and T6 containing a speaker-constant P_{min} and in T1 containing a speaker-constant P_{max}.

Apart from the six strategies reviewed in Zhu (1999), a number of tonal studies have used semitone scales to match human pitch perception. The basic interval for pitch perception is the octave, as the human auditory system perceives tones in a logarithmic way rather than a linear way. Equation (12) is used to transform hertz into semitone.

(12) $$F_0^{ST-ref} = \frac{12}{log_{10} 2} \times log_{10} \frac{x_i}{ref} \approx 39.87 \times log_{10} \frac{x_i}{ref}$$

Method 8: ST-100	ref=100 Hz
Method 9: ST-AvgF$_0$	ref=AvgF$_0$, i.e., each speaker's average pitch
Method 10: ST-x$_L$	ref=x_L, i.e., the mean of speaker-constant P_{min}
Method 11: ST-x$_H$	ref=x_H, i.e., the mean of speaker-constant P_{max}
Method 12: ST-(x$_H$+x$_L$)/2	ref=$(x_{H+}x_L)/2$
Method 13: ST-x$_{min}$	ref=x_{min}, i.e., the P_{min} of each speaker's data
Method 14: ST-x$_{max}$	ref=x_{max}, i.e., the P_{max} of each speaker's data
Method 15: ST-(x$_{min}$+x$_{max}$)/2	ref=$(x_{max}+x_{min})/2$

Many studies use the reference value of 100 Hz, thus getting F_0^{ST-100}=0 at 100 Hz, F_0^{ST-100}=12 at 200 Hz and F_0^{ST-100}=−12 at 50 Hz. Then F_0^{ST-100} is ***Method 8 ST-100***.

The change of reference value will not change the pitch shape. In other words, the slope between two points on the tonal contour is relatively independent of the semitone values; whether the reference is 1 Hz or 100 Hz, the values of their slopes are exactly the same. However, the semitone scale, centered to a fixed value (e.g., 100 Hz), cannot reduce any between-speaker variance. Some studies (e.g., Chen, 2008; Li, 2012) used semitone centered at 50 Hz for males and 100 Hz for females. Howard (1997) found that the average male pitch is around 123 Hz and that of females around 220 Hz. The split reference of males and females can reduce some sex-based differences in physiology, but the choices of 50 Hz and 100 Hz should be justified when being applied to the comparison of pitch height.

In addition to the fixed references, Andruski & Costello (2004) converted hertz to semitones relative to each speaker's average pitch, i.e., $F_0^{ST\text{-}AvgF0}$ (***Method 9 ST-AvgF$_0$***). Each speaker has a different reference based on their production data. This method makes each speaker's data comparable, not only the slope but also the pitch height. However, this method needs adjusting for the variationist study as well. If the reference is individual average pitch, the average pitch itself is inevitably affected by the tonal variation. So a reference is needed that (1) is relatively independent of the phonological variation; (2) reflects the anatomical differences and other style differences (relaxed/stressed, excited/calm). In other words, it is a speaker-constant that has low within-speaker variance.

Different unfixed references formed several different methods. ***Method 10 ST-x_L*** is semitone transformation centered to x_L, where x_L is the mean of the lowest F_0 of T2/T4/T6 in the dialects examined. ***Method 11 ST-x_H*** uses the speaker-constant P_{max}, the mean of highest F_0 of T1. ***Method 12 ST-$(x_H+x_L)/2$*** uses the mean of x_H and x_L. Similarly, ***Method 13 ST-x_{min}*** is semitone transformation centered to x_{min} of each speaker, ***Method 14 ST-x_{max}*** centered to x_{max} while ***Method 15 ST-$(x_{min}+x_{max})/2$*** is centered to the mean of x_{min} and x_{max}.

Apart from a series of semitone transformations, intonation research also used mel, bark and ERB-rate as psycho-acoustic scales, though these three scales have hardly been seen in tone research. For frequencies below 500 Hz, the main region for F_0 of the speech signal, mel and bark are nearly linear transformations of F_0 in Hz (Nolan, 2003), so they will not be discussed in this study. The ERB scale (Equivalent Rectangular Bandwidth), is proposed by Moore and Glasberg (1983). The equation used in the present study is proposed and shown in Glasberg and

Moore (1990). Its transformation is logarithmic at higher frequencies but between linear and logarithmic below 500 Hz (ibid.). This method is indexed as ***Method 16*** in the present study and calculated by equation (13).

(13) $F_0^{ERB} = 21.4 \times log_{10}(0.00437 \times F_0 + 1)$

4.3.4. Results

Following Adank et al. (2004) and Van der Harst (2011), the sixteen normalizations are evaluated on the basis of three criteria. First, it is investigated to what extent these normalizations preserve phonemic variation (Section 4.3.4.1). Second, the normalizations are tested with respect to their ability to minimize anatomical variation (Section 4.3.4.2). Finally, it is examined to what extent they preserve sociolinguistic variation (Section 4.3.4.3). All three criteria are evaluated on the basis of Linear Discriminant Analyses (LDAs). In these analyses, a certain category (i.e., phoneme, AGE or SEX) is predicted using a-, b- and c-coefficients obtained from the quadratic equation and tonal duration as predictors. The quadratic equations are calculated with the normalized or raw F_0 values. After the LDAs had been conducted, the three normalization methods that perform best in these analyses are selected in Section 4.3.4.4 and it is evaluated which method is the most optimal one for this study.

4.3.4.1. Preserving phonemic variation

LDA 1 was conducted in Wúxī and Sōngjiāng separately to investigate to what extent the normalizations preserve phonemic variation; in other words, how accurately the tonal category can be predicted on the basis of values resulting from the different normalization procedures. Besides level tone, three tones can be found in the T2, T4 and T6 of the Wúxī dialect and the Sōngjiāng dialect. They are peaking tone /131/ (Wúxī T4B, Sōngjiāng T2), rising tone /13/ (Wúxī T2, T4A and T6; Sōngjiāng T4 and T6) and falling tone /51/. The falling tone is an occasionally-used variant occurring in T2, T4 and T6 in the two regions, mainly because of the influence of PTH (see Chapter 5 and 6). So, in both regions, three tonal categories are predicted by LDA 1. A stepwise LDA was conducted in order to choose among four potential predictors: a-, b- and c-coefficients and tonal duration. In previous studies, it was never investigated which predictors play a role in differentiating between different tones.

Table 4.12 and Table 4.13 present the results for Wúxī LDA 1 and Sōngjiāng LDA 1 separately. In Wúxī data, there are 60 convex tokens (5.8%), 845 rising tokens (81.5%) and 27 falling tokens (2.6%). In Sōngjiāng data, there are 266 convex tokens (25.6%), 655 rising tokens (63.9) and 34 falling tokens (3.3%).

Table 4.12 Results for Wúxī LDA 1 (stepwise): percent correctly classified tonal shapes based on quadratic coefficients predictors: a, b and c and tonal duration

	methods	**selected predictors**	**predicted accuracy of curve shapes**			**overall accuracy**	**rank**
			peaking	**rising**	**falling**		
0	Hz	a, b, c, duration	70.0	98.9	40.7	95.4	16
1	Z-score	a, b, c, duration	71.7	98.7	48.1	95.5	13
2	FOR	a, b, c, duration	73.3	98.3	51.9	95.4	16
3	PORc1	a, b, c, duration	71.7	98.7	48.1	95.5	13
	PORc2	a, b, c, duration	71.7	98.7	48.1	95.5	13
4	LD	a, b, c, duration	76.7	98.1	51.9	95.4	16
5	T	a, b, c, duration	76.7	98.7	51.9	95.9	9
6	LZ	a, b, c, duration	78.3	98.6	44.4	95.7	12
7	LPORc1	a, b, c, duration	80.0	98.6	44.4	95.8	10
	LPORc2	a, b, c, duration	80.0	98.6	44.4	95.8	10
8	ST-100	b, c, duration	78.3	99.1	48.1	96.2	1
9	ST-AvgF_0	b, c, duration	78.3	99.1	48.1	96.2	1
10	ST-x_L	a, b, c, duration	75.0	99.1	48.1	96.0	7
11	ST-x_H	b, c, duration	78.3	99.1	48.1	96.2	1
12	ST-$(x_H+x_L)/2$	a, b, c, duration	78.3	99.1	48.1	96.2	1
13	ST-x_{min}	a, b, c, duration	76.7	99.1	44.4	96.0	7
14	ST-x_{max}	a, b, c, duration	80.0	98.9	48.1	96.2	1
15	ST-$(x_{min}+x_{max})/2$	b, c, duration	78.3	99.1	48.1	96.2	1
16	ERB	a, b, c, duration	70.0	99.1	40.7	95.5	16

Table 4.13 Results for Sōngjiāng LDA 1 (stepwise): percent correctly classified tonal shapes based on quadratic coefficients predictors: a, b and c and tonal duration

	methods	**selected predictors**	**predicted accuracy of curve shapes**			**overall accuracy**	**rank**
			peaking	**rising**	**falling**		
0	Hz	a, b, c, duration	59.8	99.1	47.1	86.4	17
1	Z-score	a, b, c, duration	87.2	97.6	61.8	93.5	2
2	FOR	a, b, c, duration	86.1	97.6	55.9	93.0	5
3	PORc1	a, b, c, duration	87.2	97.6	61.8	93.5	2
	PORc2	a, b, c, duration	97.2	97.6	61.8	93.5	2
4	LD	a, b, c, duration	88.0	98.0	64.7	94.1	1
5	T	a, b, c, duration	84.6	98.3	50.0	92.8	6
6	LZ	a, b, c, duration	73.3	97.9	50.0	89.4	15
7	LPORc1	a, b, c, duration	75.6	98.0	52.9	90.3	14
	LPORc2	a, b, c, duration	75.6	98.0	52.9	90.3	14
8	ST-100	a, b, c, duration	78.9	98.3	50.0	91.3	11
9	ST-AvgF_0	a, b, c, duration	80.8	98.2	50.0	91.7	9
10	ST-x_L	a, b, c, duration	79.3	98.0	50.0	91.2	12
11	ST-x_H	a, b, c, duration	82.7	98.3	50.0	92.3	7
12	ST-$(x_H+x_L)/2$	a, b, c, duration	79.7	98.2	50.0	91.4	10
13	ST-x_{min}	a, b, c, duration	79.3	98.0	50.0	91.2	12
14	ST-x_{max}	a, b, c, duration	84.6	98.6	52.9	93.2	4
15	ST-$(x_{min}+x_{max})/2$	a, b, c, duration	80.8	98.3	50.0	91.8	8
16	ERB	a, b, c, duration	71.4	98.5	47.1	89.2	16

On the whole, all the normalization procedures have a quite high overall accuracy of prediction. No large differences were found among different procedures. In other words, all the normalization procedures can preserve phonemic variation quite well. LZ, the procedure evaluated as the best one on the basis of the NI index, did not perform well in LDA analysis (rank 12 in Wúxī and rank 15 in Sōngjiāng).

4.3.4.2. Minimizing anatomically-based variation

In order to test the extent to which the normalization procedures minimize sex-related anatomically-based variation, LDA 2 was conducted for the rising tones in Wúxī and Sōngjiāng separately. If a normalization method successfully removes the variation, the success rate for that procedure will be at chance level, i.e., 50%. The results are given in Table 4.14 and Table 4.15.

Table 4.14 Results for Wúxī LDA 2 (stepwise): percent correctly-classified SEX based on quadratic coefficients predictors: quadratic coefficient a, b and c and tonal duration

	methods	selected predictors	predicted accuracy of SEX		overall accuracy	rank
			male	female		
0	Hz	a, b, c, duration	98.5	82.5	90.3	15
1	Z-score	-	-	-	-	1
2	FOR	b, c, duration	53.8	58.4	56.2	5
3	PORc1	-	-	-	-	1
	PORc2	-	-	-	-	1
4	LD	b, c, duration	53.5	60.5	57.1	10
5	T	a	55.3	65.0	60.3	12
6	LZ	b	44.0	65.7	55.1	4
7	LPORc1	a, c, duration	44.0	65.4	54.9	3
	LPORc2	a, c, duration	44.0	65.4	54.9	3
8	ST-100	b, duration	96.7	84.1	90.3	15
9	ST-AvgF_0	a, c, duration	49.1	64.0	56.7	6
10	ST-x_L	a, c	56.4	61.9	59.2	11
11	ST-x_H	a, b, c	56.8	66.1	61.5	13
12	ST-$(x_H+x_L)/2$	b	49.1	64.0	56.7	6
13	ST-x_{min}	a, c	64.1	65.7	64.9	14
14	ST-x_{max}	a, c, duration	49.1	64.0	56.7	6
15	ST-$(x_{min}+x_{max})/2$	a, c, duration	49.1	64.0	56.7	6
16	ERB	a, b, c, duration	97.8	83.6	90.5	17

In Table 4.14, LDA 2 cannot be completed for the procedures Z-score, PORc1 and PORc2, and they are marked with "-" in the table. Such situation occurred because these procedures removed almost all the sex-related anatomically-based variation, then further analyses are not computable. For the sake of method comparison, their predicted accuracy of sex will be calculated as chance level (50%) in Section 4.3.4.4.

In the LDA 2 procedure, it was shown that the raw values carry a great deal of anatomically-based variation, as the success rate for predicting speaker sex was quite high (around 90% in both regions). The procedure showed similar results with ST-100 and ERB in terms of the raw values, which carried much anatomically-based variation. Besides ST-100 and ERB, other normalizations are able to remove the variation successfully. In Wúxī data, Z-score, PORc1 and PORc2 performed best in removing the variation, whereas ST-x_{min} (64.9%) and ST-x_H (61.5%) removed the variation to a lesser extent. In Sōngjiāng, ST-AvgF_0 and ST-x_H predicted speaker sex closest to chance level (both 66.9%), whereas LZ, LPORc1 and LPORc2 removed the variation to a lesser extent (73.9%). In general, different normalizations (excluding ST-100 and ERB) do not have large differences in predicting speaker sex (Wúxī: 64.9%−50%=14.9%; Sōngjiāng: 73.9%−66.9%=7%). Alternatively, besides

ST-100 and ERB, the normalizations in the literature or those proposed in this study all performed well in removing sex-related anatomically-based variation.

Table 4.15 Results for Sōngjiāng LDA 2 (stepwise): percent correctly-classified SEX based on quadratic coefficients predictors: quadratic coefficient a, b and c and tonal duration

	methods	selected predictors	predicted accuracy of SEX		overall accuracy	rank
			male	**female**		
0	Hz	a, c, duration	91.4	88.4	89.9	17
1	Z-score	a, b, duration	67.6	71.3	69.4	8
2	FOR	a, b, duration	68.3	73.5	70.9	12
3	PORc1	a, b, duration	67.6	71.3	69.4	8
	PORc2	a, b, duration	67.6	71.3	69.4	8
4	LD	a, b, duration	69.5	71.0	70.2	11
5	T	b, duration	67.6	69.7	68.6	7
6	LZ	a, duration	69.5	78.4	73.9	13
7	LPORc1	a, duration	69.5	78.4	73.9	13
	LPORc2	a, duration	69.5	78.4	73.9	13
8	ST-100	a, duration	87.3	91.0	89.1	15
9	ST-AvgF$_0$	a, b, c, duration	67.9	65.8	66.9	1
10	ST-x_L	a, b, duration	69.2	67.1	68.2	6
11	ST-x_H	a, b, c	70.8	62.9	66.9	1
12	ST-$(x_H+x_L)/2$	b, duration	66.7	69.0	67.8	3
13	ST-x_{min}	b, duration	66.7	69.0	67.8	3
14	ST-x_{max}	a, b, duration	65.7	73.2	69.4	8
15	ST-$(x_{min}+x_{max})/2$	b, duration	66.7	69.0	67.8	3
16	ERB	a, duration	88.9	89.4	89.1	16

Preserving sociolinguistic variation

In order to investigate the extent to which normalization procedures preserve sociolinguistic variation, i.e., age-related variation, LDA 3 was conducted to predict whether the rising tones were spoken by the old people or by the young people in Wúxī and Sōngjiāng. Chance level is 50%, and if a procedure shows a success rate far above 50%, it demonstrates its ability to preserve sociolinguistic variation. The results for LDA 3 are given in Table 4.16 and Table 4.17 for the data of Wúxī and Sōngjiāng.

Table 4.16 Results for Wúxī LDA 3 (stepwise): percent correctly-classified AGE based on quadratic coefficients predictors: quadratic coefficient a, b and c and tonal duration

	methods	selected predictors	predicted accuracy of AGE		overall accuracy	rank
			old	young		
0	Hz	a, b, c, duration	75.8	75.5	75.7	7
1	Z-score	a, b, c, duration	79.5	64.8	72.6	14
2	FOR	a, duration	81.9	65.9	74.4	10
3	PORc1	a, b, c, duration	79.5	64.8	72.6	14
	PORc2	a, b, c, duration	79.5	64.8	72.6	14
4	LD	a, duration	81.2	67.0	74.6	8
5	T	a, b, c, duration	73.5	70.1	71.9	18
6	LZ	a, b, c, duration	78.9	67.8	73.7	11
7	LPORc1	a, duration	77.9	68.6	73.5	12
	LPORc2	a, duration	77.9	68.6	73.5	12
8	ST-100	a, b, c, duration	71.7	62.4	67.3	19
9	ST-AvgF_0	a, b, c, duration	80.6	93.3	86.5	3
10	ST-x_L	a, b, c, duration	77.9	73.9	76.0	6
11	ST-x_H	a, b, c, duration	79.5	94.3	86.4	4
12	ST-$(x_H+x_L)/2$	a, b, c, duration	78.5	91.2	84.4	5
13	ST-x_{min}	a, b, c, duration	72.8	72.0	72.5	17
14	ST-x_{max}	a, b, c, duration	80.5	95.0	87.3	1
15	ST-$(x_{min}+x_{max})/2$	a, b, c, duration	82.2	91.6	86.6	2
16	ERB	a, b, c, duration	74.5	74.7	74.6	9

Table 4.17 Results for Sōngjiāng LDA 3 (stepwise): percent correctly-classified AGE based on quadratic coefficients predictors: quadratic coefficient a, b and c and tonal duration

	methods	selected predictors	predicted accuracy of AGE		overall accuracy	rank
			old	young		
0	Hz	a, b, c, duration	86.0	94.3	89.9	6
1	Z-score	a, duration	87.2	88.2	87.7	11
2	FOR	b, c, duration	84.8	86.5	85.6	15
3	PORc1	a, duration	87.2	88.2	87.7	11
	PORc2	a, duration	87.2	88.2	87.7	11
4	LD	b, c, duration	84.5	87.5	85.9	14
5	T	a, b, c, duration	83.8	91.2	87.4	13
6	LZ	a, b, duration	82.3	89.2	85.6	15
7	LPORc1	a, b, duration	82.6	88.6	85.4	17
	LPORc2	a, b, duration	82.6	88.6	85.4	17
8	ST-100	b, c, duration	84.8	96.6	90.4	5
9	ST-AvgF_0	a, b, c, duration	89.6	98.0	93.6	3
10	ST-x_L	a, b, c, duration	84.1	96.0	89.8	9
11	ST-x_H	a, duration	92.7	99.0	95.7	1
12	ST-$(x_H+x_L)/2$	a, b, c, duration	90.9	97.6	94.1	2
13	ST-x_{min}	a, b, c, duration	84.8	95.6	89.9	6
14	ST-x_{max}	a, duration	91.2	91.2	91.2	4
15	ST-$(x_{min}+x_{max})/2$	a, duration	88.1	91.2	89.6	10
16	ERB	a, b, c, duration	85.1	95.3	89.9	6

Table 4.16 and Table 4.17 reveal some interesting results. First, the raw F_0 in hertz preserves most age-related variation, even better than the majority of normalization procedures (rank 7 in Wúxī and rank 6 in Sōngjiāng). This is reasonable as age-related variation is the steepness variation of rising tone, i.e., the variation of pitch shape, which should be reflected by the raw F_0.

Second, considering the performance of raw F_0 as a baseline, the normalization procedures can be split into two groups: the group preserving less sociolinguistic variation than the baseline and the group preserving sociolinguistic variation not less (more or equal) than the baseline. As can be seen from Table 4.16 and Table 4.17, most procedures of semitone transformation can preserve more sociolinguistic variation than baseline whereas the six methods presented in Zhu (1999), as well as T values, all remove some sociolinguistic variation in comparison with the baseline. It is not surprising that these seven methods, i.e., Z-score, FOR, POR, LD, T, LZ and LPOR, cause attrition of sociolinguistic variation because their transformations are principally composed of two steps: (1) parallel shift of tones on the coordinate using (logarithmic) mean or (speaker-constant) P_{min} (i.e., x_{min} or x_L) as reference and (2) range compression or expansion based on standard deviation or R-D range[30] (c.f., Zhu, 1999: 46[31]). The compression or expansion of the range will consequently reduce or enlarge the within-speaker variance, causing a change of speaker weight in the cross-speaker variance, i.e., the sociolinguistic variation. So these procedures cannot preserve as much sociolinguistic variation as the baseline.

With respect to another group of normalizations, i.e., the group preserving sociolinguistic variation not less than the baseline, they are mainly semitone transformations centered to varying references. They only create a parallel shift of tones on the coordinate but do not carry out range compression or expansion. The choices of references for semitones determine how much improvement can be made on the preservation of sociolinguistic variation with the procedure. Comparing the results of Wúxī and Sōngjiāng, only ST-AvgF_0, ST-x_H, ST-$(x_H+x_L)/2$, ST-x_{max} show better performances than the baseline in both regions. So it appears that the best procedure for tone normalization will be one of these. The following section 4.3.4.4

[30] R-D range is the range between two range-defining points (R-D). R-D points can be either x_{max} and x_{min}, or x_H and x_L.
[31] Zhu (1999: 46) pointed out that Z-score, POR, LD, and LZ are composed of these two steps. According to the equations of FOR, LPOR and T, these three transformations are also composed of these two steps.

summarizes and compares the performances of the sixteen normalization procedures discussed.

4.3.4.3. Summary of the LDA results: selection of ST-AvgF$_0$

In the previous sections, I presented the LDA results with respect to the preservation of phonemic variation and sociolinguistic variation and the reduction of anatomically-based variation. Table 4.18 gives a summary of those LDA results.

In Table 4.18, the LDA 1 column gives the percent change from the baseline for success in preserving phonemic variation. If a procedure showed success rates close to 100%, it can be concluded that it preserves more phonemic variation. Percentage change from the baseline is calculated by using the overall accuracy of each procedure minus the overall accuracy of the baseline, i.e., F_0 in Hz. The greater the difference, the better the procedure performed than the baseline.

The LDA 2 column gives the percentage change from the baseline for success in reducing anatomically-based variation. If a procedure successfully removes this variation, the success rate for that procedure will be at chance level (i.e., 50%), the closer the accurate rate of LDA3 is to the chance level 50%, the more variation that method removes. So the percentage change from baseline is used to compare which procedure is closest to 50% compared to the baseline. It can be calculated by using the overall accuracy of the baseline minus the overall accuracy of each procedure. The greater the difference with the baseline, the better this procedure is than the baseline. The LDA3 column gives the percent change from the baseline for success in preserving sociolinguistic variation. The calculation is the same as for LDA 1. Negative values in this column indicate that those transformations removed even more sociolinguistic variation than the baseline. As shown in Table 4.12, Table 4.13, Table 4.16 and Table 4.17, LDA 1 and LDA3 predicted the tonal shapes and speaker age group quite well, thus the accuracy rates of 16 methods do not vary much due to the ceiling effect, whereas, in LDA 2, accuracy rates of predicting speaker sex do vary considerably among different normalizations. Therefore, in addition to the original percentage difference from the baseline for each LDA, their Z-scores are presented in Table 4.18 as well in order to normalize the variation within LDAs. Then the mean change from baseline in Z-score is calculated and given in the column “mean change”. The mean is calculated because, in the present study, the three aims of the normalization procedures are considered equally important. The rank column gives the rank orders based on the mean change. The results for the two regions and three LDAs are calculated means, listed in the "total" column.

Table 4.18 Percent change in overall accuracy rate from the baseline (F_0 in Hz) for the comparisons for each procedure

	methods	Wúxī						Sōngjiāng						total	
		LDA 1		LDA 2		LDA 3		LDA 1		LDA 2		LDA 3		mean change	rank
		%	Z	%	Z	%	Z	%	Z	%	Z	%	Z		
0	Hz	-	-	-	-	-	-	-	-	-	-	-	-	-	-
1	Z-score	0.1	−1.09	40.3	0.86	−3.1	−0.65	7.1	1.11	20.5	0.36	−2.2	0.24	0.139	6
2	FOR	0	−1.40	34.1	0.33	−1.3	−0.37	6.6	0.77	19	0.14	−4.3	−1.14	−0.278	12
3	PORc1	0.1	−1.09	40.3	0.86	−3.1	−0.65	7.1	1.11	20.5	0.36	−2.2	−0.47	0.022	8
	PORc2	0.1	−1.09	40.3	0.86	−3.1	−0.65	7.1	1.11	20.5	0.36	−2.2	−0.47	0.022	8
4	LD	0	−1.40	33.2	0.26	−1.1	−0.33	7.7	1.52	19.7	0.24	−4	−1.04	−0.127	11
5	T	0.5	0.18	30	−0.02	−3.8	−0.76	6.4	0.63	21.3	0.48	−2.5	−0.56	−0.009	9
6	LZ	0.3	−0.46	35.2	0.43	−2	−0.48	3	−1.68	16	−0.31	−4.3	−1.14	−0.606	14
7	LPORc1	0.4	−0.14	35.4	0.44	−2.2	−0.51	3.9	−1.07	16	−0.31	−4.5	−1.20	−0.465	13
	LPORc2	0.4	−0.14	35.4	0.44	−2.2	−0.51	3.9	−1.07	16	−0.31	−4.5	−1.20	−0.465	13
8	ST-100	0.8	1.12	0	−2.59	−8.4	−1.49	4.9	−0.39	0.8	−2.59	0.5	0.4	−0.922	15
9	ST-AvgF_0	0.8	1.12	33.6	0.29	10.8	1.54	5.3	−0.12	23	0.74	3.7	1.42	0.832	3
10	ST-x_L	0.6	0.49	31.1	0.08	0.3	−0.11	4.8	−0.46	21.7	0.54	−0.1	0.21	0.124	7
11	ST-x_H	0.8	1.12	28.8	−0.12	10.7	1.53	5.9	0.29	23	0.74	5.8	2.09	0.941	1
12	ST-$(x_H+x_L)/2$	0.8	1.12	33.6	0.29	8.7	1.21	5	−0.32	22.1	0.6	4.2	1.58	0.747	4
13	ST-x_{min}	0.6	0.49	25.4	−0.41	−3.2	−0.67	4.8	−0.46	22.1	0.6	0	0.24	−0.034	10
14	ST-x_{max}	0.8	1.12	33.6	0.29	11.6	1.67	6.8	0.9	20.5	0.36	1.3	0.65	0.833	2
15	ST-$(x_{min}+x_{max})/2$	0.8	1.12	33.6	0.29	10.9	1.56	5.4	−0.05	22.1	0.6	−0.3	0.14	0.611	5
16	ERB	0.1	−1.09	−0.2	−2.60	−0.4	−0.33	2.8	−1.82	0.8	−2.59	0	0.24	−1.366	16

The results in Table 4.18 show that ST-x_H, ST-x_{max} and ST-AvgF_0 are the top three normalization procedures. They outperform the baseline Hz in all respects. ST-x_H is just slightly better than ST-$x_{max,}$ ST-AvgF_0. However, considering that the reference x_H is language specific, x_H is not simply P_{max} of a given speaker but P_{max} in a particular tone. This means that different languages can have different definitions for x_H, which complicates cross-linguistic comparison. ST-x_{max} and ST-AvgF_0 only have a difference of 0.001 Z-score, while the value of x_{max} is more easily influenced by falsetto register, thus being less stable than AvgF_0. For these reasons, ST-AvgF_0, i.e., semitone transformation relative to each speaker's average pitch, is the best normalization procedure and will be used to normalize F_0 values during the remainder of this study.

4.4. Phonation measurements

4.4.1. Introduction

As discussed in Section 4.1.2, MC muddy stops (全浊), including explosives and affricates, were perceived as voiced, but the acoustic instrumental studies rejected this statement. A large number of phonetic studies (Cao and Maddieson, 1992; Z. Chén, 2010; Ren, 1992; Shi, et al., 1987) discovered the breathiness features in the Wú dialects. Monosyllabic morphemes initialed with "voiced" consonants, i.e., MC muddy stops, are breathy voice, whereas those initialed with voiceless consonants are mostly modal voice. The "voiced" consonants became truly voiced in unstressed intervocalic position, but also with breathiness.

Breathiness is associated with the lowering of tone in the majority of languages (see Hombert, Ohala, & Ewan, 1979 for an overview). Since this dissertation aims to investigate the tonal variation in the Yáng register, it will be interesting to investigate the relevance of breathiness to the Yáng register. What are the phonetic bases of the Yīn-Yáng contrast preserved in the Wú dialects? Are there any variations in the phonation types? If so, what are the effects of phonation variation on the tonal variation?

As no articulatory data was collected in this study, it is impossible to examine the glottal source directly. Therefore, the judgment of the phonation types in the Wú dialects is (mainly) based on acoustic measurement of the recorded sound files, more specifically, the examination of power spectra.

The most widely used acoustical parameter is H_1-H_2, i.e., the difference between the amplitudes of the first harmonic (equivalent to F_0, H_1) and the second harmonic (H_2) in the spectrum. H_1-H_2 has been proven to correlate closely to the glottal open quotient (OQ) derived from electoglottography (EGG) (Holmberg, et al., 1995). Breathy voice is defined as a state of the glottis such that vocal folds are not closed completely during vibrating, so that they allow much more air to escape than in modal voice. Reflecting this difference in the fold pulses, breathy voice vibrations have much of the energy in the fundamental frequency (the first harmonic), where H_1 is much higher than H_2 (Ladefoged, 2003). The parameter H_1-H_2 successfully revealed spectral characteristics of contrasting phonation types in a variety of languages, but this parameter is language dependent (Esposito, 2006; 2010). Esposito (2006; 2010) and Blankenship (2002) reported that H1-H2 did not distinguish the breathy phonation from the modal voice in Mon and Tamang, though it perceived phonation differences in eight other languages. In addition, Kreiman et al. (2007) found that the relationship of H_1-H_2 to phonatory characteristics appears to be speaker-dependent, because individual speakers use different strategies to produce the same intended voice qualities. With regard to the Shànghǎi dialect, Cao & Maddieson (1992) found significant differences of H_1-H_2 between breathy and modal voices. But it was also pointed out by Chén (2010) that, in the Shànghǎi dialect, H_1-H_2 is not an effective parameter for differentiating the breathiness from the modal voice because positive H_1-H_2 was also found in the modal voice and showed no significant difference from the breathy voice.

Therefore, in addition to H_1-H_2, this study will also calculate other spectral measurements to reflect other aspects of phonation. Garellek & Keating (2011) assume that the strength of higher frequencies in the spectrum is related to the closing velocity of the vocal folds, to the presence of a posterior glottal opening, and to the simultaneity of ligamental closure, among other possible influences. Higher frequency energy is usually measured as the amplitude of H_1 relative to the amplitude of harmonics affiliated with the first (A_1), second (A_2) and third formants (A_3), as H_1-A_1, H_1-A_2, and H_1-A_3. This spectral tilt is derived from the mid-range region of the spectrum and can distinguish breathy from modal phonation fairly well, but it is hard to generalize a principal like H_1-H_2 – that, in breathy voice, it is higher than in modal voice. Esposito (2006) showed that H_1-A_1 distinguished breathy from modal in six out of ten language varieties, among which two had higher H_1-A_1 in modal voice than breathy voice while the other four had lower H_1-A_1 in modal voice. H_1-A_2 distinguished breathy from modal in seven out of ten language varieties; these seven cases all had higher H_1-A_2 in breathy voice. H_1-A_3 distinguished breathy from modal in nine out of ten languages/ vernaculars, where the modal voices of five had higher H_1-A_3, while four had lower H_1-A_3 than breathy voice. The other three parameters examined in Esposito (2006), "$((H_1+H_2)/2)$ -A_1", "H_2-H_4", and "A_2-A_3", will not be discussed in this study due to their low accuracy in differentiating phonation types.

In the following analyses, four parameters, H_1-H_2, H_1-A_1, H_1-A_2, and H_1-A_3 of morphemes in the high (modal phonation) and the Yáng register (breathy phonation) are compared.

4.4.2. Acoustic analyses on parameters for distinguishing phonation types

The acoustic data utilized in this section is a subset of read speech of isolated words in the Wúxī dialect. Speakers are 20 old people from both the urban and suburban areas. For each speaker, 45 morphemes from nine tonal subsets were analyzed, with five morphemes per subset. Among them, 15 morphemes in modal voice are from T1A, T3A and T5A, and 30 morphemes in breathy voice are from T2A, T2B, T4A, T4B, T6A and T6B. As introduced in Chapters 2 and 3, morphemes of T1A, T3A and T5A are from the Yīn register, which is in modal voice. These morphemes, rather than T1B, T3B and T5B morphemes, initial with unaspirated obstruents to avoid the influence of consonants' aspiration on the beginning of vowels. 30 morphemes from the low register have breathy voice. This section will test whether H_1-H_2, H_1-A_1, H_1-A_2, and H_1-A_3 can distinguish the 15 morphemes of T1A, T3A, T5A from the 30 morphemes of T2, T4 and T6. Considering the possible effects of vowels on these parameters, the five morphemes from the same category have five different vowels /i/, /ɛ/, /a/, /u/, and /ʌ/.

The domain of tone was split in three equal parts. Spectral slices were taken and four indexes - H_1-H_2, H_1-A_1, H_1-A_2, and H_1-A_3 - were extracted for the 1st, 2nd and 3rd slice. A script written by Chad Vicenik[32] was used to extract those 12 scores. Two sets of formant values were used depending on the speaker's sex. For males,

[32] The script is available from www.linguistics.ucla.edu/faciliti/facilities/acoustic/PraatVoiceSauceImitator.txt, retrieved on May 20, 2014

these reference values were F1 500 Hz, F2 1485 Hz, and F3 2475 Hz; for females, these values were F1 550 Hz, F2 1650 Hz, F3 2750 Hz.

A two-way ANOVA was conducted to test whether there were effects of vowel (five levels) and register (two levels) on H_1-H_2, H_1-A_1, H_1-A_2, and H_1-A_3 for each part of the tonal domain.

Table 4.19 Results of two-way ANOVA on H_1-H_2, H_1-A_1, H_1-A_2, and H_1-A_3

Index	Main effect of register	Main effect of vowel
H_1-H_2		
1	F(1, 890)=33.817, *p*=0.000	F(4, 890)=8.965, *p*=0.000
2	F(1, 890)=10.724, *p*=0.001	F(4, 890)=5.507, *p*=0.000
3	F(1, 890)=12.602, *p*=0.000	F(4, 890)=1.833, *p*=0.120
H_1-A_1		
1	F(1, 890)=44.269, *p*=0.000	F(4, 890)=8.622, *p*=0.000
2	F(1, 890)=24.265, *p*=0.000	F(4, 890)=4.812, *p*=0.001
3	F(1, 890)=4.773, *p*=0.029	F(4, 890)=1.326, *p*=0.259
H_1-A_2		
1	F(1, 890)=50.432, *p*=0.000	F(4, 890)=28.479, *p*=0.000
2	F(1, 890)=28.523, *p*=0.000	F(4, 890)=22.123, *p*=0.000
3	F(1, 890)=8.644, *p*=0.003	F(4, 890)=18.539, *p*=0.000
H_1-A_3		
1	F(1, 890)=42.802, *p*=0.000	F(4, 890)=34.974, *p*=0.000
2	F(1, 890)=15.718, *p*=0.000	F(4, 890)=34.471, *p*=0.000
3	F(1, 890)=2.789, *p*=0.095	F(4, 890)=34.008, *p*=0.000

As can be seen from Table 4.19, the main effect of vowel was significant in most cases, except for the measure of H_1-H_2, and H_1-A_1 at the last time index. The main effect of register was significant in most cases as well, except for the measure of H_1-A_3 at the last time index. In other words, the measure of H_1-A_3 at the last time index cannot independently distinguish breathy voice from modal voice. For this reason, H_1-A_3 at the last time index was not included in the following discriminant model.

A stepwise discriminant analysis (DA) was then performed on linear models to find the best predictor(s) for distinguishing breathy voice. Predictors include 11 variables, H_1-H_2, H_1-A_1 and H_1-A_2 at all the three time indices and H1-A3 at the first two time indices. For ease of reporting, the prefix "T1", "T2" or "T3" was added to indicate the time index, for example, T1_H_1-H_2 means H_1-H_2 measurement at the first time index.

The stepwise statistics table shows that four steps were taken. The most correlated independent variable T1_H_1-A_1. In the following three steps, one variable is entered for each step; the sequence being T1_H_1-A_2, T1_H_1-H_2, T3_H_1-A_2. At each step, the variable that minimizes the overall Wilks' Lambda is entered. The stepwise LDA, as an additional dependent variable, added no significant amount to the canonical R squared. The discriminate function revealed a significant association (p=0.000) between breathy-modal contrasts and four significant predictors were selected. Though there is a highly significant function of DA, Canonical Correlation (p=0.276) suggests that the model can only explain 27.6% of variance in the data. The cross validated classification showed that overall 69.9% were correctly classified (14.7% for modal voice and 96.8% for breathy voice). The correct prediction of modal voice is quite low, which means the present LDA model is not valid enough to predict the modal voice. None of the variables can successfully distinguish the breathy voice from the modal voice. This might be due to the fact that our recordings were not made in a sound-proof booth.

Chén's (2000) phonetic investigation on the acoustic features of breathy voice in Wú pointed out that H_1-H_2 is not a valid parameter for separating breathy and modal voice in the Shànghǎi dialect. Our investigation, based on a large amount of data, confirms his findings, and adds that measurements of H_1-A_1, H_1-A_2 and H_1-A_3 are not valid parameters either, not only in the Shànghǎi dialect but also in the Wúxī dialect. It is fairly safe to say that these findings are probably general in the northern Wú dialects. Chén proposed some identification markers for breathy voice in the Wú dialects: (1) the breathy voice has more gentle vibrations of power spectra than the modal voice and (2) the breathy voice may have random waveforms indicating noises in the area above 2700 Hz in the power spectra. However, both markers are quite difficult to quantify. In other words, so far, no quantifiable parameter for distinguishing breathy voice and modal voice is available.

Therefore, this study has to give up the idea of investigating the relationship between tonal variation and phonation variation. The efforts made in this study are reported for referential use in further research.

To sum, Chapter 4 described the acoustic analysis of the recorded tones. Two methodological issues in tonal acoustic analyses were resolved: (1) the domain of tone in the Wú dialects includes three portions: the (optional) prenuclear onglide, the nucleus and the (optional) final nasal consonant; and (2) the best tone normalization procedure is a semitone transformation relative to each speaker's average pitch in Hz. An attempt to fink valid parameters to distinguish breathy voice and modal voice was unsuccessful.

Chapter 5. Tonal variation in Wúxī

In this chapter, I present and discuss the results of T2, T4 and T6 variation in the Wúxī data. As this study is the first attempt to investigate tonal variation and change in the Wúxī dialect, the first task is to identify the tonal variables affected by a change in progress. This is done in Section 5.1 for the citation tone and in Section 5.3 for the tone sandhi. After having defined the tonal variables, regression analyses are conducted to interpret the variation from both the linguistic and social perspectives. Sections 5.2 and 5.4 present the analyses and interpretation for the citation tone and tone sandhi, separately. Finally, Section 5.5 gives a brief summary and the main conclusions of this chapter.

5.1. Identifying linguistic variables of citation tone

Section 2.4.4.1 reviewed several Wúxī dialect studies and found that its T2, T4 and T6 have been undergoing changes for the past few decades. After transcribing all the tokens of the current dataset, falling tones (/51/ or /41/) are found in T2, T4 and T6, which are tone patterns not recorded before. This unexpected finding is examined in Section 5.1.1. Moreover, level tones are found for T2, T4 and T6 in the current dataset as well. Although T4 was recorded as a level tone /33/ in GSDJPS (1960) and Qián (1992), T2 and T6 morphemes are also realized as /33/ in my dataset. So the level tone /33/ is examined in Section 5.1.1 as well.

The latest acoustic research on the Wúxī tones (Xú, 2007) indicated that T2, T4 and T6 were undergoing a merger. Xú (2007) used the data of Wúxī urban speakers who were around 60 years old in 2005. He found that Wúxī T2 first moved into T6 and then T4 changed from a peaking curve /131/ to a similar low rising pattern /13/ with T2/T6. Based on this finding, it is deduced that the potential linguistic variables can be identified in (1) T2 and T6, if their merger is incomplete and (2) T4, if its change is ongoing. So accordingly, two questions will be examined in order to trace these two potential linguistic variables: (1) have Wúxī T2 and T6 merged completely? (Section 5.1.2); (2) is Wúxī T4 subject to a change in progress? (Section 5.1.3)

5.1.1. Falling and level tones

Falling and level tones in T2, T4 and T6 were investigated first because on the one hand, they were almost never recorded before, and, on the other hand, because these two tones account for only a small part of the total observations. In the 5373 tokens of T2, T4 and T6, there are 212 falling tones (3.9%) and 239 level tones (4.4%). When transcribing the tones, the author found that the level and falling tones have the following features.

Table 5.1 Distribution of falling and level tokens by WORD

tone shape	% within the tone shape	WORD
falling n=212	0-0.9	伴, 荡, 稻, 豆, 度, 份, 凤, 奉, 键, 舅, 狼, 朗, 涝, 楼, 漏, 路, 瞒, 帽, 免, 盘, 球, 糖, 逃, 头, 义, 涌, 用, 邮, 云, 运, 陈, 晨, 道, 坟, 冷, 鲁, 米, 午, 友, 预, 词, 肚, 附, 湖, 篮, 牢, 李, 刘, 买, 满, 迷, 抬, 题, 勇, 余, 雨, 赵
	1-1.9	败, 曹, 聚, 埋, 卖, 毛, 谜, 吴, 罢, 大, 腐, 户, 健, 懒, 礼, 闹, 仁, 上, 寿, 寺
	2-2.9	蛋, 烂, 老, 渠, 护, 面, 排, 误, 元, 召, 诞, 弟, 第, 具, 造, 字
	3-3.9	似, 咸, 廖
level n=239	0-0.9	荡, 凤, 舅, 狼, 朗, 楼, 帽, 盘, 糖, 逃, 头, 义, 邮, 陈, 晨, 肚, 篮, 刘, 题, 赵, 埋, 大, 礼, 仁, 度, 键, 漏, 路, 瞒, 免, 球, 涌, 用, 云, 运, 坟, 鲁, 米, 词, 附, 湖, 抬, 勇, 余, 曹, 卖, 吴, 腐, 健, 渠, 元, 伴, 稻, 份, 道, 冷, 牢, 李, 买, 谜, 户, 懒, 闹, 上, 面, 排, 咸
	1-1.9	豆, 友, 预, 雨, 聚, 毛, 寿, 寺, 老, 弟, 奉, 涝, 午, 迷, 护, 误, 召, 具, 造, 似
	2-2.9	败, 廖, 满, 第
	3-3.9	/
	4-4.9	烂
	5-5.9	蛋, 字
	6-6.9	罢, 诞

First, the falling and level tokens have an unbalanced distribution among words. There are 96 different T2, T4 and T6 morphemes in three reading styles. In Table 5.1, the relative frequency and percentage of each morpheme within the falling and level tokens are listed. Only 39 morphemes were realized as falling tone more than twice, which are listed in the rows over 1% in Table 5.1. The top three frequent falling morphemes are 廖 (liào, 'a Chinese surname', 3.8%), 似 (sì, 'similar', 3.3%) and 咸 (xián, 'salty', 3.3%). For the level tone, 29 morphemes out of 96 were realized as level tone more than twice, which are also listed in the rows over 1% in the "level" part of Table 5.1, and nine of them were pronounced as /33/ more than five times, making up 39.2% of the total level tone tokens. Those nine items are listed in the rows over 2% in the “level” part of Table 5.1.

Second, the falling and level tokens are distributed unequally among speakers as shown in Table 5.2. WUym1 (29 tokens, 12.1%) and WHym4 (20 tokens, 8.4%) used falling tone most frequently, making over one fifth of the total falling observations. Regarding the level tones, three young speakers used them most frequently: WHym3 (46 tokens, 21.7%), WUym1 (22 tokens, 10.4%) and WUyf3 (18 tokens, 8.5%), making up over 40% of the total level observations.

Table 5.2 Distribution of falling and level tokens by SPEAKER

tone shape	% within the tone shape	SPEAKER
falling n=212	0-1.9	WHof5, WHom3, WUym3, WHof4, WHom2, WUom2, WUom3, WUof1, WUom4, Whom1, WHof3, WHom4, WUof4, WUof5, WHom5, WHof1, WUof2, WHyf4, WUyf2, WUyf1, WHyf1, WHym2
	2-3.9	WUof3, WUym5, WHyf2, WUyf3, WUym4, WHym1, WUom5, WHym3
	4-5.9	WUym2, WHyf5, WHof2, WHyf3
	6-7.9	WHym5, WUom1, WUyf5, WUyf4
	8.4	WHym4
	12.1	WUym1
level n=239	0-1.9	WHof5, WHom3, WUym3, WUyf2, WUym4, WHof4, WHom2, WUom2, WUom3, WHom4, WUof4, WUof5, WUof2, WUyf1, WUof3, WUof1, WUom4, WHom5, WUym5, WHom1, WHyf4, WHyf1, WHof3, WHyf2, WUym2, WUyf5
	2-3.9	WUyf3, WHym5, WUyf4, WHym1, WHym4, WHof1, WUom5, WHym2,
	4-5.9	WHof2, WUom1
	6.6	WHyf5
	8.5	WHyf3
	10.4	WUym1
	21.7	WHym3

GSDJPS (1960) and Qián (1992) found the conservative T4 is realized as a level tone /33/. However, according to the small amount of utterances and unbalanced distribution of level tones, it is unlikely that those level tones found in my dataset constitute the conservative variant that GSDJPS (1960) and Qián (1992) defined. So the level tokens are analyzed independently.

Given the unbalanced distribution of falling and level tones among words and speakers, mixed-effects modeling with crossed random effects for words and speakers was conducted on two dependent variables: *WXfalling* and *WXlevel*. *WXfalling* has two variants: falling (n=212) and non-falling observations (n=5162);

WXlevel has two variants as well: level (n=239) and non-level observations (n=5135). For each variable, two models were built, model-0 with only the random factors, and model-AGE with both the random factors and fixed factor AGE. Model-AGE is compared with model-0 to see if it can yield a better fit with the data. If so, it suggests that *WXfalling* and *WXlevel* are involved in ongoing changes. Table 5.3 lists the results of comparison:

Table 5.3 Results of likelihood ratio test on *WXfalling* (falling vs non-falling) and *WXlevel* (level vs non-level)

dependent variable	**model**	**predictor**	***log-odds***	**SE**	**t**	***p***
WXfalling (falling n=212 vs non-falling n=5162)	model-0	Intercept	0.015	0.285	−14.710	0.000***
	model-AGE	Intercept	0.005	0.367	−14.391	0.000***
		AGE (*young*): *old*	8.344	0.451	4.703	0.000***
	$\chi^2(1)=5.241$, $p=0.022^{*}$					
WXlevel (level n=239 vs non-level n=5135)	model-0	Intercept	0.011	0.257	−17.26	0.000***
	model-AGE	Intercept	0.007	0.337	−14.67	0.000***
		AGE (*young*): *old*	2.738	0.425	2.37	0.018*
	$\chi^2(1)=20.402$, $p=0.000^{***}$					

Table 5.3 shows that both *WXfalling* and *WXlevel* are constrained by AGE. The log-odds scores are greater than 1, which means young speakers use significantly more falling and level tones than old speakers. So both *WXfalling* and *WXlevel* are involved in ongoing changes. Further investigation on constraints of these two variables is done in Section 5.2.1.

5.1.2. T4

The first and most straightforward approach to investigate a change in progress is to trace it in apparent time; that is, to assess the distribution of linguistic variables across age levels (Labov, 1994: 45-46). As discussed in Section 3.1.2, geographical variation is also applied in this study as a reflection of change, the suburban Huàzhuāng being more conservative than the urban one. So the mean F_0 contours of T4A and T4B are illustrated in Figure 5.1, split by AGE and REGION. Meanwhile, three phonemic tones have been distinguished for T4 according to the author's transcription: the peaking tones /131/, low rising tone /13/ and dipping tones /213/. The mean F_0 contours are presented separately for those three phonemic tones.

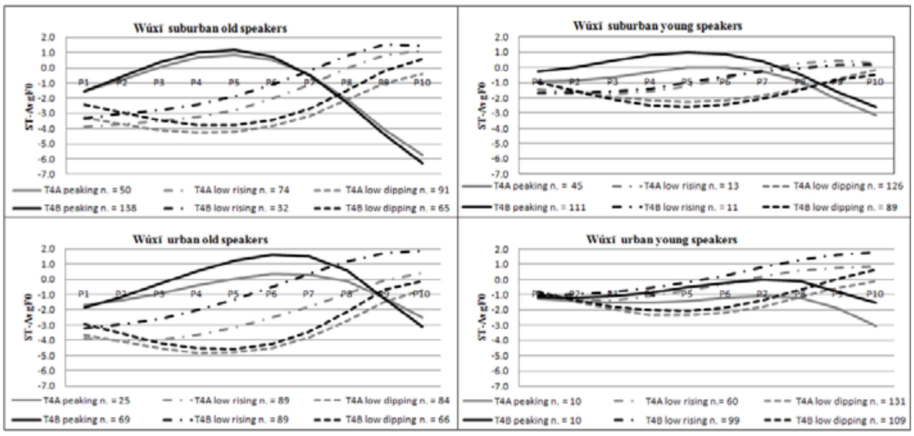


Figure 5.1 Mean F_0 contours of Wúxī T4, split by MC_AB, tone shapes (peaking, low rising and dipping), AGE and REGION

As can be seen from Figure 5.1, peaking tones occurred in all four groups - suburban old group, suburban young group, urban old group and urban young group - as well as in both T4A morphemes and T4B morphemes. Numbers in the legend of Figure 5.1 are the numbers of observations used for the calculation of mean contours. Overall, old speakers have more peaking tokens than young speakers (344 vs 144); suburban speakers have more peaking tokens than urban speakers (282 vs 176). With regard to the contour shape, the mean F_0 contours inevitably hide variation patterns, but they give an overall picture of the differences between the tonal contours and the social groups. As illustrated in Figure 5.1 the old speakers had more prime peaking tokens, in which the low rising part before the peak and the falling part following the peak are both steep. Young speakers have relatively flattened peaking tokens, where the bulging part in the middle is not obvious. From the diachronic perspective, Wúxī T4 is identified as a peaking tone by Chao (1928), and is recorded as a peaking tone /131/, level tone /33/, low rising tone /13/ or low dipping tone /214/ by some dialectological studies mentioned below. Meanwhile, T2 and T6 are also transcribed as low rising tones or low dipping tones in those studies. So based on synchronic evidence, it is concluded that Wúxī T4 is undergoing a process of merging with T2/T6. Qián (1992: 37) documented the merging of T4 with T2 and T6 in the old speakers he interviewed, who were born in the 1920s. This suggests that the T4 merger started before the promotion of PTH since 1956 and therefore it is unlikely to be contact induced. Hence the variation of Wúxī T4 is a kind of Zhuó Shǎng Guī Qù, a phonological change that can be dated back to the MC era in other Chinese dialects such as Mandarin. The peaking token is the conservative variant, while the low rising and dipping tokens are the advanced variants. This variable is called **WXT4**. Figure 5.2 (below) illustrates the distribution of three WXT4 variants, split by AGE and REGION:

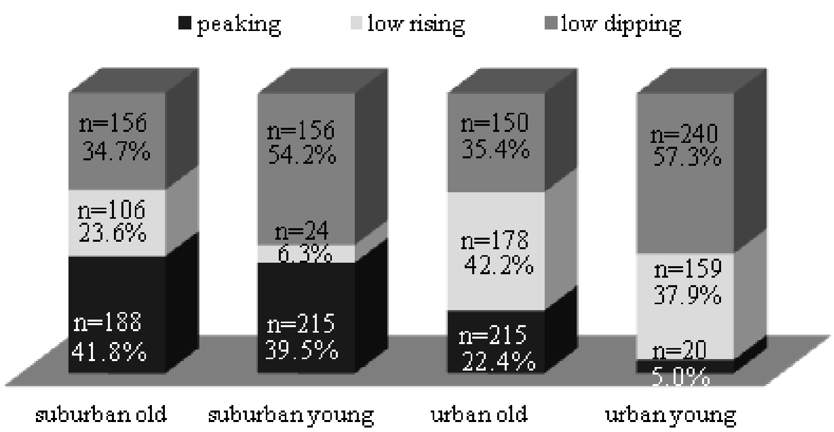


Figure 5.2 The distribution of peaking, low rising and dipping variants in WXT4, split by AGE and REGION

Regarding the distribution of three variants, Figure 5.2 illustrates that in the suburban areas old speakers have more low rising variants but less dipping variants than young speakers, and that the old and young have almost the same amount of peaking variants. This suggests that the low rising variant is changing into a dipping variant in the suburban areas, i.e., low rising > dipping. In the urban areas, old speakers have more peaking but less dipping variants than the young speakers. As old and young people in the urban areas have almost the same amount of low rising variants, it is hard to tell how the peaking T4 becomes dipping. There are two possible paths: (1) the most conservative variant peaking tone changes to dipping tone directly, i.e., peaking > dipping; or (2) the peaking tone first becomes a low rising tone and then the low rising tone changes into a dipping tone, i.e., peaking > low rising > dipping. Figure 5.2 also illustrates that within each age group, urban and suburban speakers do not differ enormously in the proportion of dipping tones, suggesting that the variation between dipping tones and non-dipping ones (low rising + peaking) is constrained by AGE but not by REGION. This dipping variant is possibly the most advanced and externally motivated as it has spread over both the urban and suburban areas at the same speed. Therefore, a hypothesis is proposed that the low rising variant is more conservative than the dipping variant and the possible trajectory of change for peaking tone into dipping tone is the second one, namely "peaking tone > low rising tone > dipping tone". This hypothesis is tested in Section 5.2 by mixed models. Section 5.2 also explains how the peaking variant changes into the low rising and/or dipping variants and whether the WXT4 variable is a gradual or an abrupt change.

5.1.3. T2 and T6

It was unexpected to find peaking tones in the transcription of T2 and T6, as this tone was not recorded in any previous study. As presented in Table 2.10 (Chapter 2), T2 was originally a low rising tone mostly transcribed as /14/ (Chao, 1928; GSDJPS, 1960; and the old variety in Qián, 1992). It then merged with T6 and became /213/ (Chan & Ren, 1989; new variety of Qián, 1992) or /113/ (Xú, 1997). In our dataset, 254 peaking tokens were transcribed, making up 7.2% of the total 3515 tokens of T2 and T6. Among the 254 peaking tokens, 78.3% (199 tokens) occurred in the reading of minimal pairs, but only 3.1% (8 tokens) occurred in the reading of the paragraph and 18.5% in the reading of the wordlist. It indicates that the majority of peaking tokens occurred in the reading of minimal pairs. It is known that reading minimal pairs is a task can increase the speaker's linguistic insecurity and therefore may result in an increase in the use of hypercorrections. A closer look at the peaking tokens in the minimal pairs finds that peaking T2 and T6 in the reading of minimal pairs only exist in pairs ("T2 & T4", "T4 & T6" or "T2, T4, & T6") where a peaking T4 token can also be found. This supports the idea that the speakers borrow peaking T4 in the minimal pairs. In other words, the peaking variants of T2 and T6 are the result of a quite different mechanism. So they will not be analyzed together with the analysis of low rising tokens of T2 and T6 in this dissertation.

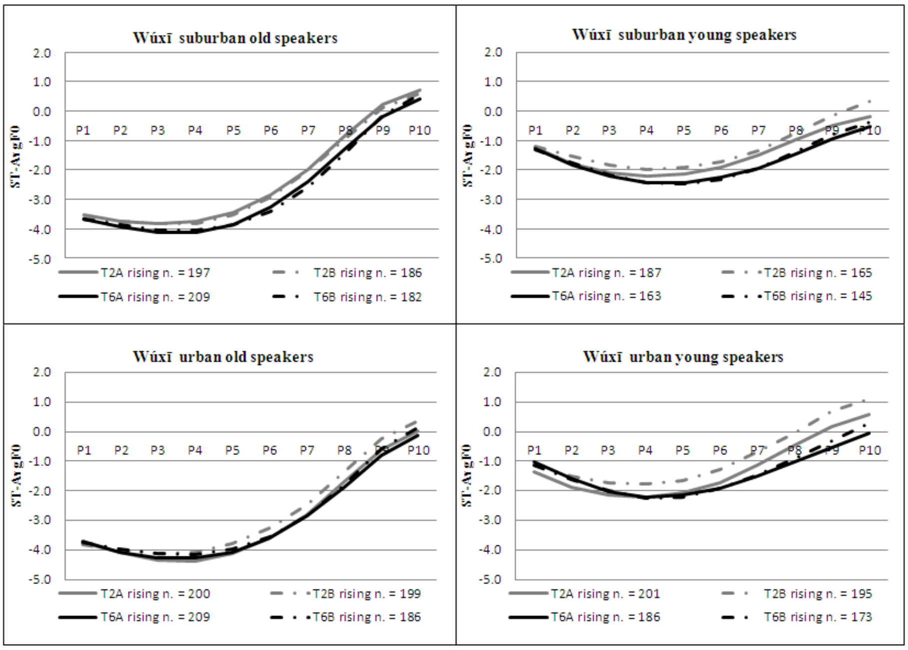


Figure 5.3 Mean F0 contours of the low rising tokens in T2A, T2B, T6A and T6B, split by AGE and REGION

In Figure 5.3, numbers in the legend are the numbers of observations used for the calculation of mean contours. With respect to the low rising T2 and T6, Figure 5.3 illustrates that no obvious difference between urban and suburban areas can be traced from the mean F_0 contours of T2 and T6 but an obvious age difference exists in the shape of low rising contours. Young people tend to use flatter slopes in rising tones than old people. However, young people's starting point and lowest point are at higher pitch heights, making the low rising contour more level so that it looks like a dipping contour. The young generation tends to use a less obvious contrast, i.e., the flattened version, in producing tonal contours. This process is called "contour loss" and is analyzed in Section 5.2.2.

Comparing T2 and T6, Chao (1928) recorded T2 as /13^{b}/ while T6 as /273^{b}/, indicating that, in the 1920s, T2 was a straight low rising tone while T6 was a dipping tone. However, T2 and T6 were merged as a low rising tone 80 years later (Xú, 2007). A slight separation of T2 from T6 in Figure 5.3 can be distinguished in the current dataset. T2 contours have similar starting points to T6 but rise straighter than T6. Figure 5.3 also shows that both T2 and T6 seem to be /113/, either the straight low rising contour (Chao's transcription $/1\underline{3}^{b}/$) or the dipping contour (Chao's transcription $/\underline{\underline{2\dot{7}3}}^{b}$). So it is hypothesized that the merging result of T2 and T6 is neither $/1\underline{3}^{b}/$ nor $/\underline{\underline{2\dot{7}3}}^{b}/$, as transcribed by Chao, but their intermediate status, a delayed low rising tone /113/. The merger not only changes T2 but also T6. After merger, T6 has a more level part before the turning point.

Three types of T2/T6 were found in the data of this study in the process of transcription: the straight low rising tone (/13/), the delayed low rising tone (/113/) and the dipping tone (/213/). However, most tokens are in between. It was hard for the author to make reliable classifications for all the tokens. In other words, the difference between T2 and T6 is too small to be distinguishable to the author's ear. This is an indication of a (near) merger between the very conservative variant of T2 (a straight low rising tone) and T6 (a dipping tone). Our large amount of acoustic data might be useful to give some clues for detecting the constraints for the merger of T2 and T6. Two indexes for separating low rising T2 and dipping T6 are: (1) the position of the P_{min} (i.e., the normalized time of the point having P_{min}, abbreviated to NT_{Pmin}); (2) the pitch difference between the starting point and P_{min} (ST_{P1}-ST_{Pmin}). The later the P_{min} occurs in the entire contour, the more dipping the contour appears. The greater the pitch difference between starting point and P_{min}, the more dipping the contour appears.

Now I attempt to test the differences between T2 and T6 using data from the old speakers. There are two reasons for using the data from old speakers. First, the merger of T2 and T6 is documented as a complete merger. If there are any clues of their historical separation in the synchronic data, these clues are more likely to be found in the old speakers than the young speakers. Second, as discussed above, T6 may have flattened its first falling part during the merger. However, the process of contour loss raises the onset of T2 and T6 to make the whole contour more level but the first falling part more steep, which is opposite to the impact of the merger. Young people show an obvious contour loss, so their data is not suitable for testing the remaining clues about the previous separation of T2 and T6.

Table 5.4 (below) presents the mean and SD of two indexes: the position of P_{min} and the pitch difference between the starting point and P_{min}, split by TONE. The results in Table 5.4 show that T2 has smaller "normalized time of P_{min}" and "ST_{P1}-ST_{Pmin}" than T6, which means T2 has an earlier turning point and a small falling part before the turning point. A new variable having four conditions – T2A, T2B, T6A and T6B – is created, which is called TONE. Then mixed-effect modeling was used to assess the effect of etymological TONE on two indexes (dependent variables). The random factors are SPEAKER and WORD. In the analyses, two models were built for each dependent variable, model-0 with only the random factors, and model-1 with both the random factors and fixed factor TONE. The two models were then compared to test the effect of TONE.

Table 5.4 Mean (M) and standard deviation (SD) of normalized time of P_{min} (NT_{Pmin}) and pitch differences (in semitones) between P_1 and P_{min}

tone	n	NT_{Pmin}		ST_{P1}-ST_{Pmin}	
		M	SD	M	SD
T2A	397	3.245	1.677	0.888	0.964
T2B	385	3.166	1.808	0.710	0.917
T6A	418	3.416	1.792	0.951	0.988
T6B	368	3.497	1.778	0.835	0.817

Then likelihood ratio tests are conducted on NT_{Pmin} and pitch differences between P_1 and P_{min} to see if MC_TONE and MC_AB together can yield a significantly better data fit to model-0. Results for NT_{Pmin} show that the difference between model-0 and model-TONE was also not significant: $\chi^2(3)=3.630$, $p=0.304$. Similarly, results for the pitch differences between P_1 and P_{min} show that the difference between model-0 and model-TONE was not significant either: $\chi^2(3)=3.572$, $p=0.312$. In sum, the acoustic data of old Wúxī speakers shows that there are no significant differences between T2 and T6 either in the position of P_{min} or the shape of falling part before P_{min}. This again supports the notion that T2 and T6 have been completely merged. Thus henceforth I refer to Wúxī T2 and T6 as WXT2/6.

5.1.4. Summary

Four variables are identified in the citation forms of T2, T4 and T6 in Section 5.1. The first variable is *WXfalling*, and the second *WXlevel*. These falling and level tones have rarely been previously documented, but are found occasionally in T2, T4 and T6 morphemes in our dataset. They are mostly used by young speakers. The third variable is WXT4. Three variants of this variable have distinct tonal shapes, in which the peaking shape is the conservative variant while the low rising and dipping variants are advanced. The fourth and final variable is *WXT2/6*, the contour loss occurring in T2 and T6. Results of mixed modelling show that T2 and T6 statistically do not have significant differences in NT_{Pmin} and ST_{P1}-ST_{Pmin}. It supports that Wúxī T2 and T6 have completely merged. So I refer to them as *WXT2/6*. The slope between P_{min} and P_{max} is used for quantifying the contour loss of *WXT2/6*.

5.2. Analyzing linguistic variables of citation tone

5.2.1. WXfalling and WXlevel

In order to further investigate constraints on *WXfalling* and *WXlevel*, logistic models with random effects were built on each binary dependent variable. Six social factors and five linguistic factors were tested by the likelihood ratio test one by one. It is noteworthy here that only one suburban area, i.e. Huàzhuāng was investigated in Wúxī (see Table 3.1). Thus, "suburban (area)" is used to replace "Huàzhuāng" in this chapter to emphasize the contrast between urban and suburban areas. A series of binary logistic regression models with mixed effects were built for each factor: model-0 with only the random factor, and model-X with both the random factor and the fixed factor, in which X denotes the only fixed factor entered in the model. Model-0 and model-X are compared by likelihood ratio test. Table 5.5 shows the results for *WXfalling* and Table 5.6 for *WXlevel*.

Table 5.5 Results of likelihood ratio test on *WXfalling (falling* 212 *vs non-falling* 5162*)*

	deviance	**Chisq**	**Chi Df**	**Pr(>Chisq)**
model-0	1447.2			
model-AGE	1442.0	5.241	1	0.022*
model-REGION	1444.2	3.037	1	0.081
model-SEX	1447.0	0.171	1	0.679
model-EDU	1447.2	0.009	1	0.924
model-PTH	1446.2	1.027	1	0.311
model-STYLE	1444.6	2.593	2	0.273
model-MORPHEME	1430.9	16.344	1	0.000***
model-MC_TONE	1421.1	26.146	2	0.000***
model-MC_AB	1445.8	1.361	1	0.243
model-HEIGHT	1442.2	5.028	5	0.413
model-DUR	1298.6	148.560	1	0.000***
model-PTH_TONE	1420.7	26.468	2	0.000***

Table 5.6 Results of likelihood ratio test on *WXlevel (level* 239 *vs non-level* 5135*)*

	deviance	Chisq	Chi Df	Pr(>Chisq)
model-0	1722.0			
model-AGE	1701.6	20.402	1	0.000***
model-REGION	1721.7	0.253	1	0.615
model-SEX	1722.0	0.001	1	0.971
model-EDU	1720.8	1.199	1	0.274
model-PTH	1720.5	1.455	1	0.228
model-STYLE	1466.5	255.410	2	0.000***
model-MORPHEME	1718.9	3.020	1	0.082
model-MC_TONE	1715.2	6.789	2	0.033*
model-MC_AB	1720.6	1.394	1	0.238
model-HEIGHT	1711.5	10.428	5	0.075
model-DUR	1597.0	124.980	1	0.000***
model-PTH_TONE	1712.5	9.464	2	0.009**

As shown in Table 5.5, AGE, MORPHEME, MC_TONE, DUR and PTH_TONE are significant for the variation of *WXfalling*. Table 5.6 shows that AGE, STYLE, MC_TONE, DUR and PTH_TONE are significant for the variation of *WXlevel*. MC_TONE and PTH_TONE are correlated because MC_T2A and MC_T2B correspond to PT2, MC_T4A corresponds to PT3, and MC_T4B, MC_T6A and MC_T6B corresponds to PT4. In order to decide which one can be entered into the final model, the contingency tables of *WXfalling* and *WXlevel* with the predictors of MC_TONE*MC_AB, i.e., T2A, T2B, T4A, T4B, T6A and T6B were examined to see the distribution of falling and level tokens in Table 5.7.

Table 5.7 Contingency tables of *WXfalling* and *WXlevel* with the predictor of MC_TONE*MC_AB

TONE	***WXfalling***		***WXlevel***	
	n	P	n	P
T2A	11	1.3%	33	4.0%
T2B	14	1.6%	29	3.5%
T4A	**56**	**6.6%**	**53**	**6.2%**
T4B	32	3.5%	28	3.0%
T6A	**63**	**7.1%**	**59**	**6.6%**
T6B	36	4.5%	37	4.7%

The falling tokens of T4A, T6A and T6B morphemes account for around a three-quarter of the total falling tokens. T4A, T6A and T6B all correspond to PT4, the falling tone /51/. So *WX falling* tones are more likely to be found in morphemes corresponding to PT4, probably due to the process of tone borrowing. T2A, T2B and T4B morphemes may also borrow their PTH corresponding tones, but those borrowings are hard to detect because these PTH corresponding tones are either PT2 /35/ or PT3 /214/, having tone shapes similar to WXT2 (/13/) and WXT4 (/13/ or

/213/). Regarding *WXlevel*, when transcribing the data I found people used level tones, which is a kind of ambiguous answer, when they were uncertain about the pronunciation. Therefore, *WXlevel* is also a kind of tone borrowing, borrowing the neutral tone of PTH (轻声, qīng shēng). Consequently, given the causes of *WXfalling* and *WXlevel,* PTH_TONE was selected to build the final model.

Interactions are also investigated by means of a likelihood ratio test. AGE*REGION is found to be significant for *WXfalling* but no interaction is found to be significant for *WXlevel*. Then mixed models were built with main effects and interactions. Table 5.8 and Table 5.9 present the outputs for *WXfalling* and *WXlevel* separately.

Table 5.8 Output of mixed model on *WXfalling* (falling n=212 vs non-falling n=5162)

predictor	***log-odds***	**SE**	**t**	**p**
Intercept	1.024	0.514	0.046	0.963
AGE (*young*): *old*	0.832	0.561	−0.328	0.743
MORPHEME (free): bound	1.885	0.235	2.698	0.007**
PTH_TONE (*PT2*): *PT3*	1.310	0.299	0.903	0.367
PT4	3.247	0.303	3.891	0.000***
DUR	1.016	0.001	10.658	0.000***
AGE*REGION *young*suburban*	0.269	0.521	−2.521	0.012*
*old*suburban*	0.567	0.574	−0.990	0.322

C=0.919, Somer's D_{xy}=0.838

In Table 5.8, the intercept is the usage of falling tokens when AGE=young, MORPHEME=free and PTH_TONE=PT4. The insignificant AGE shows that old people are not significantly different from the intercept in using falling tokens. Changing the current intercept into other possible conditions does not provide any significant AGE effect. It means actually young people do not use more falling tokens than old people if considering the effect of MORPHEME, PTH_TONE, DUR and AGE*REGION. Four significant factors are selected by the mixed model analysis of *WXfalling*. The first factor is MORPHEME: if a morpheme is bound, then the probability of a falling tone on the morpheme is higher than if a morpheme is free as its log-odds is greater than 1. Bound morphemes cannot occur on their own as independent words. They are rarely realized independently by speakers, unless on some special occasion such as dialect courses or linguistic interviews. So, speakers tend to borrow tone, either PTH tone or tone used in sandhi forms, to pronounce bound morphemes. The second factor is that morphemes corresponding to *PT4*, a high falling tone, tend to be realized into falling tones more than those corresponding to *PT2*. Morphemes corresponding to PT3 do not differ from those of

PT2 in the amount of falling tokens. The third factor is DUR: morphemes with a longer duration are more likely to be realized as non-falling tones (mainly rising and dipping tones), which is in agreement with the finding that falling tone has an intrinsically shorter duration than level and rising tones (Howie, 1976; Rose, 1980; Zhu, 1999). The fourth factor is AGE*REGION. The contingency table in Table 5.9 suggests that suburban and urban old people do not differ in the amount of falling tokens, but suburban young speakers use many more falling tokens (113 tokens) than urban speakers (43 tokens), which makes the interaction significant.

Table 5.9 Contingency table of *WXfalling* with the predictors of AGE*REGION

REGION \ AGE	***old***	***young***
urban	28	43
suburban	28	113

The major social factors that are relevant for predicting the effects of contact-induced change are the presence versus the absence of imperfect learning, intensity of contact, and speakers' attitudes (Thomason, 2010). When interviewing the suburban and urban young speakers, the author did not notice that they differed in their proficiency of local dialects. Suburban areas do not have more intensive contact with PTH than urban areas. So the only possible explanation for suburban young people's tone borrowing is that they have greater tendency towards linguistic insecurity about their local dialect. As discussed in Chapter 3, among the young generation the urban-suburban contrast runs more or less parallel with the stratification of socioeconomic status. Suburban young people as a group of "social climbers" tend to adopt the prestige forms of PTH, the common language. In order to obtain information about these issues, a questionnaire was used in the interview, in which questions enquiring language attitudes were included. The full questionnaire appears in Appendix IV. Based on the relevant data, it was found that seven suburban young speakers out of ten thought their local dialect was less pleasant to listen to than PTH while none of them judged their local dialect to be more useful than PTH. In comparison, only four urban young speakers thought their dialect was less pleasant to listen to while three speakers thought their urban dialect was more useful than PTH.

Table 5.10 Output of mixed model on *WXlevel* (level n=239 vs non-level n=5135)

predictor	*log-odds*	SE	t	p
Intercept	0.128	0.523	−3.929	0.000***
AGE (*old*): *young*	5.664	0.438	3.962	0.000***
PTH_TONE (*T2*): *T3*	0.803	0.240	−0.911	0.362
T4	1.823	0.177	3.387	0.001***
DUR	0.991	0.001	−6.332	0.000***
STYLE *(word list): paragraph*	1.328	0.232	−1.224	0.221
minimal pairs	0.056	0.302	−9.541	0.000***

C=0.913, Somer's D_{xy}=0.827

Table 5.10 shows the final mixed model on *WXlevel* contains four significant factors. First, AGE: young people tend to use more level tones than old people. Second, PTH_TONE: like *WXfalling*, morphemes corresponding to PT4 tend to be realized into level tones more than those corresponding to PT2. Morphemes corresponding to PT3 do not differ from those of PT2 in the amount of level tokens. Third, DUR: morphemes with a longer duration are less likely to be realized as level tones than non-level tones (mainly rising and dipping tones), which is also in agreement with the previous finding that level tones have intrinsically shorter durations than rising tones (Howie, 1976; Rose, 1980; Zhu, 1999). Finally, STYLE: morphemes in the reading of minimal pairs have much less probability of being realized as level tones than in the reading of a wordlist. Level tokens mainly occur in the reading of wordlists and paragraphs. As mentioned before, speakers tend to use level tone, the neutral one, when they are uncertain about its dialect pronunciation. Level tone occurs less in minimal pairs because people use another way to resolve their uncertainty, borrowing the tone of other morphemes in the same pair.

5.2.2. WXT4

WXT4 is a variable with the transcribed tonal shape as its three levels. This variable has three variants: peaking tone (conservative variant), low rising tone and dipping tone. In the dataset of 1686 total valid observations, there are 458 cases of peaking tone, 467 cases of low rising tone and 761 cases of dipping tone.

The predictors that represent linguistic, social and stylistic variables are included in the statistical model. As discussed in Chapter 3, six social factors (AGE, REGION, SEX, PTH, EDU and STYLE) and six linguistic factors (MORPHEME, MC_TONE, MC_AB, HEIGHT, DUR and PTH_TONE) are the predictors for the variation of citation tone in this study. Incidentally, MC_TONE is not a predictor for variable *WXT4* as all its tokens come from MC_T4. As discussed in Section 5.1.2, the merger of T4 with T2 and T6 is not contact induced; hence PTH_TONE is not used as a predictor for *WXT4*. EDU originally has three conditions: primary school, middle school and college. Since just three Wúxī speakers self-reported that they had only primary education, the primary school and middle school were further combined into a new condition: non-college. In analyzing the Wúxī data, EDU has two conditions: college and non-college education. The predictor conditions of other social and stylistic predictors are referred to in Section 3.1.6 and Section 3.4.7. The predictor conditions of other linguistic variables can be found in Section 3.4.7. The statistical model used for investigating WXT4 is a mixed-effects model. For variable WXT4, the random effects are SPEAKER (40 speakers) and WORD (33 words).

5.2.2.1. Peaking variant of WXT4

As discussed in Section 5.1.2, the peaking variant is undoubtedly the conservative variant. In order to further investigate whether WXT4 is a gradual change or an abrupt change, the peaking variant is first explored in more detail. In Figure 5.1 it can be observed that old and young speakers seem to have acoustic differences in the production of this conservative variant. I will first look at the exact nature of these differences and then quantitatively analyze them.

First, linear mixed models were built on P_{max} of the peaking tones, i.e., ST_{Pmax}. Six social factors and five linguistic factors are tested by the likelihood ratio test one by one. Two models were built for each factor: model-0 with only the random factor, and model-X with both the random factor and the fixed factor, in which X denotes the only fixed factor entered in the model.

Table 5.11 Results of likelihood ratio test on ST_{Pmax} of peaking WXT4 (n=458)

	deviance	**Chisq**	**Chi Df**	**Pr(>Chisq)**
model-0	**1853.9**			
model-AGE	1849.7	4.252	1	0.039*
model-REGION	1853.7	0.279	1	0.597
model-SEX	1853.5	0.436	1	0.509
model-EDU	1851.3	2.582	1	0.108
model-PTH	1851.5	2.454	1	0.117
model-STYLE	1852.4	1.562	2	0.458
model-MORPHEME	1851.9	2.071	1	0.150
model-MC_AB	1849.5	4.445	1	0.035*
model-HEIGHT	1844.0	9.899	4	0.042*
model-DUR	1851.0	2.989	1	0.084

As can be seen from Table 5.11, AGE, MC_AB, HEIGHT and PTH_TONE are the three significant main effects selected by likelihood ratio test. Interactions related to significant main effects are investigated one by one in the likelihood ratio test as well. Results show that no interactions can yield a significantly better model than the model with only significant main effects. Thus only AGE, MC_AB and HEIGHT were entered into the final model for ST_{Pmax} of the T4 peaking tones. Table 5.12 presents the output.

In the final model, only AGE and HEIGHT are significant while MC_AB hovers around significance. The output shows that young people have significantly lower (M=0.76 ST) ST_{Pmax} than old people (M=1.62 ST), which indicates that the ST_{Pmax} of the peaking tone is undergoing changes. The advanced T4 peaking has a lower peak and then looks level. The T4 peaking variant is also undergoing the contour loss process. The significance of HEIGHT suggests that if the peaking tone is carried by an open vowel, i.e., low vowel, it has significantly lower ST_{Pmax} than those carried by a close vowel. Although the three other conditions - *near close, close mid* and *open mid* - do not have significant differences from close vowels, they show the trend that open vowels go with low pitch (open: M=0.584 ST, open mid: M=0.946 ST, close mid: M=1.494 ST, near close: M=1.646 ST, close: M=1.887 ST). So it can be concluded that the effect of HEIGHT comes from the intrinsic vowel effect. The nearly significant factor MC_AB shows that T4B morphemes (M=1.51 ST) have higher ST_{Pmax} than T4A morphemes (M=0.75 ST), indicating that T4B morphemes are more conservative.

Table 5.12 Output of mixed models on ST_{Pmax} of peaking WXT4 (n=458)

predictor	*β*	SE	t	p
Intercept	1.807	0.431	4.194	0.000***
AGE (*old*): *young*	−1.019	0.487	−2.093	0.046*
MC_AB *(T4A): T4B*	0.519	0.263	1.976	0.078
HEIGHT *(close):*				
near close	0.016	0.504	0.032	0.975
close mid	−0.289	0.384	−0.752	0.473
open mid	−0.619	0.345	−1.793	0.120
open	−1.115	0.358	−3.113	0.015*

We now examined whether the position of P_{max} on the peaking contour, i.e., NT_{Pmax}, is undergoing any changes. Linear mixed models were built for each fixed factor and were compared with random-effect-only model by likelihood ratio tests. The results are listed in Table 5.13.

Table 5.13 Results of likelihood ratio test on NT_{Pmax} of peaking WXT4 (n=458)

	deviance	Chisq	Chi Df	Pr(>Chisq)
model-0	1639.3			
model-AGE	1639.2	0.079	1	0.779
model-REGION	1613.9	25.463	1	0.000***
model-SEX	1638.0	1.329	1	0.249
model-EDU	1634.7	4.592	1	0.032*
model-PTH	1637.5	1.864	1	0.172
model-STYLE	1627.1	12.242	2	0.002**
model-MORPHEME	1638.4	0.948	1	0.330
model-MC_AB	1639.1	0.204	1	0.651
model-HEIGHT	1637.6	1.762	4	0.780
model-DUR	1636.6	2.690	1	0.101

Table 5.13 shows that REGION, EDU and STYLE can yield significantly better models than model-0. Next, two-way interactions of REGION with other fixed factors, EDU with other fixed factors, and STYLE with other fixed factors were investigated one by one. The results show that only the interaction of REGION and SEX yields a better model than the model with only three significant main effects. Then, the three main factors and one interaction were entered into the final mixed model for $NT_{Pmax.}$

Table 5.14 Output of mixed models on NT_{Pmax} of peaking WXT4 (n=458)

predictor	*β*	SE	t	p
Intercept	5.173	0.279	18.542	0.000***
REGION *(suburban): urban*	1.637	0.331	4.951	0.000***
EDU *(college): non-college*	−0.088	0.221	−0.396	0.695
STYLE *(wordlist): paragraph*	1.237	0.440	2.810	0.008**
minimal pair	0.341	0.141	2.425	0.016*
REGION*SEX *suburban*male*	−0.060	0.244	−0.245	0.809
*urban*male*	−0.859	0.360	−2.383	0.022*

The output of Table 5.14 suggests that the peak of the peaking contour was moved more towards the end in the urban (M=6.18) than in the suburban areas (M=4.99). It indicates that the peaking contour with a late peak is the advanced variant. The significant interaction of REGION and SEX further suggests that in urban areas, females (M=6.56) have much later peaks than males (M=5.80), so urban females are more advanced; while suburban females (M=5.04) do not differ from suburban males significantly (M=4.94). Regarding STYLE, people tend to produce a peaking contour with an early peak when reading wordlists (M=5.07), then an average peak for minimal pairs (M=5.44), and a significantly late peak when reading paragraphs (M=5.97). The difference between wordlist reading and paragraph reading is significant, so is the difference between wordlist reading and minimal pair reading. Reading paragraphs is a less monitored style. Using the late peak in this style suggests that pronouncing a peaking tone with a late peak is a way of saving articulatory efforts.

The dependent variables ST_{P1} and ST_{P10} were further examined. Once again, linear mixed models were built on ST_{P1}. The results in Table 5.15 below show that AGE, STYLE and DUR are significant. No significant interaction was found by further mixed modeling. Then, a final model only with main effects was built. The output is listed in Table 5.16:

Table 5.15 Results of likelihood ratio test on ST_{P1} of peaking WXT4 (n=458)

	deviance	**Chisq**	**Chi Df**	**Pr(>Chisq)**
model-0	1800.6			
model-AGE	1793.8	6.777	1	0.009**
model-REGION	1798.2	2.411	1	0.121
model-SEX	1800.6	0.027	1	0.871
model-EDU	1798.5	2.122	1	0.145
model-PTH	1800.6	0.003	1	0.959
model-STYLE	1791.0	9.609	2	0.008**
model-MORPHEME	1800.0	0.656	1	0.418
model-MC_AB	1799.9	0.729	1	0.393
model-HEIGHT	1796.8	3.780	4	0.437
model-DUR	1780.8	19.823	1	0.000***

Table 5.16 Output of mixed models on ST_{P1} of peaking WXT4 (n=458)

predictor	***β***	**SE**	**t**	**p**
Intercept	−0.335	0.615	−0.545	0.587
AGE (*old*): *young*	0.970	0.448	2.168	0.038*
STYLE *(minimal pair):*				
paragraph	−0.685	0.583	−1.175	0.243
wordlist	0.080	0.199	0.401	0.688
DUR	−0.005	0.002	−3.056	0.002**

In the final model, ST_{P1} is constrained by AGE, which suggests that it is an ongoing change. The young have significantly higher ST_{P1} (M=−0.53) than the old (M= −1.63). For every millisecond increase in duration, ST_{P1} is reduced by 0.005 semitones. STYLE is not significant in the final model, although different intercepts were tried. Then linear mixed models are built on ST_{P10}.

Table 5.17 Results of likelihood ratio test on ST_{P10} of peaking WXT4 (n=458)

	deviance	Chisq	Chi Df	Pr(>Chisq)
model-0	2212.3			
model-AGE	2206.6	5.693	1	0.017*
model-REGION	2203.1	9.252	1	0.002**
model-SEX	2212.2	0.134	1	0.714
model-EDU	2211.5	0.859	1	0.354
model-PTH	2211.0	1.315	1	0.252
model-STYLE	2187.7	24.589	2	0.000***
model-MORPHEME	2209.6	2.774	1	0.096
model-MC_AB	2210.3	1.964	1	0.161
model-HEIGHT	2210.1	2.209	4	0.697
model-DUR	2142.5	69.832	1	0.000***

As can be seen in Table 5.17, AGE, REGION, STYLE and DUR were selected by likelihood ratio tests as the significant constraints. AGE*MC_AB, REGION*AGE are found to be significant interactions in the likelihood ratio tests. They were entered into the final model with four main effects. But they are both not significant in the final model. So they are removed again. The output of mixed model built with four main effects is listed in Table 5.18.

Table 5.18 Output of mixed models on ST_{P10} of peaking WXT4 (n=458)

predictor	*β*	SE	t	p
Intercept	3.262	0.775	4.209	0.000***
REGION (*urban*): *suburban*	−2.615	0.643	−4.067	0.000***
AGE (*young*): *old*	−1.905	0.644	−2.959	0.006**
STYLE *(wordlist)*: *paragraph*	−0.144	0.271	−0.530	0.596
minimal pair	0.493	0.506	0.974	0.333
DUR	−0.016	0.002	−6.785	0.000***

Like ST1, STYLE shows significances in the likelihood ratio tests but in the final model its significances are covered by other factors. STYLE is not significant with other intercepts. In the final model, only REGION, AGE and DUR are significant. Regarding REGION, urban people have a much higher offset (M=−2.72 ST) than suburban people (M=−4.63 ST). Regarding AGE, young people have a higher offset (M=−2.62) than old people (M=−5.09 ST). The peaking variant is realized with a higher offset by young and urban people, so it can be concluded that higher offset means more advanced. For DUR, similar with P_1, for every millisecond increase in duration, ST_{P10} is reduced by 0.018 ST.

So far, we have got an overview of the variation of WXT4 peaking tones. The pitch height of P_1, P_{10} and P_{max} and the normalization time of P_{max} are constrained by AGE, REGION or both. So they are all undergoing changes. The variation of WXT4 peaking tones is reflected in (1) reducing the pitch height of P_1 and P_{max}, (2) moving

P_{max} towards the rear of the contour, and (3) raising P_{10}. This process is sketched in Figure 5.4:

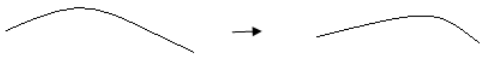


Figure 5.4 Stylized representation of the variation of WXT4 peaking tone

In Figure 5.4, the left contour is the conservative peaking tone mainly used by suburban speakers and old speakers, while the right contour is the advanced peaking tone mainly used by urban speakers and young speakers. It provides an important clue about how the peaking tone develops. The distinction of low rising tone and peaking tone lie in the falling part after the peak. The peaking tone has an obvious falling part while the low rising tone does not. Comparing with the left contour in Figure 5.4, the falling part of the right contour becomes less distinguishable from the perspective of the degree of convexity. The following section, 5.2.2.2, tests this hypothesis in the low rising and dipping variants.

5.2.2.2. Low rising and dipping variants of WXT4

To determine whether the trajectory of change is “peaking tone > low rising tone > dipping tone” or “peaking tone > dipping tone”, the location of three points are crucial: P_{min} and P_{max} in the low rising variant and P_{min} in the dipping variant. If WXT4 is involved in a gradual change through the trajectory “peaking tone > low rising tone > dipping tone”, its conservative low rising tokens should resemble the most conservative variant - the peaking tone - with early P_{max} while its advanced low rising tokens should be close to the advanced dipping variant with late P_{min}. The conservative dipping tokens should have earlier P_{min} than the advanced tokens. Therefore, three linear regression models with mixed effects are built for these three dependent variables, namely the normalized time of P_{min} (NT_{Pmin}), P_{max} (NT_{Pmax}) in the low rising WXT4 and P_{min} (NT_{Pmin}) in the dipping WXT4. The modeling methods and fixed factors for these three variables are exactly the same as those for variables of WXT4 peaking variant (see Section 5.2.1.1).

First, the position of P_{min} on the rising contour, i.e., NT_{Pmin}, is examined to see if it is undergoing any changes. Linear mixed models were built for 11 fixed factors one by one and were compared with the random-effect-only model, model-0. The results are listed in Table 5.19.

Table 5.19 Results of likelihood ratio test on NT_{Pmin} of low rising WXT4 (n=467)

	deviance	Chisq	Chi Df	Pr(>Chisq)
model-0	360.50			
model-AGE	355.88	4.613	1	0.032*
model-REGION	356.11	4.383	1	0.036*
model-SEX	360.37	0.124	1	0.725
model-EDU	359.09	1.408	1	0.235
model-PTH	360.29	0.204	1	0.651
model-STYLE	358.27	2.229	2	0.328
model-MORPHEME	358.18	2.312	1	0.128
model-MC_AB	352.92	7.577	1	0.006**
model-HEIGHT	356.83	3.667	4	0.453
model-DUR	358.82	1.681	1	0.195

Table 5.19 shows that AGE, REGION and MC_AB are the significant main effects for NT_{Pmin} of low rising WXT4. Then interactions related to significant main effects are individually investigated in the likelihood ratio test. Results show that the interaction between REGION*MC_AB yields a significantly better model than the model with only the significant main effects. So only AGE, REGION, MC_AB and REGION*MC_AB are entered into the final model for NT_{Pmin} of the T4 low rising tones. Table 5.20 presents the output:

Table 5.20 Output of mixed models on NT_{Pmin} of low rising WXT4 (n=467)

predictor	β	SE	t	p
Intercept	1.149	0.068	16.854	0.000***
AGE (*old*): *young*	0.140	0.057	2.439	0.024*
REGION (*suburban*): *urban*	0.287	0.083	3.458	0.001**
MC_AB *(T4A)*: *T4B*	0.011	0.105	0.104	0.917
REGION*MC_AB: *urban*T4B*	−0.298	0.117	−2.546	0.011*

Results in Table 5.21 show that in the low rising WXT4, young people have smaller NT_{Pmin} (M=1.23) than suburban people (M=1.36), and urban people have smaller NT_{Pmin} (M=1.18) than suburban people (M=1.31). This supports the notion that young and urban people are more advanced, because in the low rising variant, its advanced tokens bear a resemblance to the dipping variant. Significant REGION*MC_AB suggests that T4A and T4B morphemes differ greatly in urban areas but much less in the suburban area. Table 5.21 presents this interaction by the mean NT_{Pmin} in the low rising variation of WXT4, split by REGION and MC_AB.

Table 5.21 Mean NT_{Pmin} in the low rising variation of WXT4, split by REGION and MC_AB

REGION \ MC_AB	*T4A*	*T4B*
urban	1.50	1.22
suburban	1.16	1.21

Turning to the NT_{Pmax} in the low rising WXT4, neither AGE nor REGION is a significant main effect for its variation based on the results of the likelihood ratio test (AGE: p=0.928; REGION: p=0.145), so the variation of NT_{Pmax} is out of our interests and will not be discussed further.

We now examine the position of P_{min} on the dipping contour, i.e., NT_{Pmin} of dipping WXT4. Table 5.22 presents the results of the likelihood ratio test. As shown in Table 5.22, AGE is the sole significant main effect selected by likelihood ratio test. Further examinations on interaction found that AGE*STYLE is significant. Then the mixed modeling of NT_{Pmin} of dipping WXT4 was conducted and yielded the output in Table 5.23.

Table 5.22 Results of likelihood ratio test on NT_{Pmin} of dipping WXT4 (n=761)

	deviance	**Chisq**	**Chi Df**	**Pr(>Chisq)**
model-0	3402.5			
model-AGE	3389.4	13.082	1	0.000***
model-REGION	3402.3	0.209	1	0.648
model-SEX	3401.9	0.595	1	0.440
model-EDU	3402.5	0.041	1	0.839
model-PTH	3400.6	1.960	1	0.162
model-STYLE	3398.0	4.512	2	0.105
model-MORPHEME	3401.5	1.049	1	0.306
model-MC_AB	3402.5	0.050	1	0.824
model-HEIGHT	3399.2	3.297	4	0.509
model-DUR	3401.5	1.037	1	0.309

Table 5.23 Output of mixed models on NT_{Pmin} of dipping WXT4 (n=761)

predictor	***β***	**SE**	**t**	**p**
Intercept	3.886	0.184	21.166	0.000***
AGE (*old*): *young*	0.775	0.230	3.370	0.002**
AGE*STYLE: *old*paragraph*	−0.139	0.302	−0.461	0.646
*young*paragraph*	−0.768	0.285	−2.698	0.008**
*old*wordlist*	−0.380	0.182	−2.092	0.037*
*young*wordlist*	0.012	0.161	0.074	0.941

Table 5.23 shows that young people have greater NT_{Pmin} in the dipping WXT4 than old people. That is to say, the dipping tones of young people are more concave than those of old speakers. Meanwhile, the significant AGE*STYLE suggests that young speakers have earlier P_{min} when reading paragraphs than reading minimal pairs; while old speakers have earlier P_{min} when reading wordlists than reading minimal pairs, see Table 5.24.:

Table 5.24 Mean NT_{Pmin} in the dipping WXT4 split by AGE and STYLE

STYLE \ AGE	*old*	*young*
wordlist	3.64	4.62
paragraph	4.09	4.06
minimal pair	4.04	4.70

Table 5.24 shows that the age differences NT_{Pmin} are negligible in the reading of paragraphs, but are considerable in the reading of wordlists and minimal pairs. This may be interpreted as follows. The Pīnyīn system uses diacritics to mark the four tones of PTH, which draws people's attention to the citation tone of each monosyllable. The Pīnyīn system was introduced in primary school around 1960 when most old participants of the present study had already completed their primary education. Since very few of the old participants know the Pīnyīn system, they are less aware of the PTH tone than young people. All the young participants had been taught Pīnyīn system since their primary school and therefore know the PTH tone of each monosyllable. When the young people are uncertain about the dialect pronunciation, they are prone to borrow the PTH tone. Regarding WXT4, the corresponding PTH tone of T4A morphemes is falling tone /51/ but that of T4B morphemes is a low dipping tone /214/. Young people are probably influenced by PTH tone when pronouncing T4B morphemes, causing a late turning point in the dipping T4, i.e., a greater $NT_{Pmin.}$ If they are not influenced by PTH, T4B morphemes should be more conservative than T4A and have a smaller NT_{Pmin}.

Table 5.25 Mean NT_{Pmin} in the dipping WXT4 (n=761) split by AGE, STYLE and MC_AB

	old		young	
	T4A	T4B	T4A	T4B
wordlist	3.71	3.50	4.54	4.78
paragraph	4.00	4.10	3.83	4.12
minimal pair	4.27	3.89	4.63	4.75

Table 5.25 shows that young people have greater NT_{Pmin} for T4B morphemes than T4A morphemes. However, old people have nearly the same or even smaller NT_{Pmin} for T4B morphemes than T4A morphemes. This confirms the hypothesis that young people are influenced by PT3 when pronouncing T4B morphemes. This finding also suggests that although WXT4 is an internal change, it is unavoidably influenced by the contact with PTH. It is externally motivated in the late stage when WXT4 changed to dipping tone.

However, borrowing of PT3 can only partly explain young people's high NT_{Pmin} in the reading of wordlist and minimal pairs. The results for T4A presented in

Table 5.25 show that old people have greater NT_{Pmin} in paragraph reading, which supports the hypothesis of articulatory effort. However, young people have differences in the opposite direction: greater NT_{Pmin} in the wordlist. The temporary explanation is that young people might be merging T4 with T2/6 since the mean NT_{Pmin} of young people in T2/6 is 4.44, which is very close to the mean NT_{Pmin} of young people in dipping T4. In sum, young people delay P_{min} in the dipping T4 by (1) reduction of articulatory effort, (2) borrowing PT3 and (3) merging with T2/6.

To summarize Sections 5.2.2.1 and 5.2.2.2, acoustic evidence shows that the low rising variant is more conservative than the dipping tone. The low rising variant is a crucial bridge between the peaking and dipping variants. The peaking variant took four steps to become a dipping variant. Figure 5.5 sketches this process. The peaking tone shortens its falling part after P_{max} to become a direct rising tone (step 1 & 2), and then the low rising tone raises its onset and lengthens its falling part before P_{min} to become the dipping tone (step 3 & 4). Changing from step 2 to step 3 is caused by a natural coarticulation. As discussed in Section 4.1.1, it is common for a perceptually rising tone to have a falling trend at the beginning. When pronouncing a rising tone, it is easier to start from a default onset and then go down a bit to reach P_{min}. On the one hand, this step can minimize speakers' articulatory effort; on the other hand, it is phonetically easier to be assimilated to the dipping tone as its P_{min} has started being pushed backward. It is noteworthy that some speakers, especially urban young speakers, do not undergo the four steps completely. They may skip one or two steps and change from peaking to dipping tone abruptly. For instance, speaker WHym3, WHym4 and WHyf2 only have peaking and dipping tokens, but not the intermediate ones.

Figure 5.5 Stylized representation of the variation of WXT4

There are two possible interpretations for changing from peaking tone to rising tone in the first three steps. First, it may be to simplify the tonal inventory since step 3 is a low rising tone, having the same tonal shape as T2/6. Second, it could be to minimize articulatory effort, because pronouncing peaking tones demands more articulatory effort than pronouncing rising tones. Our data in Table 5.14 also supports that moving the peaking point P_{max} towards the end of peaking contour is economic articulatory. So the changing process from peaking to rising tones is to make the production easier and thus it can be concluded to be internally motivated. However, step 4, the process of changing from a low rising tone to a dipping tone, seems to be contrary to this, as it is changing from a rising tone to a more complex falling rising tone. It is found that this process of changing from a low rising tone to a dipping tone is not internally motivated but contact-induced. The dipping tone is caused by the assimilation of PTH T3 /214/.

5.2.2.3. The regression outputs of WXT4

After the separate investigations into each variant, complete regression models were conducted on WXT4. Since this variable has three variants, two binary logistic regression models with mixed effects were built separately on two binary dependent variables: model A with the dependent variable "peaking tone vs non-peaking tone" (464 vs 1222 tokens), where non-peaking tone is a combination of low rising tone and dipping tone; and model B with the dependent variable "dipping tone vs non-dipping tone" (920 vs 766 tokens), where the non-dipping tone is a combination of peaking tone and low rising tone. We now introduce and compare model A and model B. For each model, the procedure of model building is the same with previous mixed modeling procedures. The random effects are SPEAKER (40 speakers) and WORD (33 words). The differences in mixed effect logistic regressions are conducted by the function "glmer" in R due to the binary independent variables. Here a single multinomial variable was analyzed as two binomial variables to avoid difficulties to run multinomial logistic multilevel models through R because an open source is not available when this study was conducted. It it also reported that it is a standard way to run a multinomial model as series of binomial contrasts[33].

Table 5.26 and Table 5.27 present the results of likelihood ratio test on "peaking tone vs non-peaking tone" and "dipping tone vs non-dipping tone" of WXT4. They are the preliminary to the building of final model A and model B. Results show that AGE, REGION, STYLE, MORPHEME and MC_AB are the significant main effects for the variation between peaking tone and non-peaking tones. For the variation between dipping tone and non-dipping tone, the same five

[33] See the discussion in http://stackoverflow.com/questions/21082396/multinomial-logistic-multilevel-models-in-r.

main effects are found to be significant; moreover, predictor PTH, i.e., whether participants are able to speak PTH, is also a significant main effect.

Table 5.26 Results of likelihood ratio test on “peaking tone (n=458) vs non-peaking tone (n=1228)” of WXT4

	deviance	Chisq	Chi Df	Pr(>Chisq)
model-0	1458.9			
model-AGE	1455.0	3.876	1	0.049*
model-REGION	1432.1	26.765	1	0.000***
model-SEX	1458.7	0.162	1	0.688
model-EDU	1458.8	0.039	1	0.844
model-PTH	1457.6	1.286	1	0.257
model-STYLE	1451.6	7.242	2	0.027*
model-MORPHEME	1444.8	14.076	1	0.000***
model-MC_AB	1453.5	5.340	1	0.021*
model-HEIGHT	1453.8	5.066	4	0.281
model-DUR	1382.9	75.995	1	0.000

Table 5.27 Results of likelihood ratio test on “dipping tone (n=761) vs non-dipping tone (n=925)” of WXT4

	deviance	Chisq	Chi Df	Pr(>Chisq)
model-0	1841.8			
model-AGE	1826.7	15.133	1	0.000***
model-REGION	1838.4	3.413	1	0.065
model-SEX	1841.6	0.213	1	0.645
model-EDU	1841.6	0.171	1	0.679
model-PTH	1837.3	4.511	1	0.034*
model-STYLE	1793	48.867	2	0.000***
model-MORPHEME	1828.5	13.281	1	0.000***
model-MC_AB	1837	4.767	1	0.029*
model-HEIGHT	1835.5	6.265	4	0.180
model-DUR	1696.7	145.080	1	0.000***

Next, related interactions to main effects are investigated by means of a likelihood ratio test. TONE*AGE, REGION*STYLE and REGION*MORPHEME are found to be significant for the variation between peaking tone and non-peaking tones. STYLE*PTH is found to be significant for the variation between dipping and non-dipping tones. Model A and model B are then built on the basis of selected significant main effects and interactions. Table 5.28 presents the outputs of mixed effect logistic regression:

Table 5.28 Output of the mixed effect logistic regression on WXT4

dependent variable	Model A: "peaking vs non-peaking" (n=458 vs n=1228)				Model B: "non-dipping vs dipping" (n=761 vs n=925)			
predictor	***log-odds***	**SE**	**t**	***p***	***log-odds***	**SE**	**t**	***p***
Intercept	0.059	0.752	−3.758	0.000^{***}	0.011	0.742	−6.109	0.000^{***}
MC_AB *(A): B*	0.285	0.483	−2.599	0.009^{**}	0.435	0.365	−2.280	0.023^{*}
REGION *(suburban):*								
urban	4.024	0.529	2.629	0.009^{**}	2.231	0.356	2.257	0.024^{*}
AGE (*old*): *young*	3.570	0.502	2.534	0.011^{*}	5.764	0.380	4.610	0.000^{***}
STYLE *(minimal pair):*								
paragraph	5.441	0.582	2.910	0.004^{**}	25.216	0.742	4.350	0.000^{***}
wordlist	1.561	0.276	1.615	0.106	2.352	0.425	2.013	0.044^{*}
MORPHEME *(bound):*								
free	0.101	0.492	−4.657	0.000^{***}	0.214	0.375	−4.111	0.000^{***}
DUR	1.015	0.002	9.305	0.000^{***}	1.014	0.001	10.597	0.000^{***}
PTH *(cannot speak PTH):*								
can speak PTH	NA				2.842	0.606	1.724	0.085
MC_AB *AGE: *B*young*	2.637	0.331	2.930	0.003^{**}	NA			
REGION*STYLE								
*urban*paragraph*	0.511	0.610	−1.099	0.272	NA			
*urban*wordlist*	2.502	0.351	2.609	0.009^{**}	NA			
REGION*MORPHEME								
*urban*free*	4.683	0.374	4.133	0.000^{***}	NA			
STYLE*PTH								
*paragraph*yes*	NA				0.118	0.680	−3.141	0.002^{**}
*wordlist*yes*	NA				0.375	0.430	−2.280	0.023^{*}
	C=0.933, Somer's D_{xy}=0.866				C=0.896, Somer's D_{xy}=0.793			

The significant social constraints for both models include REGION, AGE, and STYLE. Urban people and young people are more likely to use the advanced variants, i.e. non-peaking tone in Model A and dipping tone in model B. It is easy to understand that young and urban people tend to use more advanced variants. For the predictor STYLE, non-peaking is more likely to occur in the reading of paragraphs while dipping tones are more likely to occur in the reading of paragraphs and wordlists than minimal pairs. The reading of paragraphs and wordlists requires less conscious self-monitoring, and therefore it is also reasonable that speakers tend to use more advanced variants when they pay less attention to their own speech.

The significant linguistic constraints for both models include TONE, MORPHEME and DURATION. T4A morphemes, bound morphemes and morphemes with longer duration are more likely to be realised in advanced variants. The effects of duration will not be discussed because it is not the focus of this study (see Section 3.4.7). It was put in the model because its effects needs to be controlled to prevent omitted variable bias. For two other linguistic constraints, it is quite interesting to find that T4A morphemes and bound morphemes are leading the change.

As discussed in Section 2.3.3, the Wú type of Zhuó Shǎng Guī Qù merges T4A and T4B morphemes together into T6. It is undoubted that WXT4 belongs to Wú type as the data shows T4A and T4B morphemes are both changing into the same dipping tone. Meanwhile, the new finding that T4A morphemes are leading the change provides us some new perspectives to understand Wú type. In the Mandarin type of Zhuó Shǎng Guī Qù, T4A morphemes merged with T3 while T4B morphemes merged with T6. Meanwhile the majority of Wú dialects have finished the merger of T4 and T6 so it is unknown whether T4A and T4B morphemes in the Wú dialects merged with T6 synchronously. The finding of this study is that T4B morphemes changes more slowly than T4A. The principal differences of T4A and T4B morphemes are their initial consonants. T4A morphemes are initialed with a voiceless obstruent plus a voiced glottal fricative /ɦ/, while T4B morphemes are initialed with voiced sonorants. It can be temporarily concluded that the stability of T4B morphemes in Zhuó Shǎng Guī Qù is correlated with its sonorant initials.

Regarding MORPHEME, free morphemes are more stable while bound morphemes undergo more changes. When bound morphemes were read in the isolated form, it is possible that speakers borrowed the tone of the sandhi form (a low level tone /22/) or the tone of other morphemes that have the same segment in it.

Besides the same five constraints, model A shows that the variation of peaking and non-peaking tones is also constrained by the interaction of TONE*AGE, REGION*STYLE and REGION*MORPHEME, while model B shows that the variation of dipping and non-dipping tones is further constrained by PTH and STYLE*PTH.

In model A, young speakers advocate non-peaking forms more than old speakers, generally speaking. The proportion of young people using non-peaking forms is 78.0%, higher than the old people’s 67.3%. But in the T4A morphemes, the age differences are not obvious (81.6% vs 84.9%). This is because both old and young people have almost finished the variation for T4A morphemes, but old people are still conservative in T4B morphemes by using many more peaking forms than young speakers. Regarding REGION*STYLE, generally speaking the percentage of those using non-peaking forms is higher in the reading of paragraphs than in the reading of minimal pairs, as shown by the significance of the main effect STYLE

(76.4% vs 79.5%). But this difference is mainly present in the urban area (67.2% vs 71.2%) and is absent in the suburban area (86.8% vs 87.8%). This suggests that conditioning effects of STYLE are more obvious in urban areas. This finding will be tested in the other variables. Finally, regarding REGION*MORPHEME, free morphemes are more conservative than bound morphemes but their differences mainly exist in the suburban area (59.8% vs 16.1%); in the urban area the differences are less pronounced (19.6% vs 6.9%). This shows that the effect of MORPHEME is stronger in suburban areas, where the change of WXT4 is in an earlier stage than in urban areas.

Regarding the special constraints of model B, i.e. PTH and STYLE*PTH, the frequency difference of dipping and non-dipping forms is small between people who are not able to speak PTH (5.3% vs 7.8%) but is quite large among those who are able to speak PTH (49.2% vs 37.7%). The advanced form, the dipping variant, is used more by people who can speak both local dialect and PTH. The significant interaction of PTH*STYLE indicates that among the participants who are able to speak PTH, STYLE variation is found. More dipping variants are found than non-dipping variants in the reading of minimal pairs (30.6% vs 15.2%) and paragraphs (5.0% vs 2.5%). However, in the reading of wordlists, non-dipping variants are used more (19.9% vs 13.6%). This indicates that speakers use more conservative variants when reading wordlists.

5.2.3. Contour loss in WXT2/6

On the basis of observations from Figure 5.3, young people are reducing the slope of low rising tone for WXT2/6 in comparison to old people. This section uses the slope between P_{min} and P_{max} on the low rising contour, i.e. m_{rising}, as an index to trace the contour loss of T2/6. Before performing a likelihood ratio test, the choice between two highly correlated predictors - MC_TONE and PTH_TONE - had to be made. Section 5.1.3 proved that among old people the merger between T2 and T6 was completed earlier than the process of contour loss. So the distinction of T2 and T6, i.e., MC_TONE, is unlikely to play a role in the variation of contour loss. Thus PTH_TONE was investigated rather than MC_TONE. Table 5.29 lists the results of the likelihood ratio test on m_{rising} of WXT2/6:

Table 5.29 Results of likelihood ratio test on m_{rising} of WXT2/6 (n=2983)

	deviance	Chisq	Chi Df	Pr(>Chisq)
model-0	24726			
model-AGE	24705	20.766	1	0.000***
model-REGION	24725	0.754	1	0.385
model-SEX	24726	0.599	1	0.439
model-EDU	24726	0.293	1	0.589
model-PTH	24724	1.770	1	0.183
model-STYLE	24719	7.711	2	0.021*
model-MORPHEME	24726	0.048	1	0.826
model-MC_AB	24725	1.467	1	0.226
model-HEIGHT	24718	8.015	5	0.155
model-DUR	24725	1.116	1	0.291
model-PTH_TONE	24709	17.540	1	0.000***

Table 5.29 shows that AGE, STYLE and PTH_TONE are significant. The significant selected by a likelihood ratio test is AGE*PTH_TONE. It is then put into the final model with all the other main effects. The output is listed in Table 5.30.

Table 5.30 Output of the mixed effect linear regression on m_{rising} of WXT2/6 (n=2983)

predictor	*β*	SE	t	*p*
Intercept	15.750	1.619	9.730	0.000***
AGE (*young*) *old*	7.623	2.169	3.515	0.001**
STYLE *(wordlist)*				
paragraph	−1.559	0.620	−2.514	0.012*
minimal pair	0.600	1.185	0.506	0.614
PTH_TONE *(PT2) PT4*	−6.776	0.906	−7.480	0.000***
AGE*PTH_TONE				
*young*PT4*	6.600	1.089	6.063	0.000***

With respect to the low rising T2/6, Figure 5.3 illustrates that no obvious difference between urban and suburban areas can be traced from the mean F_0 contours of T2 and T6 but an obvious age difference exists in the shape of low rising contours. The results of mixed models analyses of m_{rising} are in line with these observations, which indicate that young people use much flatter rising contours than old people. Meanwhile, the contour loss of T2/6 in young people is associated with the variation of speaking style (word list vs paragraph). Young people tend to pronounce rising T2/6 with less steep slope in paragraph reading for saving articulatory efforts.

Regarding PTH_TONE and AGE*PTH_TONE, it is found that old people do not differ widely between the mean m_{rising} of PT2 (M=22.94) and PT4 (M=22.71) but young people do. The mean m_{rising} of young people is 15.44 for PT2 but only 8.85 for PT4. Young people not only use flatter tones than old people but also tend to use much flatter contours for those morphemes corresponding to PT4. Since MC_T4 morphemes are involved in another type of change, they are not examined for the contour process systematically. However, MC_T4 data is needed here to explain the differences of contour loss occurring in MC_T2/6. As discussed before, MC_T4A morphemes mainly have PT4 while MC_T4B morphemes mainly have PT3. To compare with the MC_T2/6, only the PT4 morphemes in MC_T4 are taken into account. The mean m_{rising} of those morphemes is 9.22 among young people, which is close to the value of those morphemes with PT4 morphemes in MC_T2/6. This suggests that the contour loss of PT2 morphemes among young people is significantly slower than PT4 morphemes. In other words, the contour loss of morphemes with PT2 keeps its low rising pattern in the dialect pronunciation much better than morphemes with the other PTH tones. One explanation is that its PTH tone, the mid rising tone /35/, helps the dialect tone to resist the loss of rising pattern.

5.2.4. Summary and discussion

Section 5.2 has investigated the constraints of four ongoing changes in Wúxī citation tone. The main findings are as follows:

(1) Morphemes pronounced as /51/ and /33/ are probably involved in the processes of tone borrowing. Suburban young speakers are more advanced by using more falling tones than urban young speakers, which shows the latter group’s linguistic insecurity towards local dialect.

(2) The trajectory of change of *WXT4* is “peaking tone > low rising tone > low dipping tone”. *WXT4* is constrained by a few social factors and linguistic factors. *WXT4* is a historical change, being constrained by internal causation factors like MC_AB. But in its late stage, when becoming a dipping tone, WXT4 is also influenced by the PT3, becoming a dipping tone with a late turning point. Interestingly, there are differences between urban and suburban people in regard to *WXT4*. Suburban young people are more conservative towards the variation of *WXT4* than urban young people, which is contrary to their behaviors in regard to *WXfalling*. This suggests that the comparison of urban and suburban young people may be helpful to define whether an ongoing change is internally or externally motivated.

(3) WXT2/6 presents the contour loss in the citation tone. WXT2/6 is an ongoing change mainly constrained by social factors, which has never been documented before. It can be deduced that the contour loss in WXT2/6 is caused by contact with PTH. Morphemes corresponding to PT2, a high rising tone, show resistance to this contour loss not because those morphemes tend to be more

conservative, but through the strong influence of PT2. This suggests that the influence of PTH on dialect can be reflected in two aspects. First, when a dialect tone has a similar tone shape to its corresponding PTH tone, the influence of PTH exerted on dialect is reflected in the preservation of the dialect tone shapes, like the moderate contour loss for MC_T2 morphemes. Second, when the tone shape of a dialect tone differs from the tone shape of its corresponding PTH tone, the PTH influence can cause the reduction of dialect tone like the severe contour loss for MC_T6 morphemes. In addition, the contour loss of T2/6 is also constrained by speaking style. The contour loss is more serious in the reading of paragraphs than in the reading of word lists. People tend to save articulatory efforts in reading paragraphs as it is less formal than word list reading. It suggests that another motivation for contour loss is to save articulatory efforts.

5.3. Identifying the linguistic variables of tone sandhi

5.3.1. Type I

The complexity of Wúxī tone sandhi patterns found in the previous studies has been discussed thoroughly in Section 2.4.4.2. A brief review is given here. Two types of sandhi patterns were found in the Wúxī dialect. Section 5.3.1 identifies the variation in Type I constructions while Section 5.3.2 identifies Type II. In Chan & Ren (1989), four patterns of Wúxī citation tone were distinguished. According to Chan & Ren's (1989) analysis, Type I has typically "left-dominant" sandhi patterns. It undergoes tone deletion and spread: tone deletion eliminates all but the tone of the leftmost syllable, and then this tonal melody is extended onto the entire sandhi domain. In other words, T2.X in Type I uses the citation pattern of T2; T4.X in Type I uses the citation pattern of T4; and T6.X in Type I uses the citation pattern of T6. Type I comprises four constructions: (1) reduplicated verbs; (2) verb followed by resultative or directional complements; (3) expressions with "number and classifier"; and (4) reduplicated nouns only used in child-directed speech. The first three types of words were tested in the reading of wordlists. 46 stimuli for T2.X, T4.X and T6.X are shown in Tables 3.9 to 3.11 in Chapter 3.

According to Qián (1992), Cáo (2003) and Xú (2007), T2.X uses /22.55/, /22.44/ or /33.55/, i.e., a low level tone plus a high level tone if neglecting the transcribers' differences; T4.X uses /24.31/, /13.21/ or /35.31/, i.e., a low rising tone plus a low falling tone; and T6.X uses /21.23/, /22.24/ or /33.34/, i.e., a low level tone plus a low rising tone. Considering the phonetic variation, T2.X is mostly transcribed into /22.44/ in the data of the present study. T4.X is mostly transcribed into /24.31/ while T6.X is mostly transcribed into /22.24/. The slashes "//" indicate that the transcriptions are phonemic, not considering the phonetic variations within each category. The transcription results of our own data on Type I constructions are highly in line with the results found in previous research. However, 68 tokens differ from the majority in the total 1840 valid observations, with 17 tokens using /24.31/ for T2.X, 11 tokens using /22.24/ for T4.X, three tokens using /24.31/ for T6.X and 37 using /22.44/ for T6.X.

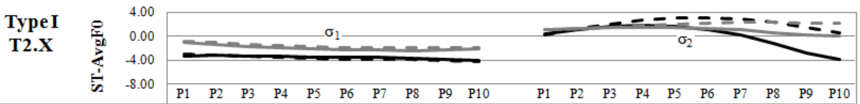
First, T2.X, T4.X and T6.X were investigated separately by mixed models to see if these exceptional tokens are involved in any ongoing changes. For each of them, model-0 was built with only the random factors. Model-AGE, model-REGION, model-AGE*REGION were compared with model-0 one by one to see if any fixed factor or interaction indicating ongoing changes is significant. The results are presented in Table 5.31:

Table 5.31 Results of likelihood ratio test on Type I inconsistencies in T2.X, T4.X, and T6.X

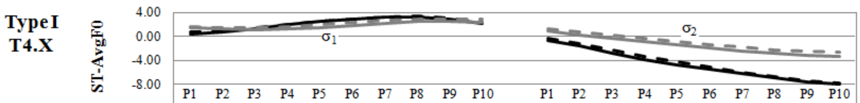
	deviance	Chisq	Chi Df	Pr(>Chisq)
T2.X (n=377)				
model-0	-1375.9			
model-AGE	-1376.2	0.284	1	0.594
model-REGION	-1375.9	0.032	1	0.859
model-AGE*REGION	-1377.4	1.475	3	0.688
T4.X (n=689)				
model-0	-181.4			
model-AGE	-183.5	2.069	1	0.150
model-REGION	-182.7	1.239	1	0.266
model-AGE*REGION	-186.1	4.725	3	0.193
T6.X (n=678)				
model-0	-415.3			
model-AGE	-415.9	0.562	1	0.454
model-REGION	-418.7	3.372	1	0.124
model-AGE*REGION	-421.7	6.352	3	0.096

The results in Table 5.31 show that AGE, REGION and their interaction cannot yield a better model. It suggests that these exceptional tokens in Type I are not involved in ongoing changes and are less interesting in view of tonal change.

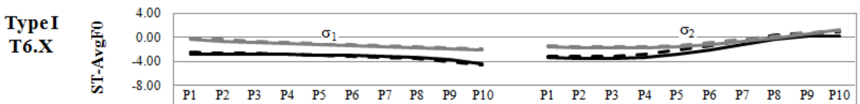
To investigate the "regulars", the mean F_0 contours are presented separately for T2.X, T4.X and T6.X in Figure 5.6. Age differences for F_0 mean contours show up everywhere, except for the σ_1 of T4.X. The difference between urban and suburban areas mainly shows up in the σ_2 of T2.X.



Type I T2.X	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
urban old n=98	-2.94	-3.13	-3.29	-3.46	-3.64	-3.79	-3.85	-3.89	-4.07	-4.17	0.27	1.28	2.06	2.65	3.03	3.13	2.92	2.38	1.48	0.64
suburban old n=98	-3.29	-3.25	-3.28	-3.37	-3.47	-3.55	-3.61	-3.69	-3.85	-4.05	0.44	1.08	1.60	1.80	1.64	1.11	0.16	-1.16	-2.73	-3.96
urban young n=84	-0.78	-1.10	-1.38	-1.63	-1.83	-1.95	-2.00	-1.99	-1.96	-1.85	1.10	1.30	1.53	1.77	2.01	2.21	2.34	2.39	2.25	2.15
suburban young n=97	-1.09	-1.42	-1.68	-1.90	-2.10	-2.26	-2.36	-2.38	-2.25	-2.02	1.10	1.27	1.45	1.53	1.50	1.34	1.05	0.61	0.14	-0.02



Type I T4.X	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
urban old n=171	0.67	0.96	1.42	1.95	2.47	2.92	3.26	3.40	3.08	2.41	-0.27	-1.17	-2.24	-3.30	-4.28	-5.16	-5.98	-6.74	-7.47	-7.80
suburban old n=171	0.46	0.80	1.34	1.93	2.49	2.93	3.22	3.28	2.87	2.18	-0.67	-1.64	-2.73	-3.80	-4.71	-5.51	-6.22	-6.86	-7.58	-7.92
urban young n=174	1.59	1.49	1.51	1.68	1.97	2.33	2.69	2.92	2.94	2.80	1.27	0.79	0.28	-0.26	-0.78	-1.31	-1.79	-2.22	-2.54	-2.62
suburban young n=173	1.43	1.24	1.19	1.28	1.50	1.83	2.17	2.45	2.55	2.43	0.84	0.27	-0.27	-0.87	-1.46	-1.99	-2.45	-2.87	-3.20	-3.29



Type I T6.X	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
urban old n=177	-2.53	-2.58	-2.67	-2.81	-2.96	-3.12	-3.29	-3.52	-4.00	-4.55	-3.19	-3.19	-3.09	-2.73	-2.14	-1.33	-0.41	0.36	0.82	0.94
suburban old n=165	-2.77	-2.79	-2.80	-2.86	-2.92	-3.00	-3.11	-3.30	-3.72	-4.38	-3.41	-3.53	-3.53	-3.32	-2.82	-2.03	-1.13	-0.28	0.24	0.24
urban young n=167	-0.10	-0.37	-0.61	-0.85	-1.07	-1.26	-1.43	-1.58	-1.80	-1.99	-1.40	-1.60	-1.65	-1.60	-1.36	-0.94	-0.38	0.22	0.75	1.18
suburban young n=169	-0.31	-0.61	-0.82	-1.01	-1.20	-1.38	-1.55	-1.71	-1.90	-2.05	-1.53	-1.71	-1.78	-1.72	-1.53	-1.19	-0.71	-0.09	0.61	1.22

Figure 5.6 Mean F_0 contours of T2.X, T4.X and T6.X, split by AGE and REGION

Regarding T2.X, the low level contours of its σ_1 are around 2 ST higher among young people than old people. Meanwhile, there are differences between old and young people and between urban and suburban people in σ_2. First, urban young speakers have the flattest contours while a falling trend can be distinguished in the second half of contours of the urban old speakers. This falling trend is more obvious in the suburban area, where the productions of old speakers are visibly falling. Regarding T4.X, old people have deep falling tone in σ_2 while young people have a 4 ST higher offset. Regarding T6.X, young speakers have higher pitch than old speakers in σ_1, which is similar to T2.X. Young speakers raised the onset and then make the low rising contour flatter than the old speakers in σ_2.

In sum, there are two ongoing changes in Type I words. On the whole, young people have much smaller pitch differences between σ_1 and σ_2 than old people. It seems to be similar with the contour loss process distinguished in the citation tone. Meanwhile, everyone except suburban old people have higher offset in the σ_2 of Type I T2.X, losing the falling trend.

5.3.2. Type II

Besides four constructions of Type I, common compounds and phrases use Type II patterns. According to Chan & Ren (1989), Type II sandhi undergoes an extra step of pattern substitution before tone deletion and spread. T2.X has /22.44/ in Type I, which is substituted for /24.31/; T6.X has /22.24/ in Type I, which is substituted for /22.44/. T4.X is a bit more complicated. Its /24.31/ of Type I is replaced by /22.24/ and /22.44/ in Type II. The two-way split pattern for T4.X is conditioned by the MC tonal categories on σ_2. If the σ_2 of the tone sandhi domain is *Shǎng* and *Qù*, i.e., T3, T4, T5 and T6, the substitution is /22.24/; if σ_2 is *Píng* and *Rù*, i.e., T1, T2, T7 or T8, the substitution is /22.44/. The stimuli for Type II compounds and phrases are listed in Tables 3.6 to 3.8 in Chapter 3. The present dataset is able to test the validity of those substitutive rules used in T4.X. Results of T2.X, T4.X and T6.X are reported in Sections 5.3.2.1 to 5.3.2.3 separately. Table 2.11 (Chapter 2) summarizes the tone sandhi patterns of T2.X, T4.X and T6.X. Cáo (2003) and Xú (2007) found T2.X, T4.X and T6.X use at least two different sandhi patterns: /22.24/, /22.44/ and /24.31/.

From the transcription of all the observations, the sandhi pattern of /55.31/ is found occasionally in T2.X, T4.X and T6.X alongside /22.24/, /22.44/ and /24.31/. There are 46 observations of /55.31/ out of the total 5530: five for T2.X, 10 for T4.X and 31 for T6.X. /55.31/ was recorded in Qián (1992) but not in the other studies (see Table 2.11). So in Section 5.3.1, /55.31/ is first examined to see if it is involved in any ongoing changes. Likewise, model-0 was built with only random factors WORD (137 levels) and SPEAKER (40 levels) on the variable /55.31/ (yes vs no). Model-AGE and model-REGION were built and compared with model-0 successively.

Results show that AGE is significant ($\chi^2(1)=30.913$, $p=0.000^{***}$), while REGION is not ($\chi^2(1)=0.047$, $p=0.828$). Young speakers are more likely to use /55.31/ than old speakers. A mixed model containing AGE and the interaction of AGE and REGION as fixed factors was then built and compared with model-AGE. The chi square value is not significant ($p=0.772$). Therefore, it is an ongoing change constrained by AGE but not REGION.

Next, the linguistic constraints of /55.31/ are investigated. Since the new variant /55.31/ occurred rather infrequently in percentage terms, the common feature of the /55.31/ tokens was investigated using qualitative analysis. /55.31/

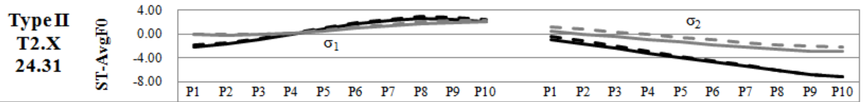
is the sandhi pattern used frequently by T3.X and T5.X for Type II constructions and T1.X for Type I constructions (Cáo 2003; Xú 2007). In other words, /55.31/ is a pattern used by the phrases of the Yīn register whose nucleus of σ_1 is in the modal voice. In our dataset, all the T2.X, T4.X and T6.X phrases using /55.31/ patterns lost their breathy voice and are in modal voice for their first syllables. For example, 议论 'discuss' use [i] instead of [ɦi̤]; 愿望 'wishes' use [nʊ] instead of [nʊ̤]; 寺庙 'temples' use [sɨ] instead of [sɦɨ̤] and 负数 'negative numbers' uses [fu] instead of [fɦṳ] in their /55.31/ realizations. It seems that the usage of the /55.31/ pattern in the Yáng register is correlated with the loss of breathy voice. When a phrase in the Yáng register replaces its breathy with model voice, it is more likely to be realized in a sandhi pattern originally used in the Yīn register. Meanwhile, young people, the group of people using /55.31/ pattern in T2.X, T4.X and T6.X the most, are also the group which lost their breathy voice in pronouncing words of the Yáng register. So it is hypothesized that dropping the breathy feature of the nucleus is the reason for those phrases of the Yáng register to switch to /55.31/, a sandhi pattern used in the Yīn register. However, this cannot be tested acoustically in the present study because no valid acoustic index for the phonation contrast in the Wú dialects is found in Section 4.4 (Chapter 4). Further research should be undertaken to explore this issue.

5.3.2.1. WXT2.X

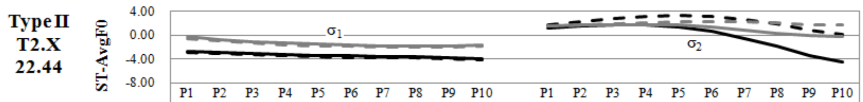
2017 valid T2.X tokens are observed in the current dataset. Besides five cases of /55.31/ discussed above, 80.3% (n=1616) of the remaining tokens use /24.31/; 18.2% (n=367) use /22.44/ while 1.4% (n=29) use /22.24/, according to my transcription. The majority of tokens of T2.X use /24.31/, which is in line with Chan & Ren (1989), Cáo (2003) and Xú (2007). However, /22.44/ and /22.24/ are also found. Given the percentage distribution, the major variation of T2.X is the variation between /24.31/ and /22.44/. The contrast between /24.31/ and /22.44/ for T2.X phrases was mainly lexically conditioned. Among the 51 total phrases, 43 of them were realized as /24.31/. In the total 1700 valid observations of those 43 phrases, 91.4% of them use /24.31/. Those 43 phrases are grouped together because each of them has at least 30 valid observations of /24.31/ by 40 speakers. The remaining eight phrases have 312 valid observations, 75.3% of which use /22.44/. Each of those eight phrases has at least 20 valid observations of /22.44/ by 40 speakers. Table 5.32 lists the lexical split of /24.31/ and /22.44/ for Type II T2.X:

Table 5.32 The lexical split of /24.31/ and /22.44/ for Type II T2.X

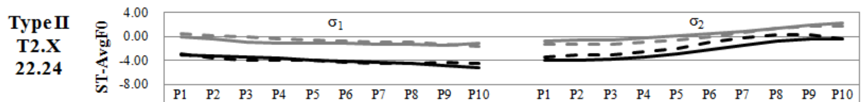
	/24.31/	/22.44/
T2.T1	曹操 ‘Cao Cao, a famous historical figure in China’, 调羹 ‘spoon’, 停当 ‘be completed’, 辰光 ‘time’, 台灯 ‘desk lamp’, 迷宫 ‘maze’,提高 ‘to raise or improve’, 洋机 ‘sewing machine’, 阳台 ‘balcony’, 莲心 ‘lotus seed’	层单 ‘sheet’, 馋胚 ‘glutton’
T2.T2	排球 ‘volleyball’, 名堂 ‘variety’, 梧桐 ‘Chinese parasol tree’, 胡咙 ‘throat’, 洋盘 ‘one not smart enough’	浑堂 ‘bathhouse’
T2.T3	团长 ‘team leader’, 篮板 ‘backboard’, 难板 ‘seldom’, 门坎 ‘door sill’, 元宝 ‘(gold or silver) ingot’	词组 ‘phrase’
T2.T4	渠道 ‘channel’	厘米 ‘centimeter’
T2.T5	咸菜 ‘brined vegetables’, 牢靠 ‘secure’, 文旦 ‘pomelo’	残废 ‘disabled’
T2.T6	余地 ‘leeway’, 闲话 ‘gossip’, 原旧 ‘as usual’	迷露 ‘fog’, 和调 ‘to play a joke’
T2.T7	湖北 ‘a province in China’, 难得 ‘rare’, 毛笔 ‘writing brush’	
T2.T8	题目 ‘topic’, 埋没 ‘to bury’, 寒热 ‘fever’	



	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
urban old n=382	-1.76	-1.39	-0.74	0.09	1.02	1.88	2.55	2.93	2.89	2.53	-0.28	-1.07	-1.94	-2.82	-3.69	-4.49	-5.25	-6.01	-6.76	-7.15
suburban old n=382	-2.16	-1.68	-0.90	0.01	0.94	1.75	2.35	2.64	2.52	2.06	-0.87	-1.56	-2.36	-3.17	-3.94	-4.68	-5.39	-6.09	-6.76	-7.13
urban young n=421	0.04	-0.09	-0.02	0.24	0.64	1.15	1.67	2.09	2.32	2.34	1.26	0.86	0.41	-0.06	-0.52	-0.96	-1.38	-1.73	-1.98	-2.07
suburban young n=431	-0.01	-0.15	-0.07	0.16	0.53	0.99	1.46	1.83	2.02	2.04	0.56	0.07	-0.40	-0.87	-1.33	-1.75	-2.14	-2.50	-2.79	-2.88



	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
urban old n=120	-2.79	-3.00	-3.24	-3.45	-3.60	-3.68	-3.73	-3.83	-3.99	-4.06	1.73	2.27	2.76	3.14	3.32	3.13	2.67	1.90	0.92	0.14
suburban old n=113	-2.71	-2.91	-3.06	-3.20	-3.31	-3.41	-3.50	-3.58	-3.78	-3.95	1.23	1.61	1.83	1.81	1.43	0.66	-0.48	-1.83	-3.32	-4.38
urban young n=70	-0.47	-0.85	-1.19	-1.53	-1.73	-1.84	-1.89	-1.92	-1.90	-1.82	1.59	1.79	1.97	2.13	2.28	2.24	2.24	2.08	1.79	1.72
suburban young n=64	-0.26	-0.67	-1.03	-1.31	-1.51	-1.64	-1.73	-1.78	-1.82	-1.66	1.64	1.71	1.81	1.85	1.77	1.45	0.93	0.36	-0.05	-0.15



	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
urban old n=5	-2.86	-3.56	-3.91	-3.98	-3.99	-4.20	-4.38	-4.42	-4.27	-4.45	-3.42	-3.10	-2.95	-2.59	-1.93	-0.92	-0.14	0.42	0.30	-0.34
suburban old n=10	-3.01	-3.28	-3.46	-3.65	-3.88	-4.09	-4.32	-4.52	-4.84	-5.11	-3.98	-3.94	-3.71	-3.36	-2.83	-2.07	-1.37	-0.77	-0.38	-0.30
urban young n=6	0.43	0.20	-0.05	-0.31	-0.55	-0.72	-0.85	-0.97	-1.23	-1.62	-1.32	-1.34	-1.21	-0.94	-0.60	-0.02	0.73	1.38	1.77	1.83
suburban young n=8	-0.03	-0.39	-0.82	-1.04	-1.08	-1.14	-1.21	-1.32	-1.46	-1.07	-0.75	-0.62	-0.47	-0.20	0.21	0.57	0.88	1.35	1.93	2.34

Figure 5.7 Mean F_0 contours of Type II T2.X, split by sandhi pattern, AGE and REGION

The mean F_0 contours are presented for three sandhi patterns found in Type II. Each one is split by AGE and REGION. As can be seen from Figure 5.7, F_0 mean contours have age differences mainly in the σ_2 of /24.31/, the σ_1 and σ_2 of /22.44/ and the σ_1 and σ_2 of /22.24/. The difference between urban and suburban areas mainly exists in the σ_2 of /22.44/; urban speakers have a higher offset. These trends are very similar to the ongoing changes found in Figure 5.6: (1) young people have much smaller pitch differences between σ_1 and σ_2 than old people, and (2) young people and urban people end with a higher offset in the σ_2 of /22.44/.

Although the lexical split of /24.31/ and /22.44/ for T4.X is rather regular, this split is now further investigated using other social or linguistic factors. Once again, mixed models were built one by one to compare with random-effects-only model. The results are listed in Table 5.33.

Table 5.33 Results of likelihood ratio test on the split of /24.31/ (*n.*=1616) and /22.44/ (n=367) for Type II T2.X

	deviance	Chisq	Chi Df	Pr(>Chisq)
model-0	820.13			
model-AGE	785.20	34.930	1	0.000***
model-REGION	819.62	0.514	1	0.474
model-SEX	820.00	0.134	1	0.715
model-EDU	818.63	1.503	1	0.220
model-PTH	818.09	2.042	1	0.153
model-STYLE	816.44	3.690	1	0.055
model-σ_2_MC_TONE	814.49	5.642	7	0.582
model-σ_1_MC_AB	820.16	0.000	1	1.000
model- LEXICAL ATTESTEDNESS	818.27	1.859	1	0.173
model-σ_1_HEIGHT	809.79	10.345	5	0.066
model-σ_2_HEIGHT	811.42	8.705	6	0.191
model-σ_2_PTH_TONE	819.99	0.138	3	0.987
model-σ_1_DUR	819.81	0.316	1	0.574
model-σ_2_DUR	741.02	79.111	1	0.000***
model-σ_2_CODA	819.75	0.379	1	0.538

Table 5.33 shows that only AGE and σ_2_DUR are significant for the variation of /24.31/ and /22.44/. Next, interactions of AGE and other main effects are investigated using a likelihood ratio test. It is found that the interaction of AGE and LEXICAL ATTESTEDNESS is significant. So AGE, σ_2_DUR and AGE:LEXICAL ATTESTEDNESS were entered into the final model together. The output is listed in Table 5.34:

Table 5.34 Output of logistic model with mixed-effects on the split of /24.31/ (n=1616) and /22.44/ (n=367) for Type II T2.X

predictor	*log-odds*	SE	t	*p*
Intercept	0.001	1.714	−4.243	0.000***
AGE (*old*): *young*	0.917	0.450	−0.191	0.848
σ_2_DUR	1.020	0.003	7.229	0.000***
AGE:LEXICAL ATTESTEDNESS				
old PTH & Wú lexical item*	0.128	1.979	−1.037	0.300
*young*PTH & Wú lexical item*	0.020	1.990	−1.968	0.049*

C=0.983, Somer's D_{xy}=0.966

Table 5.34 shows that σ_2_DUR is significant. The value of log-odds suggests that the longer duration σ_2 has, the more likely to use /22.44/. AGE is not a significant main effect but its interaction with LEXICAL ATTESTEDNESS is. The age difference for Wú lexical items is not significant (18.6% vs 19.8%) but significant for PTH & Wú lexical items (58.1% vs 66.6%). Young people use more /24.31/ in the shared lexical items of PTH and Wú than old people, which is an ongoing change. However, the age difference for /22.44/ mainly exists in the eight phrases predominantly using /22.44/ (Table 5.35). It means that this ongoing change cannot explain the substantial split of /24.31/ and /22.44/ but indicates another trend of changes. Young people are adapting /24.31/ for /22.44/ especially for those /22.44/ phrases. As this trend is more obvious in shared lexical items of PTH and Wú, the explanation is that the PTH tone of σ_1 in T2.X is a rising tone /35/. Young people, who are more influenced by PTH, tend to use a rising tone in the σ_1 of PTH & Wú lexical items and then adopt /24.31/ for the whole phrase. This ongoing change shows the influence of PTH tone on the variation of sandhi patterns in the dialect.

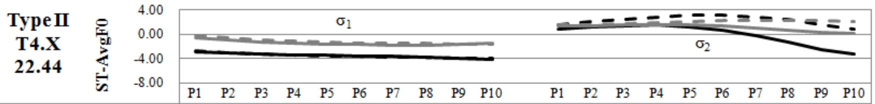
5.3.2.2. WXT4.X

There are 1739 valid T4.X observations in the current dataset. Besides the 10 cases of /55.31/ discussed above, 56.9% (n=990) of the remaining tokens use /22.44/, while 43.1% (n=749) use /22.24/, according to my transcription. The variation between /22.44/ and /22.24/ for T4.X is in line with the findings in Chan & Ren (1989), Cáo (2003) and Xú (2007). An exploratory analysis shows that it is lexically conditioned. Among the 44 phrases, 24 of them were realized almost exclusively as /22.44/; 94.2% of their 988 valid observations (n=937) use /22.44/. Each of the 24 phrases has at least 30 valid observations of /22.44/ among 40 speakers. The other 18 words are /22.24/ dominant: 93.2% of their 749 valid observations use /22.24/. Table 5.35 lists the lexical split of /22.24/ and

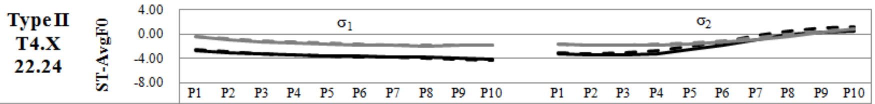
/22.44/ for Type II T4.X. Chan & Ren (1989) claimed that this two-way split (/22.44/ vs /22.24/) is conditioned by σ_2, which indicates the residue of an unfinished historical shift from last to first syllable dominance. If the σ_2 of the tone sandhi domain is *Shǎng* and *Qù*, i.e., T3, T4, T5 and T6, the substitution is /22.24/; if the σ_2 is *Píng* and *Rù*, i.e., T1, T2, T7 or T8, the substitution is /22.44/. However, it is obvious that exceptions to this substitution rule are found in our data, which are indicated by grey shading in Table 5.35. The substitution rule of Chan & Ren (1989) is further investigated to test its validity in our data. A new predictor σ_2_MC_TONE (collapsed) was added to build the mixed model. It collapses MC_T1, T2, T7 and T8 into one condition and collapses MC_T3 to MC_T6 into another.

Table 5.35 The lexical split of /22.44/ and /22.24/ for Type II T4.X

	/22.44/	/22.24/
T4.T1	罢工 'strike', 像腔 'to behave properly', 午休 'noon break',上腔 'to retort sarcastically', 买方 'buyer', 雨披 'poncho', 懒胚 'lazy bone', 眼泡 'eyelid'	老官 'husband'
T4.T2	冷盆 'cold dishes', 领头 'collar'	理由 'reason', 羽毛 'feather'
T4.T3	户口 'registered residence'	老虎 'tiger', 老酒 'liquor', 耳朵 'ear'
T4.T4	腐乳 'preserved beancurd', 丈姆 'mother-in-law'	上下 'up and down', 蚂蚁 'ant'
T4.T5	负数 'negative number', 鳝片 'eel slice'	上进 'enterprising', 礼拜 'week', 上照 'photogenic', 上劲 'enthusiastic about doing something', 惹气 'angry', 造句 'to make sentences', 有劲 'interesting'
T4.T6	近视 'myopic', 部队 'army'	上路 'reliable', 懒惰 'lazy', 道地 'considerable'
T4.T7	道德 'morality', 挺刮 '(clothes) neat', 米色 'cream color'	
T4.T8	聚集 'to gather', 序幕 'prologue', 满月 'the achievement of the first month of life', 暖热 'warm'	



	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
urban old n=250	-2.73	-3.04	-3.27	-3.44	-3.56	-3.65	-3.70	-3.75	-3.85	-3.95	1.65	2.11	2.56	2.92	3.13	3.14	2.91	2.42	1.64	0.85
suburban old n=252	-2.80	-3.01	-3.17	-3.31	-3.46	-3.57	-3.66	-3.75	-3.90	-4.05	0.94	1.24	1.46	1.50	1.28	0.73	-0.13	-1.25	-2.43	-3.18
urban young n=247	-0.25	-0.61	-0.88	-1.10	-1.26	-1.38	-1.46	-1.52	-1.53	-1.55	1.51	1.60	1.76	1.95	2.13	2.27	2.37	2.37	2.24	2.11
suburban young n=239	-0.50	-0.89	-1.18	-1.40	-1.57	-1.69	-1.74	-1.71	-1.59	-1.42	1.40	1.47	1.56	1.60	1.53	1.38	1.11	0.76	0.39	0.23



	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
urban old n=187	-2.54	-2.93	-3.26	-3.44	-3.56	-3.68	-3.79	-3.90	-4.10	-4.31	-3.09	-3.17	-3.04	-2.67	-2.05	-1.18	-0.27	0.53	1.04	1.16
suburban old n=186	-2.63	-3.03	-3.27	-3.40	-3.51	-3.60	-3.68	-3.75	-3.91	-4.12	-3.13	-3.37	-3.40	-3.13	-2.56	-1.76	-0.85	-0.07	0.39	0.43
urban young n=184	-0.36	-0.72	-1.00	-1.24	-1.45	-1.61	-1.71	-1.75	-1.79	-1.86	-1.65	-1.77	-1.76	-1.66	-1.41	-1.02	-0.51	0.04	0.55	0.95
suburban young n=192	-0.44	-0.86	-1.19	-1.41	-1.58	-1.71	-1.83	-1.89	-1.87	-1.83	-1.57	-1.75	-1.82	-1.77	-1.61	-1.33	-0.87	-0.31	0.33	0.81

Figure 5.8 Mean F_0 contours of Type II T4.X, split by sandhi pattern, AGE and REGION

The mean F_0 contours are presented for the two sandhi patterns found in Type II T4.X. Each one is split by AGE and REGION. The raw frequencies are indicated in the legend. As can be seen from Figure 5.8. F_0 mean contours have age differences mainly in the σ_1 and σ_2 of both patterns. The difference between urban and suburban areas mainly exists in the σ_2 of /22.44/, the same as was identified in Type I T2.X, and the /22.44/ of Type II T2.X.

Although the lexical split of /22.44/ and /22.24/ for T4.X is regular, this split is further investigated by studying the impact of other social or linguistic factors. Mixed models were built one by one to compare with the random-effects-only model. The results are listed in Table 5.36.

Table 5.36 shows, besides σ_2_DUR, four significant predictors - σ_2_MC_TONE, σ_2_PTH_TONE, σ_2_CODA and σ_2_MC_TONE (collapsed). These four predictors are intrinsically correlated. The relationship between MC_TONE and PTH_TONE can be found in Table 3.15 (Chapter 3). σ_2_CODA collapses MC_T1-6 into one condition and MC_T7 and MC_T8 into another. To avoid the problem of collinearity, these four predictors cannot be entered into the final model jointly. As Table 5.36 shows, σ_2_MC_TONE (collapsed) yields a chi square difference of 25.983 with one degree of freedom, which has the smallest *p* value. The model built on σ_2_MC_TONE, σ_2_PTH_TONE and σ_2_CODA shows similar conditioning patterns but with higher deviation. Therefore, σ_2_MC_TONE (collapsed) was entered into the final model instead of them. Its interaction with other independent main effects was then investigated, but none was significant. So σ_2_MC_TONE (collapsed) and $σ_2$_DUR were entered into the final model. The output of the final model is listed in Table 5.37.

Table 5.36 Results of likelihood ratio test on the split of /22.44/ (n=990) and /22.24/ (n=749) for Type II T4.X

	deviance	Chisq	Chi Df	Pr(>Chisq)
model-0	798.78			
model-AGE	798.33	0.453	1	0.501
model-REGION	798.71	0.067	1	0.796
model-SEX	798.07	0.712	1	0.399
model-EDU	797.96	0.817	1	0.366
model-PTH	798.51	0.274	1	0.601
model-STYLE	797.24	1.536	1	0.215
model-$σ_2$_MC_TONE	763.25	35.535	7	0.000^{***}
model-$σ_1$_MC_AB	798.78	0.004	1	0.952
model-LEXICAL ATTESTEDNESS	795.53	3.247	1	0.072
model-$σ_1$_HEIGHT	794.90	3.879	4	0.423
model-$σ_2$_HEIGHT	788.02	10.765	6	0.096
model-$σ_2$_PTH_TONE	779.74	19.043	3	0.000^{***}
model-$σ_1$_DUR	797.88	0.901	1	0.342
model-$σ_2$_DUR	760.19	38.590	1	0.000^{***}
model-$σ_2$_CODA	791.28	7.499	1	0.006^{**}
model-$σ_2$_MC_TONE (collapsed)	772.80	25.983	1	0.000^{***}

Table 5.37 Output of logistic model with mixed-effects on the split of /22.44/ (n=990) and /22.24/ (*n*=749) for Type II T4.X

predictor	*log-odds*	SE	t	*p*
Intercept	0.001	0.925	−7.434	0.000^{***}
$σ_2$_MC_TONE (collapsed) *(T1, T2, T7, T8)*: *T3-T6*	341.009	1.057	5.516	0.000^{***}
$σ_2$_DUR	1.014	0.002	6.193	0.000^{***}

C=0.984, Somer's D_{xy}=0.967

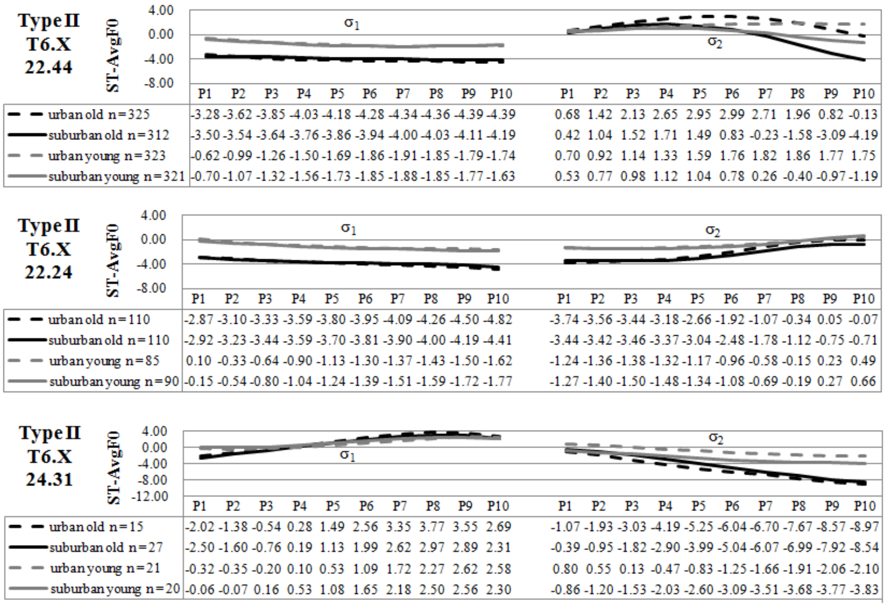
The substitution rule proposed by Chan & Ren (1989, see the beginning of Section 5.3.2) is confirmed in the data. If the $σ_2$ of a phrase belongs to MC_T3, MC_T4, MC_T5 or MC_T6, the phrase is more likely to use /22.24/. If $σ_2$ has longer duration, the phrase is more likely to be /22.24/ than to be /22.44/, which is in agreement with the finding that rising tones intrinsically have a longer duration than level tones (Howie, 1976; Rose, 1980; Zhu, 1999). This output also suggests that the split of /22.44/ and /22.24/ is not involved in any ongoing changes. Only two ongoing changes have been detected in T4.X, i.e., a higher offset in the $σ_2$ of /22.44/ and the contour loss of /22.44/ and /22.24/. They will be discussed in Section 5.4 below.

5.3.2.3. *WXT6.X*

There are 1763 valid T6.X observations in the current dataset. Besides 31 cases of /55.31/ discussed above, 72.9% (n=1286) of the remaining tokens use /22.44/, 22.4% (n=394) use /22.24/ and 4.7% (n=83) use /24.31/, according to my transcription. The majority tokens of T6.X are /22.44/, which is in line with the findings of Chan & Ren (1989), Cáo (2003) and Xú (2007). However, /22.24/ and /24.31/ are also found. Given the percentage distribution, the major variation of T6.X is the variation between /22.44/ and /22.24/, and it mainly shows up in the lexical words. As shown in the right column of Table 5.38, 10 of the total 45 phrases were realized mainly as /22.24/. Each of those 10 phrases have at least 30 valid observations of /22.24/ among 40 speakers; on average, 92.8% (*n*=387) use /22.24/. Another 33 phrases have 1267 valid observations; 99.0% of them use /22.44/. Table 5.38 lists the lexical split of /22.44/ and /22.24/ for Type II T6.X. Besides /22.44/ and /22.24/, /24.31/ is also found in Type II T6.X. This pattern is exclusively used by the 40 Wúxī speakers: 浪费 'waste' and 料重 '(burden) heavy'.

Table 5.38 The lexical split of /22.44/ and /22.24/ for Type II T6.X

	/22.44/	/22.24/
T6.T1	召开 'to convoke', 自家 'oneself', 便当 'convenient', 闹钟 'alarm clock'	
T6.T2	蛋黄 'yolk', 寿头 'a fool', 旧年 'last year', 烂糊 'pulpy', 弄堂 'alley', 用场 'function'	调查 'to investigate'
T6.T3	字典 'dictionary', 事体 'affair; matter', 面孔 'face'	具体 'specific', 效果 'effect'
T6.T4	护士 'nurse', 运道 'fortune', 号码 'number', 闹猛 'bustling with noise and excitement'	导演 'directors of film, drama, etc.', 具有 'have, process', 万岁 'long live'
T6.T5	大蒜 'garlic', 硬劲 'manage to do sth with difficulty', 卖相 'appearance'	夜快 'at night fall'
T6.T6	寺庙 'temple', 烂饭 'rice cooked with too much water', 夜饭 'supper'	预料 'to predict', 议论 'to discuss', 自愿 'voluntarily', 愿望 'desire, wish'
T6.T7	硬扎 'firm, strong', 败笔 'a faulty expression in writing', 预测 'to forecast'	
T6.T8	附属 'affiliated', 赖学 'to play truant', 卖力 'to work hard'	



Type II T6.X 22.44

	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
urban old n=325	-3.28	-3.62	-3.85	-4.03	-4.18	-4.28	-4.34	-4.36	-4.39	-4.39	0.68	1.42	2.13	2.65	2.95	2.99	2.71	1.96	0.82	-0.13
suburban old n=312	-3.50	-3.54	-3.64	-3.76	-3.86	-3.94	-4.00	-4.03	-4.11	-4.19	0.42	1.04	1.52	1.71	1.49	0.83	-0.23	-1.58	-3.09	-4.19
urban young n=323	-0.62	-0.99	-1.26	-1.50	-1.69	-1.86	-1.91	-1.85	-1.79	-1.74	0.70	0.92	1.14	1.33	1.59	1.76	1.82	1.86	1.77	1.75
suburban young n=321	-0.70	-1.07	-1.32	-1.56	-1.73	-1.85	-1.88	-1.85	-1.77	-1.63	0.53	0.77	0.98	1.12	1.04	0.78	0.26	-0.40	-0.97	-1.19

Type II T6.X 22.24

	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
urban old n=110	-2.87	-3.10	-3.33	-3.59	-3.80	-3.95	-4.09	-4.26	-4.50	-4.82	-3.74	-3.56	-3.44	-3.18	-2.66	-1.92	-1.07	-0.34	0.05	-0.07
suburban old n=110	-2.92	-3.23	-3.44	-3.59	-3.70	-3.81	-3.90	-4.00	-4.19	-4.41	-3.44	-3.42	-3.46	-3.37	-3.04	-2.48	-1.78	-1.12	-0.75	-0.71
urban young n=85	0.10	-0.33	-0.64	-0.90	-1.13	-1.30	-1.37	-1.43	-1.50	-1.62	-1.24	-1.36	-1.38	-1.32	-1.17	-0.96	-0.58	-0.15	0.23	0.49
suburban young n=90	-0.15	-0.54	-0.80	-1.04	-1.24	-1.39	-1.51	-1.59	-1.72	-1.77	-1.27	-1.40	-1.50	-1.48	-1.34	-1.08	-0.69	-0.19	0.27	0.66

Type II T6.X 24.31

	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
urban old n=15	-2.02	-1.38	-0.54	0.28	1.49	2.56	3.35	3.77	3.55	2.69	-1.07	-1.93	-3.03	-4.19	-5.25	-6.04	-6.70	-7.67	-8.57	-8.97
suburban old n=27	-2.50	-1.60	-0.76	0.19	1.13	1.99	2.62	2.97	2.89	2.31	-0.39	-0.95	-1.82	-2.90	-3.99	-5.04	-6.07	-6.99	-7.92	-8.54
urban young n=21	-0.32	-0.35	-0.20	0.10	0.53	1.09	1.72	2.27	2.62	2.58	0.80	0.55	0.13	-0.47	-0.83	-1.25	-1.66	-1.91	-2.06	-2.10
suburban young n=20	-0.06	-0.07	0.16	0.53	1.08	1.65	2.18	2.50	2.56	2.30	-0.86	-1.20	-1.53	-2.03	-2.60	-3.09	-3.51	-3.68	-3.77	-3.83

Figure 5.9 Mean F_0 contours of Type II T6.X, split by sandhi pattern, AGE and REGION

The mean F_0 contours are presented for three sandhi patterns found in Type II T6.X. Each one is split by AGE and REGION. The raw frequencies are indicated in the legend. As can be seen from Figure 5.9, age differences for F_0 mean contours are mainly found in the σ_1 and σ_2 of /22.24/ and /22.44/, as well as in the σ_2 of /24.31/. The difference between urban and suburban areas mainly exists in the σ_2 of /22.44/. These findings are exactly the same as the ongoing changes found in Type I T2.X, T4.X and T6.X and Type II T2.X and T4.X. According to the observation, the loss of falling pattern of the σ_2 of /22.44/ and contour loss of /22.24/, /22.44/ and /24.31/ are identical among T2.X, T4.X and T6.X. So Section 5.4 will analyze them in the dataset including all the Type II tokens of T2.X, T4.X and T6.X. σ_1_MC_TONE will be examined as one of the predictors.

Although the lexical split of /22.44/ and /22.24/ for T6.X is regular, this split is further investigated using other social or linguistic factors of interest. Mixed models were once again built one by one to compare to random-effects-only model. The results are listed in Table 5.39:

Table 5.39 Results of likelihood ratio test on the split of /22.44/ (n=1286) and /22.24/ (n=394) for Type II T6.X

	deviance	Chisq	Chi Df	Pr(>Chisq)
model-0	248.30			
model-AGE	234.38	13.918	1	0.000***
model-REGION	246.34	1.961	1	0.161
model-SEX	248.04	0.264	1	0.607
model-EDU	247.79	0.514	1	0.473
model-PTH	244.28	4.021	1	0.045*
model-STYLE	248.25	0.057	1	0.811
model-σ_2_MC_TONE	NA			
model-σ_1_MC_AB	248.30	0.001	1	0.970
model-LEXICAL ATTESTEDNESS	248.28	0.021	1	0.884
model-σ_1_HEIGHT	NA			
model-σ_2_HEIGHT	251.71	0.000	6	1.000
model-σ_2_PTH_TONE	NA			
model-σ_1_DUR	247.94	0.361	1	0.548
model-σ_2_DUR	230.56	17.748	1	0.000***
model-σ_2_CODA	247.98	0.325	1	0.569
model-σ_2_MC_TONE (collapsed)	247.63	0.668	1	0.414
model-σ_1_HEIGHT (collapsed)	248.30	0.0006	1	0.980

It can be observed in Table 5.39 that Model-σ_2_MC_TONE, model-PTH_TONE and model-σ_1_HEIGHT cannot be built because of empty cells in the /22.24/ pattern. No phrase of /22.24/ has its σ_2 using MC_T1, MC_T7 and MC_T8, and therefore no phrase of /22.24/ has its σ_1 using PT1, which developed from MC_T1 and MC_T7. Predictor σ_2_MC_TONE was therefore collapsed according to Chan & Ren's (1989) substitution rule found in T4.X. The result of model-σ_2_MC_TONE (collapsed) is also presented in Table 5.39, which is not significantly different. Incidentally, because of the low frequency of occurrence of /22.24/ in T6.X, no phrase uses near close, close, mid, close mid and mid vowels in the /22.24/ pattern, and so σ_1_HEIGHT is collapsed. The conditions of *close, near close, close mid* and *mid vowel* of σ_1_HEIGHT are collapsed into a new condition, "*close and mid vowels*", while *open mid* and *open vowels* of σ_1_HEIGHT are collapsed into a new condition, "*open vowel*". Model-σ_1_HEIGHT (collapsed) was also built but it does not yield a better fit of the data.

Table 5.39 shows that AGE, PTH and σ_2_DUR are significant. AGE and PTH are strongly correlated as only a few old speakers who participated in our experiments reported they were not able to speak PTH, while all young speakers can. According to the model-comparison, model-AGE yields a better fit of the data. Two mixed models with only the significant main effects were then built, one using predictors AGE and $σ_2$_DUR and another one using PTH and $σ_2$_DUR. The results show that using AGE yields a higher C and Somers' D_{xy} of the model, which fits the data better than PTH. So AGE instead of PTH is used to build the final model. Then the interaction of AGE with other independent main effects is investigated but none show significance. So AGE and $σ_2$_DUR were entered into the final model. The output of final model is listed in Table 5.40:

Table 5.40 Output of logistic model with mixed-effects on the split of /22.44/ (n=1286) and /22.24/ (n=394) for Type II T6.X

predictor	*log-odds*	SE	t	*p*
Intercept	0.000	5.540	−3.079	0.002^{**}
AGE(*old*):*young*	0.146	0.649	−2.964	0.003^{**}
$σ_2$_DUR	1.019	0.006	3.273	0.001^{**}

C=0.999, Somer's D_{xy}=0.998

AGE is significant. Young people use less /22.24/ than old people (175 tokens vs 220 tokens). When examining the distribution of /22.24/ by word, it is found that young people use less /22.24/ mainly because they use /55.31/ pattern for 21 tokens belonging to the /22.24/ group and use /22.44/ for 20 tokens belonging to the /22.24/ group. In other words, the /22.24/ group of Type II T6.X is less stable among young people. Meanwhile, $σ_2$_DUR is significant; its log-odds suggests that the longer σ2_DUR, the more likely the speakers use /22.44/.

5.3.3. Summary and discussion

Section 5.3 has aimed to give an overview of the sandhi patterns used in the Wúxī dialect and identify the ongoing changes constrained by AGE, REGION or both. The main findings can be summarized as follows:

(1) Our data confirms that Type I compounds and phrases in Chan & Ren (1989)'s classification (details in Section 2.4.4) use different sandhi patterns with Type II phrases, although their first syllables are from the same MC tonal category.

(2) The /55.31/ pattern is found in T2.X, T4.X and T6.X. It was recorded in Qián (1992) but not in any other study. Qián (1992) found this pattern only in T4.T4, T4.T5 and T4.T7. Our investigation confirms his finding and also shows the occurrence of /55.31/ is an ongoing change mainly used by young people. It is hypothesized that dropping the breathy feature of the nucleus is the cause of those phrases of the Yáng register to switch to /55.31/, a sandhi pattern originally used in the Yīn register.

(3) For Type I constructions, T2.X uses /22.44/, T4.X uses /24.31/ and T6.X uses /22.24/. This is in line with previous findings.

(4) For Type II T2.X phrases, there is a lexically constrained split of /24.31/ and /22.44/. This split is also an ongoing change. Young people use more /24.31/ as they borrowed MC_T2's corresponding PTH tone /35/ to pronounce σ_1.

(5) For Type II T4.X phrases, the substitution rule raised by Chan & Ren (1989) is confirmed in the data. If the σ_2 of a phrase belongs to MC_T3, MC_T4, MC_T5 or MC_T6, the phrase is more likely to use /22.24/. If σ_2 has longer duration, the phrase is more likely to be /22.24/ than /22.44/.

(6) For Type II T6.X phrases, besides two phrases using /24.31/, the other tokens predominantly use /22.44/ and /22.24/. This split is also constrained lexically. Meanwhile, ongoing changes are found that young people are replacing /22.24/ with /22.44/ and /55.31/.

(7) Young people have smaller pitch differences between σ_1 and σ_2 than old people, reflecting contour loss in the sandhi pattern. More concretely, young people raise the overall pitch height in the σ_1 of /22.24/ and /22.44/, while they use rising contours with a less steep slope in the σ_2 of /22.24/. Young people also tend to end with a higher offset in the σ_2 of /22.44/ and /24.31/. These processes are spread over T2.X, T4.X and T6.X as well as Type I and Type II sandhi patterns distinguished by Chan & Ren (1989).

In Section 5.4, the conditioning factors of contour loss found in sandhi forms are investigated. To build the mixed models, five indexes were used to quantify the process of contour loss. The contour loss of σ_1 in /22.24/ is quantified and called *MeanST_(22).24*. The parentheses indicate the syllable involved in variation, i.e., the σ_1 of /22.24/ in this case. "MeanST" is calculated for the level tone to get a value that represents the pitch height of the whole contour. The F_0 values in ST of ten points abstracted from one contour are averaged. Likewise, the contour loss of the σ_1 in /22.44/ is denoted as *MeanST_(22).44.*

The contour loss of σ_2 in /22.24/ is denoted as *SLOPE_22.(24)*, the contour loss of σ_2 in /22.44/ is denoted as *SLOPE_22.(44),* and the contour loss of σ_2 in /24.31/ is denoted as *SLOPE_24.(31)*. I use parentheses to indicate the syllable involved in variation. "SLOPE" is calculated using different algorithms for rising contours (σ_2 in /22.24/) and falling contours (σ_2 in /22.44/ and /24.31/). The slope between P_1 and

P_{max} is calculated for the steepness of rising contours while the slope between P_{max} and P_{10} is calculated for the steepness of falling contours. The slope between P_{max} and onset or offset is calculated to present the steepness of rising or falling contours because, irrespective of being a rising tone or a falling tone, P_{max} is the target for tone realization (Zhū 2010: 277). According to the observations in Figures 5.7 to 5.10, *MeanST_(22).24, MeanST_(22).44, SLOPE_22.(24)* and *SLOPE_24.(31)* are mainly constrained by AGE, but *SLOPE_22.(44)* is constrained by both AGE and REGION. So in Section 5.4, *SLOPE_22.(44)* is discussed first.

5.4. Analyzing the linguistic variables of tone sandhi

In this section, five indexes of contour loss process are investigated across T2.X, T4.X and T6.X. Then σ_1_PTH_TONE and σ_1_MC_TONE are added as independent variables to build mixed models, to see if the variation can be predicted by the tonal categories of σ_1.

5.4.1. SLOPE_22.(44)

First, linear mixed models were built on *SLOPE_22.(44)*. Six social factors and 11 linguistic factors were tested by the likelihood ratio test one by one. A series of mixed models were built for each factor: model-0 with only the random factor, and model-X with both the random factor and the fixed factor, in which X denotes the only fixed factor entered in the model.

Table 5.41 Results of likelihood ratio test on *SLOPE_22.(44)* (n=2643)

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	29338			
model-AGE	29322	16.459	1	0.000***
model-REGION	29331	6.983	1	0.008**
model-SEX	29337	1.003	1	0.317
model-EDU	29179	159.9	1	0.000***
model-PTH	29178	160.25	1	0.000***
model-STYLE	29325	13.053	1	0.000***
model-LEXICAL ATTESTEDNESS	29338	0.001	1	0.979
model-σ_1_MC_TONE	29331	7.134	2	0.028*
model-σ_1_PTH_TONE	29325	13.259	2	0.001**
model-σ_1_MC_AB	29337	1.7	1	0.192
model-σ_1_HEIGHT	29326	11.991	6	0.062
model-σ_1_DUR	29338	0.023	1	0.879
model-σ_2_MC_TONE	29328	10.051	7	0.186
model-σ_2_PTH_TONE	29338	0.928	3	0.819
model-σ_2_HEIGHT	29332	6.344	6	0.386
model-σ_2_DUR	29318	20.388	1	0.000***
model-σ_2_CODA	29333	5.654	1	0.017*

As shown in Table 5.41, AGE, REGION, EDU, PTH, σ_2_DUR σ_2_CODA, σ_1_PTH_TONE and σ_1_MC_TONE are significant. As σ_1_PTH_TONE and σ_1_MC_TONE are highly correlated, given the lower deviation σ_1_PTH_TONE yielded, σ_1_PTH_TONE is entered into the final model rather than σ_1_MC_TONE. The constraining effect of EDU and PTH may come from AGE, as all the young people are PTH speakers and more young people are college educated. However, further likelihood ratio tests show that they are independent of AGE as they also show significant effects in the data of old people alone (model-EDU: $\chi^2(1)=148.27$ $p=0.000^{***}$; model-PTH: $\chi^2(1)=148.41$, $p=0.000^{***}$). Four interactions - AGE*STYLE, AGE:σ_2_CODA, AGE*REGION and REGION*STYLE - are significant. However, after putting all of them into the final model, only the first two show a significant difference. So, the other two are removed from the final model.

Table 5.42 shows that urban speakers have higher *SLOPE_22.(44)*, i.e., a falling tone in a less steep slope, than suburban speakers. This can also be observed in Figure 5.6 to Figure 5.9. Speakers make less steep falls for *SLOPE_22.(44)* in the reading of paragraphs than of wordlists. σ_2_DUR shows that the longer the duration of σ_2, the steeper the falling contour is. This trend is also reflected in σ_2_CODA. σ_2_CODA shows that smooth syllables have steeper falling tones than the checked syllables in the old people.

Table 5.42 Output of mixed model on *SLOPE_22.(44)* (n=2643)

predictor	*β*	SE	t	*p*
Intercept	-15.633	8.939	−1.749	0.085
REGION *(suburban)*: *urban*	16.901	4.391	3.849	0.000^{***}
AGE (*old*): *young*	3.763	6.639	0.567	0.572
EDU *(college): non-college*	5.668	4.645	1.220	0.231
PTH *(cannot speak): can*	−8.556	7.035	−1.216	0.232
STYLE *(paragraph): wordlist*	-16.781	3.447	−4.869	0.000^{***}
σ_1_PTH_TONE *(PT2):*				
PT3	3.532	2.473	1.428	0.157
PT4	−3.567	1.927	−1.851	0.067
σ_2_DUR	−0.034	0.010	−3.261	0.000^{***}
σ_2_CODA (*checked syllable*)				
smooth syllable	−9.009	2.324	−3.877	0.000^{***}
AGE*STYLE				
young:wordlist	8.952	4.250	2.106	0.035^{*}
AGE*σ_2_CODA				
young: smooth syllable	10.668	2.840	3.757	0.000^{***}

AGE is not significant for the intercept but is involved in two significant interactions. Generally speaking, young speakers have larger *SLOPE_22.(44)* than old speakers.

The significant AGE*STYLE shows that old people have greater style differences than young people. This is also reflected in Table 5.43. Regarding AGE*σ_2_CODA, old speakers have higher *SLOPE_22.(44)* in the reading of checked syllables than in the reading of smooth syllables but young speakers do not show obvious difference. In other words, old speakers tend to differ in *SLOPE_22.(44)* between smooth and checked syllables, but young people have lost this distinction. Table 5.44 lists the mean scores of *SLOPE_22.(44)* split by AGE and σ_2_CODA, which can also present this contrast. This contrast might be correlated with another issue, that young speakers are losing the glottal stop of checked syllable (Cao, 2012). However, at this point it is hard to say if they have a causal connection.

Table 5.43 Mean *SLOPE_22.(44)* in WXD, split by AGE and STYLE

STYLE \ AGE	*old*	*young*
wordlist	−43.68	−21.54
paragraph	−26.31	−14.18

Table 5.44 Mean *SLOPE_22.(44)* in WXD, split by AGE and σ_2_CODA

σ2 CODA \ AGE	*old*	*young*
smooth syllable	−44.71	−20.83
checked syllable	−33.74	−21.75

5.4.2. *MeanST_(22).44*

In the same way, *MeanST_(22).44* is examined by the mean pitch of ten points on the contour. The observations of Figure 5.6 to Figure 5.9 imply that, on the whole, the level tone of σ_1 is higher among young than old people.

Table 5.45 shows that AGE, EDU, PTH and σ_1_HEIGHT are significant main effects. EDU and PTH may correlate with AGE. However, further likelihood ratio tests show that they are independent of AGE as they are also significant in the data of old people alone (model-EDU: $\chi^2(1)=25.755$, $p=0.000^{***}$; model-PTH: $\chi^2(1)=27.004$, $p=0.000^{***}$). Then AGE, EDU, PTH and σ_1_HEIGHT are entered into the final mixed model together. The significant interaction is AGE*STYLE.

Table 5.45 Results of likelihood ratio test on the σ₁_MeanST for *MeanST_(22).44* (n=2643)

	deviance	Chisq	ChiDf	Pr(>Chisq)
model-0	10486			
model-AGE	10460	25.667	1	0.000***
model-REGION	10486	0.001	1	0.974
model-SEX	10485	0.017	1	0.895
model-EDU	10459	26.261	1	0.000***
model-PTH	10459	26.229	1	0.000***
model-STYLE	10483	2.635	1	0.105
model-LEXICAL ATTESTEDNESS	10483	2.621	1	0.105
model-σ_1_MC_TONE	10480	5.731	2	0.057
model-σ_1_PTH_TONE	10481	4.762	2	0.092
model-σ_1_MC_AB	10485	0.631	1	0.427
model-σ_1_HEIGHT	10455	30.358	5	0.000***
model-σ_1_DUR	10483	2.118	1	0.146
model-σ_2_MC_TONE	10479	6.527	7	0.48
model-σ_2_PTH_TONE	10485	0.16	3	0.984
model-σ_2_HEIGHT	10484	1.267	5	0.938
model-σ_2_DUR	10485	0.105	1	0.746
model-σ_2_CODA	10483	2.857	1	0.091

Table 5.46 Output of mixed model of *MeanST_(22).44* (n=2643)

predictor	***β***	**SE**	**t**	***p***
Intercept	−3.025	0.707	−4.277	0.000***
AGE (*old*): *young*	3.301	0.439	7.518	0.000***
EDU *(college): non-college*	0.348	0.401	0.868	0.391
PTH *(cannot speak): can*	−1.028	0.621	−1.655	0.107
σ_1_HEIGHT *(close):*				
near close	−0.198	0.192	−1.028	0.307
close mid	−0.509	0.211	−0.515	0.518
mid	−0.102	0.335	−0.306	0.760
open mid	−0.494	0.180	−2.741	0.007**
open	−0.949	0.180	−5.259	0.000***
STYLE *(paragraph): wordlist*	0.290	0.273	1.063	0.290
AGE*STYLE *young:wordlist*	−0.765	0.183	−4.194	0.000***

Table 5.46 confirms that young people use higher pitch to pronounce the σ_1 of *MeanST_(22).44.* Besides, the intrinsic vowel height also plays a role. Open mid and open vowels have lower pitch heights than close vowels. This confirms that close vowels (low F_1) have a higher F_0 than low vowels or close vowels. STYLE is not significant as a main effect but its interaction is significant. Table 5.47 below shows that old people do not differ considerabley for *MeanST_(22).44* between two styles (wordlist reading vs paragraph reading) but young people show significant differences. Young people tend to use higher pitch in the reading of paragraph to save articulatory efforts.

Table 5.47 Mean *MeanST_(22).44* in WXD, split by AGE and STYLE

STYLE \ AGE	*old*	*young*
wordlist	−3.67	−1.51
paragraph	−3.82	−0.91

MeanST_(22).24

MeanST_(22).24 is defined as the mean pitch of ten points on the contour. Figure 5.6 to Figure 5.9 also illustrate that, generally, the level tone of σ_1 is higher among young people than old people.

Table 5.48 Results of likelihood ratio test on *MeanST_(22).24* (n=1172)

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	5471.3			
model-AGE	5445.7	25.57	1	0.000***
model-REGION	5471.2	0.009	1	0.924
model-SEX	5471.2	0.028	1	0.868
model-EDU	5430.1	41.194	1	0.000***
model-PTH	5430.5	40.744	1	0.000***
model-STYLE	5470.5	0.782	1	0.377
model-LEXICAL ATTESTEDNESS	5471.2	0.038	1	0.847
model-σ_1_MC_TONE	5471.1	0.191	1	0.663
model-σ_1_PTH_TONE	5470.9	0.324	1	0.57
model-σ_1_MC_AB	5471.0	0.267	1	0.606
model-σ_1_HEIGHT	5451.0	20.296	4	0.000***
model-σ_1_DUR	5469.3	1.967	1	0.161
model-σ_2_MC_TONE	5458.2	13.097	7	0.07
model-σ_2_PTH_TONE	5467.7	3.52	3	0.318
model-σ_2_HEIGHT	5466.2	5.019	5	0.414
model-σ_2_DUR	5471.0	0.225	1	0.635
model-σ_2_CODA	5471.1	0.187	1	0.666

Table 5.48 shows that AGE, EDU, PTH, and σ_1_HEIGHT are significant main effects. EDU and PTH are also independent of AGE as they are also significant in the data of old people along (model-EDU: $\chi^2(1)=44.013$, $p=0.000^{***}$; model-PTH: $\chi^2(1)=45.138$, $p=0.000^{***}$). Like for *MeanST_(22).44*, AGE*STYLE is also significant.

Table 5.49 shows that AGE, σ_1_HEIGHT, STYLE and AGE*STYLE are significant. Young people have higher pitch contours in the σ_1 of /22.44/. Open mid and open vowels have significantly lower pitch heights than close vowels. EDU and PTH do not show any significance in the final model. They are not significant either when using other intercepts. Their effects are covered by AGE and σ_1_HEIGHT. Significant STYLE and AGE*STYLE show that old people use higher pitch contour in the wordlist reading (positive β) but young people use lower pitch contour in the wordlist reading (1.196−1.540=−0.344), which can be seen more obviously in Table 5.50 below. Again, young people tend to save articulatory efforts in less formal reading style.

Table 5.49 Output of mixed model on *MeanST_(22).24* (n=1172)

predictor	***β***	**SE**	**t**	***p***
Intercept	−3.592	0.602	−5.963	0.000***
AGE (*old*): *young*	3.895	0.473	8.238	0.000***
EDU *(non-college): college*	−0.069	0.409	−0.168	0.868
PTH *(cannot speak): can*	−0.855	0.633	−1.351	0.185
σ_1_HEIGHT *(close):*				
near close	−0.240	0.142	−1.682	0.099
close mid	−0.573	0.370	−1.548	0.128
open mid	−0.357	0.161	−2.223	0.031*
open	−0.884	0.178	−4.978	0.000***
STYLE *(paragraph): wordlist*	1.196	0.287	4.169	0.000***
AGE*STYLE *young:wordlist*	−1.540	0.240	−6.408	0.000***

Table 5.50 Mean *MeanST_(22).24* in WXD, split by AGE and STYLE

STYLE \ AGE	*old*	*young*
wordlist	−3.44	−1.28
paragraph	−4.24	−0.44

5.4.3. SLOPE_22.(24)

Linear mixed models are built on *SLOPE_22.(24)*. Likelihood ratio tests are performed as before.

Table 5.51 Results of likelihood ratio test on *SLOPE_22.(24)* (n=1172)

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	14466			
model-AGE	14451	15.224	1	0.000***
model-REGION	14466	0.739	1	0.39
model-SEX	14466	0.188	1	0.664
model-EDU	14337	129.31	1	0.000***
model-PTH	14337	129.19	1	0.000***
model-STYLE	14464	2.054	1	0.152
model-LEXICAL ATTESTEDNESS	14460	6.805	1	0.009**
model-σ_1_MC_TONE	14464	2.19	1	0.139
model-σ_1_PTH_TONE	14465	1.365	1	0.243
model-σ_1_MC_AB	14465	1.846	1	0.174
model-σ_1_HEIGHT	14460	6.736	4	0.151
model-σ_1_DUR	14464	2.003	1	0.157
model-σ_2_MC_TONE	14460	6.893	7	0.44
model-σ_2_PTH_TONE	14466	0.643	3	0.887
model-σ_2_HEIGHT	14459	7.445	5	0.19
model-σ_2_DUR	14465	1.3	1	0.254
model-σ_2_CODA	14466	0.603	1	0.437

Table 5.51 shows that AGE, EDU, PTH and LEXICAL ATTESTEDNESS are significant main effects. EDU and PTH are independent of AGE as they are also significant in the data of old people along (model-EDU: $\chi^2(1)=131.22$, $p=0.000^{***}$; model-PTH: $\chi^2(1)=131.35$, $p=0.000^{***}$). Again, no interaction is significant for *SLOPE_22.(24)*.

Table 5.52 Output of mixed model on *SLOPE_22.(24)* (n=1172)

predictor	*β*	SE	t	*p*
Intercept	27.561	4.508	6.114	0.000^{***}
AGE (*old*): *young*	−10.903	2.702	−4.035	0.000^{***}
EDU *(non-college): college*	−0.260	2.675	−0.097	0.923
PTH *(cannot speak): can*	2.784	4.140	0.673	0.506
LEXICAL ATTESTEDNESS *(PTH&Wú lexical item): Wú lexial item*	4.120	1.485	2.775	0.008^{**}

Table 5.52 indicates that the rising tone used by young people is less steep than those of old speakers in the σ_2 of /22.24/. The constraining effects of EDU and PTH are not significant when entered simultaneously with AGE and LEXICAL ATTESTEDNESS. *Wú lexical items* have much higher slopes in the σ_2 of /22.24/ than *PTH&Wú* lexical items, indicating conservative features. *SLOPE_22.(24)* is not constrained by STYLE but it is relevant to *MeanST_(22).24* because the likelihood ratio test shows that, in sandhi pattern /22.24/, the mean pitch of its σ_1 is responsible for the steepness of its σ_2 ($\chi^2(1)=13.020$, $p=0.000^{***}$). The higher the pitch of σ_1, the less steep σ_2's rising slope. In other words, the contour loss σ_2 is relevant with the contour loss of σ_1. The contour loss of σ_1 is constrained by STYLE then *SLOPE_22.(24)* is indirectly constrained by STYLE.

5.4.4. *SLOPE_24.(31)*

Finally, linear mixed models were built for *SLOPE_24.(31)* to investigate the contour loss in /24.31/. First, likelihood ratio tests are conducted on it in Table 5.53.

Table 5.53 shows that AGE, EDU, PTH, STYLE, σ_2_DUR and σ_2_CODA are significant main effects. EDU and PTH are independent of AGE as they are also significant in the data of old people along (model-EDU: $\chi^2(1)=160.01$, $p=0.000^{***}$; model-PTH: $\chi^2(1)=159.5$, $p=0.000^{***}$). AGE*σ_2_CODA is the only significant interaction found for *SLOPE_24.(31)*. Then six main effects and one interaction are entered together into the final model and yield the results in Table 5.54.

Table 5.53 Results of likelihood ratio test on *SLOPE_24.(31)* (n=1699)

	deviance	Chisq	ChiDf	Pr(>Chisq)
model-0	22448			
model-AGE	22433	15.318	1	0.000***
model-REGION	22447	1.755	1	0.185
model-SEX	22448	0.58	1	0.447
model-EDU	22288	159.87	1	0.000***
model-PTH	22287	161.65	1	0.000***
model-STYLE	22425	23.181	1	0.000***
model-LEXICAL ATTESTEDNESS	22445	3.504	1	0.061
model-σ_1_MC_TONE	22443	5.736	2	0.057
model-σ_1_PTH_TONE	22443	5.248	2	0.073
model-σ_1_MC_AB	22448	0.188	1	0.665
model-σ_1_HEIGHT	22440	8.677	5	0.123
model-σ_1_DUR	22445	3.397	1	0.065
model-σ_2_MC_TONE	22436	12.166	7	0.095
model-σ_2_PTH_TONE	22442	6.037	3	0.11
model-σ_2_HEIGHT	22440	8.459	5	0.133
model-σ_2_DUR	22247	201.66	1	0.000***
model-σ_2_CODA	17974	4474.6	1	0.000***

Table 5.54 Output of mixed model on *SLOPE_24.(31)* (n=1699)

predictor	***β***	**SE**	**t**	***p***
Intercept	-94.765	9.065	−10.454	0.000***
AGE (*old*): *young*	29.931	5.108	5.859	0.000***
EDU *(college): non-college*	7.202	4.351	1.655	0.107
PTH *(cannot speak): can*	7.442	6.768	1.100	0.279
STYLE *(paragraph): wordlist*	−11.066	3.828	-2.891	0.005**
σ_2_CODA (*checked syllable*) smooth syllable	10.355	4.104	2.523	0.014*
σ_2_DUR	0.196	0.016	12.003	0.000***
AGE*σ_2_CODA *young:smooth syllable*	−11.222	3.295	-3.406	0.001***

Table 5.54 shows that young people have reduced the steepness of falling tones in comparison with old speakers. Like *SLOPE_22.(44)*, falling tokens are less steep in the reading of paragraphs than in the reading of word lists. Smooth syllables have larger *SLOPE_24.(31)* than checked syllables, which contrasts to its effect on *SLOPE_22.(44)* (see Section 5.4.1). Like *SLOPE_22.(44)*, σ_2_DUR is significant but plays an opposite role: the longer the duration, the more contour loss *SLOPE_24.(31)* has. This suggests that the contour loss of the σ_2 of /24.31/ may have a different mechanism to the contour loss of the σ_2 of /22.44/, though they are both falling tones.

Table 5.55 Mean *SLOPE_(24).31* in WXD, split by AGE and σ_2_CODA

σ2_CODA / **AGE**	*old*	*young*
smooth syllable	−51.01	−35.65
checked syllable	−67.53	-42.10

/22.42/, a low level tone plus a high falling tone, is in a "conflicting tonal context", "in which adjacent phonetic units have very different values along that phonetic dimension" (Y. Xu, 1994). It was found by Xu (1994) that a phonetic unit deviated more from its underlying form when produced in a conflicting context. It implies that to fully realize such an underlying form in a conflicting phonetic context requires extra efforts. In the checked syllable, the duration of which is much shorter than that of the smooth syllable, it is even harder to fully realize the high falling tone for the σ_2. So the checked syllable of /22.42/ tends to reduce the steepness of its falling trend to increase the naturalness and to minimize articulatory effort. However, /24.31/ is in "a compatible context" (ibid.), the falling tone of its σ_2 agrees with the high ending of its σ_1. To reach the same pitch decline in a shorter time, undoubtedly, the slope has to increase. So the contrary effects of σ_2_CODA and σ_2_DUR are related to tonal shapes of adjacent syllables. It also helps us further understand that *SLOPE_22.(44)* is phonetically motivated to change from [22.42] to [22.44]. So it is an internally motivated change.

5.4.5. Summary and discussion

Section 5.4 aims to examine five ongoing changes of Wúxī sandhi patterns by mixed models: *SLOPE_22.(44)*, *MeanST_(22).44*, *MeanST_(22).24*, *SLOPE_22.(24)*, *and SLOPE_24.(31)*. The main findings can be summarized as follows:

(1) Besides *SLOPE_22.(24)*, the other four variables are constrained by STYLE, either being one of main effects or showing interaction with AGE. This confirms that contour loss is mainly motivated by articulatory reduction in the less formal style (i.e., less monitored). The steepness of the σ_2 of /22.24/, i.e. *SLOPE_22.(24)* is constrained by STYLE indirectly due to its relevance with *MeanST_(22).24*.

(2) σ_1_HEIGHT is found to be a significant predictor for *MeanST_(22).44* and *MeanST_(22).24*, two variables being quantified by pitch height. On the one hand, it confirms the intrinsic vowel effect (Hombert, et al., 1979; Lehiste, 1970); on the other hand, it suggests that σ_1_HEIGHT or σ_2_HEIGHT is more likely to be significant predictors for the dependent variable associated to F_0 rather than slope or pitch shapes.

(3) *SLOPE_22.(44)* and *SLOPE_24.(31)* show similarities in the development. The conservative form of *SLOPE_22.(44)* has a falling trend in the end of its σ_2

while its advanced form is a high level tone. The conservative form of *SLOPE_24.(31)* is a falling tone with steep slope (a negative score in high absolute value) while its advanced form is a falling tone with less steep slope. σ_2_CODA and σ_2_DUR are both found to be predictors for those two variables but both show inverse effects. Checked syllables tend to have contour loss for *SLOPE_22.(44)* but are more resistant to the contour loss of *SLOPE_24.(31).* The longer the duration σ_2 has, the more contour loss *SLOPE_24.(31)* has but the less contour loss *SLOPE_22.(44)* has. σ_2_CODA and σ_2_DUR play different roles on the contour loss of the σ_2 of /24.31/ and /22.44/ because the former is in a "compatible tonal context" while the latter is in a "conflicting tonal context" (Y. Xu, 1994).

5.5. Summary and discussion of Chapter 5

With reference to Goal I, four ongoing changes in citation tones were identified and described in Section 5.1: *WXfalling, WXlevel, WXT4* and *WXT2/6* and analyzed in Section 5.2. In addition, the contour loss of three sandhi patterns - /22.44/, /22.24/ and /24.31/ - were described and analyzed in Sections 5.3 and 5.4.

WXfalling and *WXlevel* refer to the small number of falling (n=212) and level (n=239) tones in our Wúxī data, which are probably borrowings of PTH falling tone and neutral tone. These borrowings are easily detected because there are major distinctions of tonal shapes or pitch height between the borrowing tone and the original tone. However, some non-falling or non-level realizations in our dataset can also be the result of tone borrowing, but are hard to detect when the dialect tone and PTH tone share the same tonal shape. The majority of non-falling or non-level occurrences in T2, T4 and T6 are rising. When analyzing *WXT2/6* and *WXT4* in Section 5.2, it was found that, especially among young people, the rising contours of T2 tokens were also influenced by the corresponding PTH T2, a mid rising tone. Therefore, *WXfalling* and *WXlevel* make up only a portion of tone borrowings in T2, T4 and T6 and are called "explicit borrowing" while the borrowing of PT2 in *WXT4* and *WXT2/6* are called "implicit borrowing". Mixed models were only built for explicit borrowings in this chapter. The analyses of *WXfalling* shows that suburban young people borrow more falling tones than urban young people (See Table 5.8), reflecting that suburban young people have a greater tendency towards linguistic insecurity about their local dialect.

WXT4 has three discrete variants: peaking, low rising and low dipping tones. As shown in Figure 5.5, peaking tone is the most conservative variant, low rising tone is the intermediate one and low dipping tone is the final stage. In order to further investigate whether *WXT4* is a gradual change or an abrupt change, each of three variants was explored by mixed modeling. The independent variables include ST_{Pmax}, ST_{P1}, ST_{P10} and NT_{Pmax} of peaking T4, NT_{Pmin} of low rising and dipping T4. The results of mixed models built on those variables showed that NT_{Pmax} of peaking T4 and NT_{Pmin} of dipping T4 are constrained by STYLE (Table 5.14 and Table 5.23). This suggests that the P_{max} of peaking variant and P_{min} of dipping variant are moving backwards for saving articulatory efforts. Meanwhile, the delayed NT_{Pmin} of dipping T4 is also motivated by borrowing PT3 and merging with T2/6. So, generally speaking, *WXT4* is a gradual change in its early stages (Step 1 to 3 in Figure 5.5) and thus it is concluded that WXT4 is triggered internally. However, in its final stages, due to the contact with PTH, WXT4 unavoidably involved some borrowings of PT3 as the delayed rising WXT4 (/113/) and the dipping PT3 (/213/) are very close in terms of tonal shapes and pitch height.

Since those three variants of *WXT4* are three different tonal shapes, no uniform index like slope or pitch height is available to separate them, so two logistic mixed models were built. The results of Table 5.27 also confirm that *WXT4* is involved in a change in progress constrained by internal factors (e.g., MC_AB, MORPHEME, DUR) and external factors (e.g. AGE, REGION, PTH). *WXT4* is a good example to reveal the interplay of internal factors and external factors in terms of Goal II. It is triggered internally while its transmission is constrained externally by REGION, where its advanced forms are spread more slowly in suburban areas than in urban areas.

WXT2/6 refers to the contour loss process in which Wúxī T2/6 rising contours are getting flattened in the young generation. The index for this variable is m_{rising}, which is the slope between P_{min} and P_{max} on the low rising contour. A mixed model built on m_{rising} shows that *WXT2/6* is constrained by AGE, STYLE and PTH_TONE. The conditioning effects of STYLE show that the contour loss of WXT2/6 is more likely to occur in the paragraph reading than in word list reading, as a result of saving speech effort. However, it is hard to judge here whether *WXT2/6* is contact-driven since no regional contrasts and PTH constrains are found in Wúxī data. So whether contour loss is externally constrained needs to be further investigated in the Shànghǎi data.

Then turning to ongoing changes in tone sandhi, contour loss of three sandhi patterns - /22.44/, /22.24/ and /24.31/ - were investigated. These are quantified by five dependent variables: *SLOPE_22.(44)*, *MeanST_(22).44*, *MeanST_(22).24*, *SLOPE_22.(24)*, *and SLOPE_24.(31)*.

The mixed model analysis in Section 5.4 already showed the special status of *SLOPE_22.(44)*. It is constrained both by AGE (in the interactions) and REGION, whereas the other four ongoing changes identified in sandhi forms are only constrained by REGION. Suburban speakers use more conservative forms, the falling tone in the σ_2 of /22.44/, than urban speakers. But for the other four ongoing changes, no urban-suburban differences can be distinguished.

To link this evidence with citation tone, it has been found that REGION is a significant predictor for *WXT4* but not for *WXT2/6*. Suburban speakers tend to be conservative in their use of *WXT4* while suburban and urban speakers do not differ greatly for *WXT2/6*. Meanwhile, suburban young speakers are more advanced than urban young speakers for *WXfalling*. These findings together suggest that *WXT4*, *WXT2/6* and *WXfalling* follow three different patterns in the urban-suburban dimension. The internally triggered *WXT4* seems to be associated with the suburban conservativeness while the externally triggered *WXfalling* seems to be associated with the suburban lead.

Chapter 6. Tonal variation in Shànghǎi

In this chapter, I present and discuss the results of T2, T4 and T6 variation in the Shànghǎi data. Following the structure of Chapter 5, the first task is to identify the tonal variables involved in ongoing changes in the dialects spoken in Shànghǎi, including the Shànghǎi urban dialect (hereafter SHUD), the Nánhuì dialect (hereafter NHD), the Bǎoshān dialect (hereafter BSD) and the Sōngjiāng dialect (hereafter SJD). The identification of citation tone is done in Section 6.1 and the tone sandhi work is done in Section 6.3. After having defined the tonal variables, regression analyses are conducted to interpret the variation patterns from both linguistic and social perspectives. Sections 6.2 and 6.4 present the analyses and interpretation for the citation tone and tone sandhi separately. Finally, Section 6.5 gives a brief summary and the main conclusions of this chapter.

6.1. Identifying the linguistic variables of citation tone

Section 2.5.2.1 reviewed several Shànghǎi dialect studies and found that: (1) T2, T4 and T6 in SHUD and NHD have merged into a rising tone; and (2) T4 and T6 in SJD and NHD have merged into a rising tone while T2 is a mid falling tone /31/ independently (Qián, 1988).

After transcribing all the tokens of the current dataset, high falling tone (/51/) is occasionally (<5%) found in T2, T4 and T6 in all the four regions. These tone patterns were not documented before in Shànghǎi. Moreover, level tone /33/ is found in the four regions as well with very low frequency (<2%). The high falling tone and mid level tone were also identified in Wúxī and found to be borrowings from PTH. Is Shànghǎi also undergoing such borrowings? Section 6.1.1 examines this issue. Besides SHUD, the tones of Shànghǎi suburban dialects have been rarely studied acoustically. Thus, another goal of this section is to describe T2, T4 and T6 of the urban area and three suburban dialects on the basis of acoustic data and to identify ongoing changes in these regions. SHUD, NHD, BSD and SJD are presented in Sections 6.1.2 to 6.1.5 separately.

6.1.1. High falling and mid level tones

The distribution of all tonal types transcribed for T2, T4 and T6 morphemes in four Shànghǎi regions are presented in Table 6.1:

Table 6.1 Distribution of tonal types for T2, T4 and T6 morphemes in Shànghǎi

	SHUD		NHD		BSD		SJD	
	old	young	old	young	old	young	old	young
falling	40	24	30	35	25	36	51	64
/51/	3.3%	1.8%	2.2%	2.8%	1.8%	2.94%	3.8%	4.7%
level	17	11	16	30	6	2	17	19
/33/	1.4%	0.8%	1.2%	1.8%	0.4%	0.16%	1.3%	1.4%
rising	1158	1323	1290	1210	781	864	907	863
/113/, /13/	95.3%	97.4%	96.6%	95.4%	57.5%	70.47%	67.0%	63.9%
peaking	0	0	0	0	564	324	379	405
/131/	0	0	0	0	40.2%	26.43%	28.0%	30.0%
total	1215	1358	1336	1275	1376	1226	1342	1351

Table 6.1 lists both the frequency and percentage of each tonal category in each region, split by age. Results show that both high falling tone and mid level tone count for only a small part of the total realizations, between 0.2% and 4.7%. Comparing old and young, it is apparent from this table that they do not differ widely in the amount of /51/ and /33/. These differences were tested in mixed model and the results are listed in Table 6.2:

Table 6.2 Results of AGE effect on the variation of falling vs non-falling tones and level vs non-level tones in SHUD, NHD, BSD and SJD

dependent variable		**Deviance**	**Chisq**	**Chi Df**	**Pr(>Chisq)**
SHUD	model-0	510.37			
falling vs non-falling (n=64 vs n=2509)	model-AGE	507.76	2.612	1	0.106
SHUD	model-0	277.82			
level vs non-level (n=28 vs n=2545)	model-AGE	276.34	1.478	1	0.224
NHD	model-0	586.33			
falling vs non-falling (n=65 vs n=2546)	model-AGE	585.96	0.371	1	0.542
NHD	model-0	382.82			
level vs non-level (n=46 vs n=2565)	model-AGE	382.26	0.561	1	0.454
BSD	model-0	769.83			
falling vs non-falling (n=61 vs n=2541)	model-AGE	767.86	1.979	1	0.160
BSD	model-0	101.78			
level vs non-level (n=8 vs n=2594)	model-AGE	100.85	0.930	1	0.335
SJD	model-0	836.00			
falling vs non-falling (n=115 vs n=2578)	model-AGE	835.52	0.488	1	0.485
SJD	model-0	368.83			
level vs non-level (n=36 vs n=2657)	model-AGE	368.71	0.116	1	0.733

Two mixed models were built for each variable, model-0 using only the random factors as predictors, and model-AGE using both main effect AGE and random factors as predictors. If model-AGE fits the data significantly better than model-0, AGE is a significant predictor. Results in Table 6.2 confirm that there are no significant differences between old and young people in the use of falling or level tones in all four regions. Contrary to Wúxī, there is no change in progress of falling and level tones in Shànghǎi.

It was found in Wúxī that falling and level tokens were more likely to occur in PTH T4 morphemes, suggesting that falling and level tokens are PTH borrowings. This issue is also examined in the Shànghǎi data. The same method of likelihood ratio test was used. For each dependent variable, two mixed models were built, model-0 using only the random factors, and model- PTH_TONE using both PTH_TONE and random factors as predictors. The results are listed in Table 6.3:

Table 6.3 Results of PTHtone effect on the variation of falling vs non-fallling tones and level vs non-level tones in SHUD, NHD, BSD and SJD

dependent variable		**deviance**	**Chisq**	**Chi Df**	**Pr(>Chisq)**
SHUD	model-0	510.37			
falling vs non-falling (n=64 vs n=2509)	model-PTH_TONE	497.67	12.705	2	0.002**
SHUD	model-0	277.82			
level vs non-level (n=28 vs n=2545)	model-PTH_TONE	272.4	5.425	2	0.066
NHD	model-0	586.33			
falling vs non-falling (n=65 vs n=2546)	model-PTH_TONE	579.15	7.181	2	0.028*
NHD	model-0	382.82			
level vs non-level (n=46 vs n=2565)	model-PTH_TONE	376.44	6.379	2	0.041*
BSD	model-0	769.83			
falling vs non-falling (n=61 vs n=2541)	model-PTH_TONE	766.59	3.246	2	0.197
BSD	model-0	101.78			
level vs non-level (n=8 vs n=2594)	model-PTH_TONE	101.21	0.568	2	0.753
SJD	model-0	836.00			
falling vs non-falling (n=115 vs n=2578)	model-PTH_TONE	835.64	0.361	2	0.835
SJD	model-0	368.83			
level vs non-level (n=36 vs n=2657)	model-PTH_TONE	368.08	0.747	2	0.689

Table 6.3 shows that PTH_TONE has a significant effect on three dependent variables: falling tones in SHUD, falling tones in NHD, and level tones in NHD. The effect of PTH_TONE on level tones in SHUD is fairly close to significance. For

those four variables, PTH T4 morphemes are more likely to be realized as falling or level tones than PTH T2 morphemes. However, PTH_TONE does not have a significant effect in SJD and BSD. These results suggest that the factors influencing the realizations of falling tones and level tones are not the same in the different areas of Shànghǎi. The falling and level tones in SHUD and NHD are likely to be PTH borrowings, whereas this is not the case in Bǎoshān and Sōngjiāng. These two regions are suburbs and further away from the Shànghǎi urban area, so they might be less influenced by PTH.

6.1.2. Shànghǎi urban dialect (SHUD)

SHUD is one of the most well studied Chinese dialects. The merger of its T2, T4 and T6 has been well documented since Chao (1928). As can be seen in Table 6.1, besides the small amount of falling and level tones, only rising tone /13/ is found in the urban data. To visualize this rising tone, F_0 contours of T2A, T2B, T4A, T4B, T6A and T6B morphemes are illustrated in Figure 6.1, split by AGE.

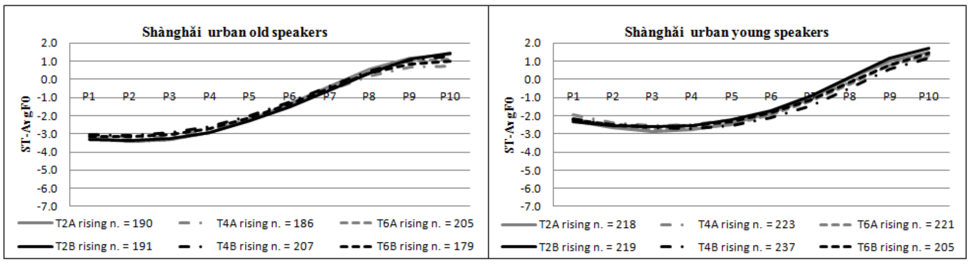


Figure 6.1 Mean F_0 contours of rising tones in SHUD, split by AGE, MC_TONE and MC_AB

It can be seen from Figure 6.1 that old and young people have completely merged T2, T4 and T6 into /13/. It can also be seen that the rising contours of old and young people are quite similar. Unlike T2/6 of the Wúxī dialect, young people keep the steepness of rising contours and do not have contour loss. Finally, it can also be observed from Figure 6.1 that young people have slightly higher onset and later P_{min}, causing their entire contour to be a delayed rising contour, like /113/.

These observations were then analyzed statistically. First, in order to test the merger of T2, T4 and T6, three indexes - the slope between P_{min} and P_{max}, (m_{rising}), the pitch height of P_{min} (ST_{Pmin}) and P_{max} (ST_{Pmax}) - were tested to see if they are conditioned by MC_TONE. Two mixed models were built for each index/dependent variable: (1) model-0 using random factors as predictors, and (2) model-MC_TONE using MC_TONE and random factors as predictors.

Table 6.4 The merger of T2, T4 and T6 in SHUD (n=1158)

dependent variable		deviance	Chisq	Chi Df	Pr(>Chisq)
SHUD m_{rising}	model-0	18148			
	model-MC_TONE	18144	4.694	2	0.096
SHUD ST_{Pmin}	model-0	7902.7			
	model-MC_TONE	7900.8	1.890	2	0.389
SHUD ST_{Pmax}	model-0	9887.6			
	model-MC_TONE	9885.4	2.211	2	0.331

As seen from Table 6.4, none of the three indexes is significant, so the merger of T2, T4 and T6 of SHUD is confirmed in my dataset. These three tones are hereafter called SHUD T2/4/6.

In order to identify ongoing changes in T2/4/6, m_{rising} was used as an index to test whether young people are undergoing contour loss in T2/4/6 as they are in T2/6 of the Wúxī dialect. Pitch height of onset (*ST1*), ST_{Pmin} and the normalized time of P_{min} (NT_{Pmin}) were used as indexes to test whether young people are changing the rising tone /13/ to a delayed rising tone /113/ or even concave tone /213/.

Table 6.5 Identifying ongoing changes in SHUD T2/4/6 with the comparison of models with and without AGE (*n*=1158)

dependent variable		Deviance	Chisq	Chi Df	Pr(>Chisq)
SHUD m_{rising}	model-0	18148			
	model-AGE	18148	0.024	1	0.878
SHUD ST_1	model-0	7864.7			
	model-AGE	7860.4	4.278	1	0.039^{*}
SHUD ST_{Pmin}	model-0	7902.7			
	model-AGE	7899.9	2.777	1	0.096
SHUD NT_{Pmin}	model-0	7504.7			
	model-AGE	7492.4	12.265	1	0.000^{***}

Table 6.5 shows that m_{rising} is not conditioned by AGE. Shànghǎi young people keep the same steepness of rising T2/4/6 as old people and thus do not have contour loss. From Table 6.5, it can also be seen that no age difference is found in ST_{Pmin}. ST_1 is conditioned by AGE, and young people have higher onset (M=−2.21 ST) than old people (M=−3.22 ST). The pitch differences are around 1 ST, which is not great enough to change the transcription of onset from /1/ to /2/. NT_{Pmin} is conditioned by AGE, as young people have later P_{min} (M=3.06) than old people (M=2.18). Old people start the rising trend at around 22% of the entire duration while young people delay to 30%. Young people are changing a direct rising tone /13/ into a delayed rising tone. This is the first ongoing change identified in SHUD, and is called

SHUD_Delay/13/. This variable is further examined in Section 6.2.

6.1.3. Bǎoshān dialect (BSD)

Bǎoshān T4 and T6 have been merged into a rising tone since the 1920s, according to Chao (1928). Qián (1992) confirmed this merger and transcribed them as /213/ in the old variety and /113/ in the new variety. Bǎoshān T2 was always a peaking tone and kept distinct from T4 and T6 in the previous studies (Chao, 1928; Qián, 1992). The transcriptions of my dataset confirm Qián's (1992) findings but also indicate some new findings. First, F_0 contours of Bǎoshān T2, T4 and T6 morphemes are illustrated in Figure 6.2, split by tone shapes (peaking and low rising), AGE and MC_AB.

The legend of Figure 6.2 shows that the majority of Bǎoshān T4 and T6 morphemes are rising tones and those rising contours have merged into /113/ among old and young people. Then these two tones are called T4/6. In T2 and T4/6, obvious differences of tonal shapes can be distinguished between old and young people. Whatever the tonal shape, young people are undergoing contour loss because their peaking tones have a lower degree of convexity than those of old people and their rising tones show a less steep slope.

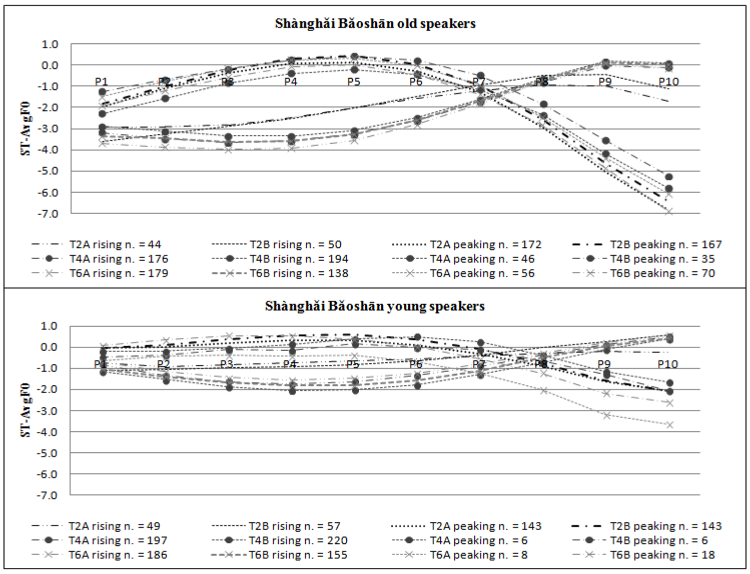


Figure 6.2 Mean F_0 contours of Bǎoshān T2, T4 and T6, split by tone shapes (peaking and low rising), AGE and MC_AB

Since the delayed rising contours are found in both SHUD and NHD, BSD is also investigated to see whether the P_{min} of its rising T4/6 is later in young people. As before, model-0 and model-AGE were built for NT_{Pmin}. The results show that young people have significantly larger NT_{Pmin} than old people (4.49 vs 3.34), $\chi^2(1)=10.975$, $p=0.000^{***}$. This suggests that though T4/6 is already a delay rising tone /113/ in old people, young people are moving P_{min} further backward like SHUD and NHD. This variation is called *BSD_Delay/113/*.

Considering the tonal shapes, though the majority of T4/6 morphemes are rising tones (85.5%), some peaking contours are also found in T4/6: 207 tokens pronounced by old people and 38 tokens pronounced by young people. Meanwhile, the majority of Bǎoshān T2 morphemes (75.8%) are a peaking tone /131/, but rising contours of T2 also occur in our data: 94 tokens pronounced by old people and 106 tokens pronounced by young people. According to frequency, young people do not seem to use more rising contours in T2 than old people, but do use less peaking tone in T4/6.

In order to interpret these splits, mixed models were built to see whether any internal or external factors can predict the splits. The dependent variables are the tone shapes of T2 (peaking vs rising tone) and tone shapes of T4/6 (peaking vs rising tone). For each dependent variable, a series of binary logistic regression models with mixed effects were built, model-0 with only the random factor and model-X with both the random factor and one fixed factor in our design, which include AGE, SEX, EDU, PTH, STYLE, MORPHEME, MC_AB, HEIGHT, DUR and PTH_TONE.

Table 6.6 Results of likelihood ratio tests on the split of peaking tones (*n*=200) and rising tones (*n*=2757) for T2 of BSD

	deviance	**Chisq**	**Chi Df**	**Pr(>Chisq)**
model-0	626.34			
model-AGE	626.33	0.013	1	0.910
model-SEX	625.30	1.039	1	0.308
model-EDU	623.78	2.558	2	0.278
model-PTH	625.04	1.300	1	0.254
model-STYLE	574.66	51.675	2	0.000^{***}
model-MORPHEME	626.14	0.200	1	0.654
model-MC_AB	625.46	0.878	1	0.349
model-HEIGHT	620.27	6.069	5	0.300
model-DUR	567.34	58.996	1	0.000^{***}

Table 6.6 shows that only two factors - STYLE and DUR - are significant for T2. The factor STYLE shows that there are significantly more rising tones in the reading of minimal pairs than in the reading of word lists (89 vs 39). The unexpected rising contours in T2 mostly occur in the minimal pairs where speakers have trouble

distinguishing the tonal difference of T2 and T4/6 morphemes, and borrow the rising tone of T4/6 morphemes to pronounce T2 morphemes that are in the same pair. The factor DUR suggests that rising tone /113/ has a longer duration than peaking tone /131/. This was not addressed in previous studies concerning intrinsic duration differences of different tonal shapes.

Table 6.7 Results of likelihood ratio tests on the split on the split of peaking tones (*n*=245) and rising tones (*n*=1445) for T4/6 of BSD

	deviance	**Chisq**	**Chi Df**	**Pr(>Chisq)**
model-0	1113			
model-AGE	1087.9	25.024	1	0.000***
model-SEX	1112.5	0.443	1	0.506
model-EDU	1108.3	4.651	2	0.098
model-PTH	1110.7	2.258	1	0.133
model-STYLE	1107.2	5.778	2	0.056
model-MORPHEME	1101.4	11.584	1	0.000***
model-MC_TONE	1110.6	2.325	1	0.127
model-MC_AB	1112.8	0.132	1	0.717
model-HEIGHT	1102.7	10.268	5	0.068
model-DUR	1048.8	64.180	1	0.000***
model-PTH_TONE	1108.7	4.244	2	0.120

In terms of the convex contours identified in T4/6, three factors are significant. First, DUR indicates again that rising tone /113/ has a longer duration than peaking tone /131/, which might indicate the intrinsic durational difference between rising and peaking tones. AGE shows that old people use more peaking contours in T4/6 than young people (207 vs 38, see also Figure 6.2). Interestingly, old people show more variances but young people are more consistent in using rising contours in T4/6. MORPHEME suggests that bound morphemes are more likely to be realized as peaking contours. STYLE is not significant here. Old people tend to use peaking tones for bound morphemes, which account for 55% of the total peaking tokens observed in T4/6. Bound morphemes are rarely used independently by speakers apart from on some special occasions. When old people were asked to pronounce those bound morphemes, they tended to use a peaking tone, i.e., the tone not used in the urban area. But young people tended to use a rising tone /113/, which is also used in the urban area and other suburbs and therefore is less characteristic of Bǎoshān.

To sum up this section, T4 and T6 of BSD have merged into a rising tone /113/ and T2 is a peaking tone /131/. Rising T2 and peaking T4/6 are also found in my dataset but they are not ongoing changes. The rising T2 is mostly realized in minimal pairs, which is caused by borrowing the rising contours of T4/6 in the same pair. The peaking T4/6 is mostly used by old speakers for bound morphemes. Two ongoing changes are identified in BSD as they are characteristic of young speakers: contour loss and delayed P_{min} of rising contours. Their conditioning effects are examined in Section 6.2.

6.1.4. Sōngjiāng dialect (SJD)

Sōngjiāng T4 started to merge with T6 in the 1920s, according to Chao (1928). At that time, T4A merged into T6 and was realized as /62/, while T4B was realized as /^{7b}67/. In the 1980s, Qián (1992) found that the merger was complete in both old and young people and transcribed it as /113/. Sōngjiāng T2 was a rising tone ending with a falling trend, which was transcribed as /13^{b7}/ in the 1920s (Chao 1928). It developed into a peaking tone /231/ in the 1980s, both in the old and young generations (Qián 1992). The transcriptions of my dataset confirm Qián's (1992) findings but also indicate some new findings. F_0 contours of T2, T4 and T6 morphemes are illustrated in Figure 6.3, split by tone shapes (peaking and low rising), AGE and MC_AB.

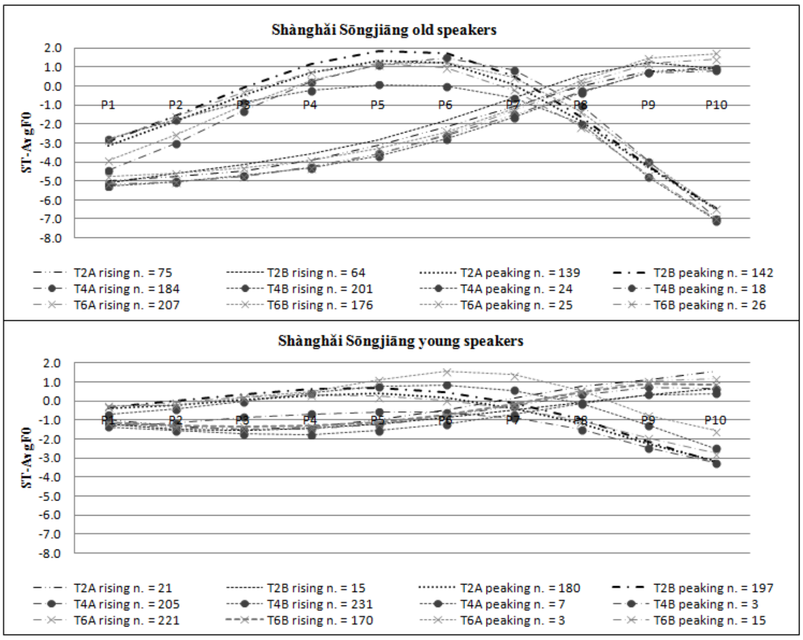


Figure 6.3 Mean F_0 contours of Sōngjiāng T2, T4 and T6, split by tone shapes (peaking and low rising), AGE and MC_AB

As can be seen in Figure 6.3, Sōngjiāng young people are undergoing contour loss in both rising and peaking tones. This is the first ongoing change identified in Sōngjiāng.

Looking at the rising contours, the majority of Sōngjiāng T4A, T4B, T6A and T6B morphemes have rising contours. Those rising contours have merged and are henceforth called T4/6. Old people have direct rising contours for T4/6, /13/, while young people have delayed rising contours, /113/. Based on the delayed process found in SHUD, NHD and BSD, SJD is also investigated to see whether the P_{min} of its rising T4/6 is later in young people. Model-0 and model-AGE were built for NT_{Pmin} (n=1595). The results show that young people have significantly larger NT_{Pmin} (M=3.65) than old people (M=2.03), $\chi^2(1)=14.22$, $p=0.000^{***}$. The process of delaying P_{min} in rising contours is also found in SJD and it is called *SJD_Delay/13/*. This process is found in all four Shànghǎi dialects and is analyzed in Section 6.2.

Although the majority of Sōngjiāng T4/6 morphemes have a rising tone, some peaking contours are found in T4/6: 93 tokens used by old people and 28 tokens used by young people. Meanwhile, the majority of Bǎoshān T2 morphemes (79.0%) are realized as a peaking tone /131/, but rising contours of T2 also occur in our data: 139 tokens used by old people and 36 tokens used by young people. Young people are more consistent in using peaking tone in T2 and rising tone in T4/6. In order to interpret these splits, mixed models were built to see whether any social or linguistic factors can predict the splits. The dependent variables are tone shapes of T2 (peaking vs rising tone) and tone shapes of T4/6 (peaking vs rising tone). For each dependent variable, a series of binary logistic regression models with mixed effects was built, model-0 with only the random factor and model-X with both the random factor and one fixed factor in our design, which include AGE, SEX, EDU, PTH, STYLE, MORPHEME, MC_AB, HEIGHT, DUR and PTH_TONE.

Table 6.8 shows that three factors - AGE, STYLE and DUR - are significant. All Sōngjiāng speakers reported that they can speak PTH, so model-PTH cannot be built and is indicated as "NA" in Table 6.12. The significant effect of AGE shows that young people use more peaking tones in T2. Regarding STYLE, more rising tones are found in the reading of minimal pairs, which means the unexpected rising contours in T2 mostly show up in the minimal pairs. Speakers cannot completely distinguish the tonal difference of T2 and T4/6 morphemes, and borrow the rising tone of T4/6 morphemes to pronounce T2 morphemes in the same pair.

Table 6.8 Results of likelihood ratio tests on the split of peaking tones (*n*=658) and rising tones (*n*=175) for T2 of SJD

	deviance	Chisq	Chi Df	Pr(>Chisq)
model-0	687.99			
model-AGE	674.14	13.854	1	0.000***
model-SEX	687.96	0.030	1	0.862
model-EDU	684.77	3.223	2	0.200
model-PTH	NA			
model-STYLE	664.21	23.785	2	0.000***
model-MORPHEME	687.89	0.103	1	0.748
model-MC_AB	687.99	0.005	1	0.945
model-HEIGHT	681.77	6.222	5	0.285
model-DUR	597.61	90.389	1	0.000***

The interaction of AGE and STYLE was investigated by means of a likelihood ratio test, and turned out to be not significant (*p*=0.173). The factor DUR, like in BSD, suggests that rising tone /113/ has a longer duration than peaking tone /131/.

Table 6.9 Results of likelihood ratio tests on the split of peaking tones (*n*=121) and rising tones (*n*=1595) for T4/6 of SJD

	deviance	Chisq	Chi Df	Pr(>Chisq)
model-0	759.90			
model-AGE	749.84	10.064	1	0.002**
model-SEX	759.83	0.066	1	0.798
model-EDU	752.67	7.235	2	0.027*
model-PTH	NA			
model-STYLE	754.24	5.664	2	0.059
model-MORPHEME	759.82	0.085	1	0.770
model-MC_TONE	759.38	0.518	1	0.472
model-MC_AB	759.87	0.033	1	0.855
model-HEIGHT	756.39	3.506	5	0.623
model-DUR	688.75	71.153	1	0.000***
model-PTH_TONE	756.17	4.728	2	0.094

Regarding peaking T4/6, three factors are significant: AGE, EDU and DUR. AGE shows that old people use fewer rising contours /113/ - the expected tone shapes in T4/6 - than young people, but use more convex contours /131/. EDU is correlated with AGE. People who received higher education are mostly young people. DUR here also indicates that rising tone /113/ has a longer duration than peaking tone /131/. In addition, the main effect of STYLE is almost significant (*p*=0.059). Model-STYLE suggests that more unexpected peaking tones occur in the reading of minimal pairs than in the reading of word lists (*p*=0.046). This means that speakers may also have difficulties in distinguishing the tonal differences between T2 and T4/6, and so borrow the peaking tone of T2 morphemes to pronounce T4/6 morphemes in the same pair.

6.1.5. Nánhuì dialect (NHD)

In the 1920s, Chao (1928) found that T2, T4 and T6 of NHD were three independent rising tones, starting at different pitch heights and rising in different slopes. He used the precursor of the five-point scale to notate tone shapes and transcribed T2 as /14/, T4 as /$^{3b}\underline{23}$/ and T6 as /$\widehat{\underline{3}^{b}5}$/. In the 1980s, Qián (1992) found T2, T4 and T6 were still kept separate by old speakers as T2 was /113/, T4 was /323/ and T6 was /35/, though in the young generation those three tones had been merged into /113/. However, as can be seen from Table 6.1, no /35/ or /323/ was found in our data. Besides a few high falling and mid level tones, the vast majority (over 95%) of tonal tokens were transcribed as rising tones, /13/ or /113/. /13/ is used mostly by old speakers and thus is the most conservative form in our data. So, our data cannot provide a real time window into the changing process from /35/ or /323/ to /13/, but can only track the changing process from /13/ to /113/. First, the contingency table of transcribed /13/ and /113/, split by AGE, MC_TONE and MC_AB, is provided in Table 6.10 to give an overview of their distributions.

Table 6.10 Contingency table of /13/ and /113/ in NHD, split by AGE, MC_TONE and MC_AB

MC_TONE* MC_AB	/13/ old n=562	 young n=188	/113/ old n=728	 young n=1022	**TOTAL** n=2500
T2A	42	26	173	167	408
T2B	43	32	171	172	418
T4A	122	35	89	170	416
T4B	66	3	158	210	437
T6A	163	45	62	162	432
T6B	126	47	75	141	389

As can be seen in Table 6.10, young people have a considerable reduction of using /13/ in comparison with old people. Within /13/ morphemes, T4A, T6A and T6B morphemes are more likely to be realized as /13/ than T2A, T2B and T4B morphemes, among old and young alike. However, during transcription, the author found it was difficult to separate /13/ and /113/ clearly because the only difference between them is whether the rising tendency is direct or delayed. Continuous variables and acoustic analysis are needed here to investigate the distribution and reduction of /13/. In order to seek proper continuous variables, F_0 contours of T2A, T2B, T4A, T4B, T6A and T6B morphemes are illustrated in Figure 6.4, split by AGE:

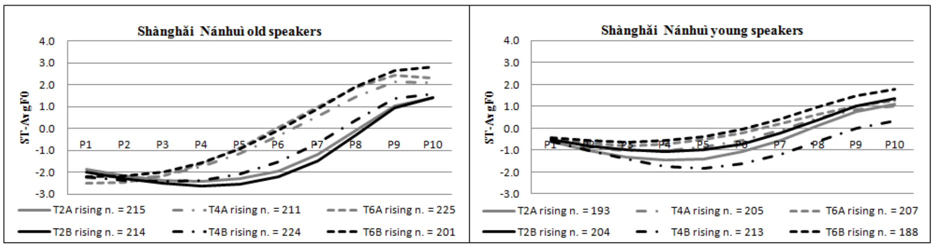


Figure 6.4 Mean F_0 contours of rising tones in NHD, split by AGE, MC_TONE and MC_AB

Figure 6.4 illustrates three interesting points. First, among old speakers, the mean contours of T4A, T6A and T6B have a direct rising tendency while the mean contours of T2A, T2B and T4B have a delayed rising tendency, in which the differences can be quantified by the position of P_{min}, i.e., NT_{Pmin} and the slope between P_{min} and P_{10}, i.e., m_{rising},. Second, among young speakers, the mean contours of T4B, T6A and T6B are getting closer to those of T2A and T2B. Third, T2A, T2B and T4B contours of young speakers seem to be more flattened and more concave-shaped than those of old speakers, especially for the contour of T4B.

The first and second observations together indicate the merger of T4A, T6A and T6B (/13/) into T4B, T2A and T2B (/113/). As discussed in Section 2.5.2.1, Qián (1992: 43-44) found that T6 in the old variety was realized as /35/ but became /113/ in the new variety, just as with T2. Based on tone values, T6 merged into T2 as T6 took over T2's tone values. But Qián (1992: 44) thought it was T2 and T4 merged into T6, which contradicted his own observations.

Qián's observation was done in NHZ. Now I use my own data, which is from NHH, to trace what was going on in T2, T4 and T6 of the Nánhuì dialect. The left graph in Figure 6.2 shows that T4A has merged with T6 while T4B has merged with T2 for old people, i.e., T4A>T6, T4B>T2. Chén (2007) also observed this and called it the "bilateral diffusion of T4". Chen (2007) asserted that "T4A>T6" belongs to "Zhuó Shǎng Guī Qù", the major phonological change found in many Chinese dialects, while "T4B>T2" is a local change. Our data confirmed Chen's (2007) observation, and furthermore indicated the change after the merger of "T4A>T6" and "T4B>T2". As seen in the young people's data in Figure 6.5, after the "bilateral diffusion of T4", T6 merged with T2 but not vice versa. Finally, T2, T4 and T6 were merged and the process of merger is sketched in Figure 6.5:

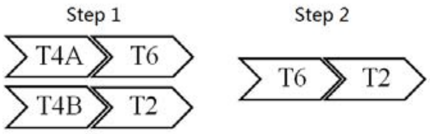


Figure 6.5 The merging process of T2, T4 and T6 in NHD

To test whether old speakers mainly use /13/ for T4A, T6A and T6B, while use /113/ for T4B, T2A and T2B, m_{rising} and NT_{Pmin} were investigated to see if they were conditioned by MC_TONE*MC_AB in the data of old speakers. Though /13/ and /113/ are both mid-rising tones, /13/ is observed to have a steeper rising slope than /113/ in Figure 6.5. Thus it is expected that, first, m_{rising} is significantly higher in T4A, T6A and T6B. T4B, T2A and T2B morphemes have delayed rising tendency, and second, their P_{min}s occur later than direct rising contours, so they are expected to have higher NT_{Pmin}.

Mixed models were built using the data of old people to test these two hypotheses. The dependent variables are m_{rising} and NT_{Pmin}. For each of them, two models were built, model-0 having only random factors and model-TONE adding the interaction of MC_TONE and MC_AB as a fixed factor on the basis of model-0. The results are given in Table 6.11:

Table 6.11 Split of /13/ (n =562) and /113/ (n=728) in NHD T2, T4 and T6 among old speakers

dependent variable	Model	predictor	β	SE	t	p
m_{rising}	**model-0**	Intercept	23.812	1.763	13.510	0.000***
	model-TONE	Intercept	21.262	1.946	10.928	0.000***
		TONE (baseline: *T2A*)				
		T2B	−0.591	1.337	−0.442	0.660
		T4A	4.549	1.327	3.429	0.000***
		T4B	0.698	1.329	0.525	0.601
		T6A	5.745	1.311	4.383	0.000***
		T6B	4.427	1.348	3.283	0.002**
		$\chi^2(5)=36.577$, $p=0.000^{***}$				
NT_{Pmin}	**model-0**	Intercept	3.027	0.297	10.200	0.000***
	model-TONE	Intercept	3.906	0.334	11.709	0.000***
		TONE (baseline: *T2A*)				
		T2B	0.041	0.266	0.153	0.879
		T4A	−1.495	0.260	−5.737	0.000***
		T4B	−0.520	0.263	−1.973	0.052
		T6A	−1.742	0.259	−6.722	0.000***
		T6B	−1.433	0.267	−5.372	0.000***
		$\chi^2(5)=66.701$, $p=0.000^{***}$				

The results of mixed models built on m_{rising} and NT_{Pmin} confirm the above two hypotheses. Table 6.11 shows that T4A, T6A and T6B have significantly higher m_{rising} (positive βs) and lower NT_{Pmin} (negative βs) than T2A morphemes while T2B and T4B morphemes do not differ from T2A for these variables. These results indicate that there are two tones in the old variety of NHD: /13/ and /113/. T4A, T6A and T6B mostly use /13/, while T2A, T2B and T4B mostly use /113/. Based on this statistical test, T2A, T2B and T4B are categorized as the /113/ group while T4A, T6A and T6B are categorized as the /13/ group.

Confirming the split of /13/ and /113/ is the basis for further investigations into ongoing changes, i.e. the reduction of /13/ group and the contour loss of /113/ group. To identify whether /13/ group is involved in any ongoing changes, NT_{Pmin} and m_{rising} were further analyzed by mixed models to see if and how they are conditioned by AGE and the interaction between AGE and tone category (/13/ vs /113/). Table 6.12 lists the outputs of mixed models for these two variables.

Table 6.12 Ongoing merger of 13/ (n =750) and /113/ (n=1750) in NHD T4A, T6A and T6B

	dependent variable: NT_{Pmin}				**dependent variable: m_{rising}**			
	β	**SE**	**t**	**p**	**β**	**SE**	**t**	**p**
Intercept	2.360	0.233	10.135	0.000^{***}	26.170	1.675	15.625	0.000^{***}
AGE (*old*): young	0.747	0.319	2.337	0.030^{*}	−9.727	2.340	−4.157	0.000^{***}
TONE (*/13/*): */113/*	1.416	0.116	12.206	0.000^{***}	−4.916	0.681	−7.222	0.000^{***}
AGE*TONE *young*/113/*	−0.454	0.116	−3.920	0.000^{***}	6.285	0.799	7.863	0.000^{***}

The output of NT_{Pmin} shows that two main effects and their interaction are significant. First, significant TONE indicates that tone /113/ has larger NT_{Pmin} than tone /13/ in the old people, confirming again the split of /13/ and /113/ in NHD. Second, the significant AGE shows that young people have larger NT_{Pmin} than old people for tone /13/. Young people's tone /13/ is moving its P_{min} backward. Moreover, the significant effect of *young*/113/* shows that the NT_{Pmin} difference between young people and old people is significantly different in TONE=/13/ vs TONE=/113/. Considering the value of β, young people have larger NT_{Pmin} (β=0.747) than old people for /13/ but much less so (β=0.747−0.454) in tone /113/, indicating the NT_{Pmin} difference between /13/ and /113/ is shrinking in young people. Young people have late P_{min} in /13/ while they are reducing the NT_{Pmin} differences between /13/ and /113/. Together this confirms that young people are migrating from /13/ towards /113/.

Similarly, the output of m_{rising} shows that both two main effects and their interaction are significant. First, significant TONE indicates that tone /113/ has smaller m_{rising} than tone /13/ in the old people, confirming again the split of /13/ and /113/ in NHD. Second, the significant AGE shows that young people have smaller m_{rising} than old people for tone /13/, indicating that young people are reducing the slope of /13/ to get closer to /113/. Third, the significant effect of *young*/113/* shows that the m_{rising} difference between young people and old people is significantly different when TONE=/13/ vs TONE=/113/. Considering the value of β, young people have smaller m_{rising} (β=−9.727) than old people for /13/ but much less so (β=−9.727+6.285) in tone /113/, indicating the slope difference between /13/ and /113/ is shrinking in young people. Young people are dropping the steepness of /13/ as well as reducing the slope differences between /13/ and /113/, which together confirm that they are migrating /13/ towards /113/. The delayed NT_{Pmin} and small m_{rising} together indicate that young people are merging /13/ and /113/ in NHD.

Turning to the tone /113/, it has already been mentioned above that young people have larger NT_{Pmin} and smaller m_{rising} for /113/ than old people but the differences are not huge. NT_{Pmin} and m_{rising} were examined by likelihood ratio tests to see if they are conditioned by AGE significantly in the dataset of /113/. Meanwhile, as can be observed in Figure 6.3, the contour of T4B is separating from T2A and T2B contours in the data of young people. T2 and T4B contours are analyzed separately. The results for T2 show that the age differences are neither significant for m_{rising} ($\chi^2(1)=1.875$, $p=0.171$) nor NT_{Pmin} ($\chi^2(1)=2.156$, $p=0.142$). These results suggest no significant contour loss exists in T2. The observed "flattened" T2 contour of young people in Figure 6.4 is mainly caused by its high onset. Why is contour loss found in SHUD, BSD and SJD but not in NHD? The temporary explanation is NHD is undergoing another ongoing change: the merger of /13/ and /113/. During this process, its /113/ remains relatively stable to wait for the accomplishment of /13/ migration. It seems that the affordable variation for one tonal inventory is limited.

Regarding the contour of T4B, young people show a clear trend of using dipping T4B with small m_{rising} and large NT_{Pmin} among six tonal contours, which make T4B distinguishable from the other rising contours. Thus it is hypothesized that T4B variation is caused by the borrowing of PT3 (/214/). The borrowing of PT3 is also found in other Shànghǎi dialects.

To sum up Section 6.1.5, young people in NHD are found to replace /13/ with /113/. Since this replacement is also realized by delaying P_{min}, it is called

NHD_Delay/13/. NHD_Delay/13/ not only causes a delayed rising tone but also simplifies the tonal inventory. It is analyzed in Section 6.2.4.

6.2. Analyzing the linguistic variables of citation tone

6.2.1. Delayed rising tone

Section 6.1 identified four ongoing changes in four regions: *SHUD_Delay/13/, BSD_Delay/113/, SJD_Delay/13/ and NHD_Delay/13/.* These are the processes of young people moving P_{min} towards the end of a rising contour either from the direct rising tone /13/ used by old people in SHUD, NHD and SJD or the delayed rising tone /113/ used by old people in BSD. Those processes are collectively referred to as *DelayRising.*

DelayRising has never previously been documented in the Shànghǎi dialects, so the first task is to identify whether it is an internal or a contact-driven change. To test if *DelayRising* is motivated by the contact with PTH, speakers who state they can speak PTH are compared with those who cannot. If two groups of people are not different in NT_{Pmin}, the variation of NT_{Pmin} is more likely to be internally motivated. As all young SHD speakers report being able to speak PTH, our analysis is fouced on the group of old speakers. 7 of the 40 old Shànghǎi speakers do not speak PTH: 2 from the urban area, 2 from Bǎoshān and 3 from Nánhuì. In Sōngjiāng, everybody claims to speak PTH. If old people who can speak PTH have larger NT_{Pmin} than those who cannot, it suggests the NT_{Pmin} is externally motivated. If there is no difference for NT_{Pmin} between those speaking PTH and those not, then *DelayRising* is probably internally motivated.

The statistical method used here is once again the likelihood ratio test. First, two models were compared, model-0 with only random effects of SPEAKER and WORD and model-PTH with the main effect PTH and two random effects. There was no significant effect of PTH on NT_{Pmin}, $\chi^2(1)=0.421$, $p=0.516$. Then interactions of PTH with all the other main effects were checked one by one, again by comparing two models: model-PTH+X (X denotes a fixed factor in our design) and model-PTH*X, which not only includes the main effects but also the interaction of PTH with another fixed factor. The results are listed in Table 6.13.

As seen from Table 6.13, no interaction of PTH with a fixed factor in our design is significant. The absence of a PTH effect implies that old people in Shànghǎi who can speak PTH do not differ from those who cannot speak PTH for

$NT_{Pmin,}$ i.e., the position of P_{min} of rising contours /13/ or /113/.

Table 6.13 Results of likelihood ratio test on NT_{Pmin} (*n.* = 7884)

model-PTH+X vs model-PTH*X	likelihood ratio test
X=REGION	$\chi^2(2)=0.614$, p=0.735
X=SEX	$\chi^2(1)=0.762$, $p=0.383$
X=EDU	$\chi^2(2)=0.887$, $p=0.642$
X=STYLE	$\chi^2(2)=4.593$, $p=0.101$
X=MORPHEME	$\chi^2(1)=0.012$, $p=0.913$
X=MC_TONE	$\chi^2(2)=4.774$, $p=0.091$
X=MC_AB	$\chi^2(1)=0.170$, $p=0.680$
X=PTH_TONE	$\chi^2(2)=0.552$, $p=0.759$
X=HEIGHT	$\chi^2(5)=7.682$, $p=0.175$
X=DUR	$\chi^2(1)=1.808$, $p=0.179$

This suggests that those delaying processes identified in four regions are probably internally motivated. One more piece of evidence for this is that the results of merging T2, T4 and T6 in Wúxī and Shànghǎi are always /113/ or /213/, a delayed low rising tone. The delayed low rising tone seems to have an intrinsic advantage favored by lax tones in the lower register. This is probably caused by articulatory-based constraints. What is the internal motivation for the changes of /13/ into /113/? Speakers have a tendency to minimize articulatory effort. In *SHUD_Delay/13/, BSD_Delay/113/, SJD_Delay/13/ and NHD_Delay/13/*, the advanced variant with delayed P_{min} always has a longer duration than the conservative variant with earlier P_{min}. To reach the same pitch height, longer duration can reduce the load of pitch movement per time unit and hence reduce the articulatory effort.

Next, mixed models were built for NT_{Pmin} in each region to look for conditioning effects. It was decided to not analyze *DelayRising* of four regions in one model because their developments are imbalanced in MC_TONE. As discussed above, *SHUD_Delay/13/* and *SJD_Delay/13/* are in a relatively early stage as old people in these places still use /13/. *NHD_Delay/13/* is at an intermediate stage as old people there use /13/ for T6 but /113/ for T2 and T4. *BSD_Delay/113/* is the most advanced because no /13/ was found in its dataset. In order to build one model for all regions, it is necessary to include the interaction MC_TONE*REGION. However, *SHUD_Delay/13/* and *NHD_Delay/13/* have T2 rising contours in their dataset while *BSD_Delay/113/* and *SJD-Delay/13/* do not. As T2 in BSD and SJD are mostly convex tones, the T2*BSD and T2*SJD cells are empty, rendering the investigation of MC_TONE*REGION impossible. Four models were built separately for the four dialects, in Sections 6.2.1.1 to 6.2.1.4 respectively. The same method of model building was applied. First, the likelihood ratio test was used to select significant main effects and interactions. Then a final model was built on those selected fixed factors.

As discussed in Section 6.1.5, T4B morphemes in NHD are quite special because they are probably involved in borrowing PT3 in the young generation. The borrowing of PT3 is also found in the other three dialects. To investigate if T4B morphemes are involved in borrowings, a new variable TONE with 6 conditions - T2A, T2B, T4A, T4B, T6A and T6B - was used as a replacement for MC_TONE and MC_AB, because the factor TONE equals MC_TONE*MC_AB.

6.2.1.1. SHUD

Table 6.14 lists the results of likelihood ratio test on NT_{Pmin} in SHUD.

Table 6.14 Results of likelihood ratio test on NT_{Pmin} in SHUD (*n.* =2426)

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	7504.7			
model-AGE	7492.4	12.265	1	0.000^{***}
model-SEX	7504.7	0.031	1	0.859
model-EDU	7502.6	2.150	1	0.143
model-STYLE	7497.7	6.990	2	0.030^{*}
model-TONE	7492.9	11.807	5	0.038^{*}
model-MORPHEME	7504.7	0.004	1	0.951
model-DUR	7500.4	4.319	1	0.038^{*}

AGE, STYLE, TONE and DUR are the significant main effects selected by likelihood ratio tests. Further analyses of interactions show that AGE*TONE is significant. The results of the final model, built on those four main effects and one interaction, are listed in Table 6.15.

As expected, young people use a larger NT_{Pmin} in SHUD than old people. There are also two new findings. First, T2A morphemes in paragraph reading have larger NT_{Pmin} than in the reading of word lists among young people. This indicates that STYLE is responsible for the delayed rising process in SHUD and the delayed rising tone is more likely to occur in the less formal style. Second, young people pronounce T4B morphemes with larger NT_{Pmin} than T2A morphemes in the reading of word lists. At the same time, morphemes of other tonal categories do not differ from T2A. In addition, the significant effect of *old*T4B* shows that the NT_{Pmin} difference between young people and old people significantly differ for TONE=T2A vs TONE=T4B. Old people's T4B morphemes do not differ from T2A morphemes too much (0.377−0.413=0.036) in terms of NT_{Pmin}. The main effect of TONE and its interaction with AGE together indicate that young people pronounce T4B with especially delayed NT_{Pmin}. This suggests T4B morphemes are probably involved in the borrowing of PT3, a dipping tone.

Table 6.15 Output of mixed model on NT_{Pmin} in SHUD (*n.* =2426)

Predictor	***β***	**SE**	**t**	**p**
Intercept	2.729	0.224	12.209	0.000***
AGE (*young*): *old*	−0.781	0.227	−3.437	0.002**
STYLE *(wordlist): paragraph*	0.116	0.058	2.013	0.044*
minimal pair	0.028	0.119	0.238	0.813
TONE (*T2A*): *T2B*	−0.238	0.152	−1.565	0.120
T4A	0.109	0.150	0.728	0.468
T4B	0.377	0.150	2.511	0.013*
T6A	0.108	0.150	0.717	0.475
T6B	−0.043	0.154	−0.280	0.780
DUR	0.001	0.001	1.532	0.126
AGE*TONE				
*old*T2B*	0.084	0.154	0.542	0.588
*old*T4A*	−0.013	0.156	−0.083	0.934
*old*T4B*	−0.413	0.152	−2.725	0.006**
*old*T6A*	−0.052	0.154	−0.337	0.736
*old*T6B*	0.131	0.158	0.831	0.406

Table 6.16 Mean NT_{Pmin} in SHUD, split by AGE and TONE

TONE \ AGE	***old***	***young***
T2A	2.19	3.02
T2B	2.02	2.76
T4A	2.26	3.12
T4B	2.14	3.40
T6A	2.24	3.12
T6B	2.24	2.93

The final model of SHUD suggests two things: (1) the delayed rising in SHUD is associated with the variation of speaking style (word list vs paragraph). Using delayed rising tone in paragraph reading suggests *DelayRising* is motivated by saving articulatory effort, which further suggests that this process is internally motivated; (2) young people not only tend to use delayed rising tones but also tend to use extremely large NT_{Pmin} for T4B. This is probably caused by the borrowing of PT3 because T4B is the only MC tonal category in the lower register that corresponds to PT3, a dipping tone with large NT_{Pmin}. These two findings are also found in the SJD and BSD. The analyses are presented below.

6.2.1.2. BSD

Table 6.17 lists the results of likelihood ratio tests on NT_{Pmin} in BSD. AGE, STYLE and DUR are the significant main effects selected by likelihood ratio tests. Further analyses on interaction found that AGE*SEX and AGE*TONE are significant. A final model built with AGE, STYLE, DUR, SEX, AGE*SEX and AGE*TONE gives the following output in Table 6.18.

Table 6.17 Results of likelihood ratio tests on NT_{Pmin} in BSD (n=1439)

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	6106			
model-AGE	6095	10.975	1	0.001***
model-SEX	6105.5	0.475	1	0.491
model-EDU	6105.2	0.788	2	0.674
model-STYLE	6053	53.057	2	0.000***
model-TONE	6102.5	3.496	3	0.321
model-MORPHEME	6106	0.040	1	0.841
model-DUR	6084.9	21.073	1	0.000***

Table 6.18 Output of mixed model on NT_{Pmin} in BSD (n=1439)

predictor	***β***	**SE**	**t**	**p**
Intercept	5.486	0.429	12.781	0.000***
AGE (*young*): *old*	0.070	0.412	0.169	0.867
STYLE *(wordlist):*				
minimal pair	0.478	0.119	4.012	0.000***
paragraph	0.695	0.230	3.017	0.003**
DUR	−0.006	0.001	−5.668	0.000***
SEX *(female)*: *male*	0.419	0.375	1.117	0.281
TONE (*T4A*):				
T4B	0.437	0.233	1.880	0.063
T6A	0.192	0.235	0.815	0.417
T6B	0.467	0.247	1.895	0.060
AGE*SEX				
*old*male*	−1.753	0.532	−3.295	0.004**
AGE*TONE				
*old*T4B*	−0.599	0.282	−2.121	0.034*
*old*T6A*	−0.043	0.273	−0.156	0.876
*old*T6B*	−0.560	0.299	−1.869	0.062

STYLE, DUR, TONE, AGE*SEX and AGE*TONE are significant factors in the final model. To understand the insignificant effect of AGE and SEX, the baseline category of Table 6.18 needs to be explained. The baseline category is the category with all main effects assigned the baseline values. To read Table 6.18, the baseline for AGE is young people, the baseline for STYLE is word list, the baseline for SEX is female, and the baseline for TONE is T4A. Then the baseline category for the whole model is NT_{Pmin} of T4A morphemes is the reading of word lists by young women. The insignificant AGE and SEX occurring in Table 6.18 means that in the reading of word lists, young women do not differ from old women for NT_{Pmin} in T4A morphemes, and furthermore do not differ from young men.

Turning to the significant effects, STYLE suggests that morphemes in paragraph readings have significantly larger NT_{Pmin} than in the reading of word lists and minimal pairs. The larger NT_{Pmin} in paragraph reading support the notion that *DelayRising* is associated with the reduction of articulatory effort. The significant DUR implies that the longer duration a morpheme has, the more likely it has a delayed P_{min}. The main effect of TONE is nearly significant in the condition of T4B (p=0.063) while the interaction *old*T4B* is also significant. It means young people tend to use larger NT_{Pmin} for T4B morphemes (positive β for T4B*)* than T4A morphemes while this gap in old people is significantly smaller (negative β for *old*T4B*). The mean scores of NT_{Pmin} split by TONE and AGE in Table 6.19 present this situation more clearly. They suggest that the T4B of young people is also influenced by PT3.

Table 6.19 Mean NT_{Pmin} in BSD, split by AGE and TONE

TONE \ AGE	***old***	***young***
T4A	3.43	4.27
T4B	2.99	4.51
T6A	3.66	4.53
T6B	3.22	4.68

Interestingly, SEX is responsible for the NT_{Pmin} variation in BSD. As seen in Table 6.20 below, old males are also the most conservative people in using the earliest P_{min}; old females lead the change in old people. The SEX difference in the young generation is as great as in old people. Young males even have larger NT_{Pmin} than young females.

Table 6.20 Mean NT_{Pmin} in BSD, split by AGE and SEX

SEX \ AGE	*old*	*young*
female	3.87	4.23
male	2.76	4.75

6.2.1.3. SJD

Table 6.21 gives the results of likelihood ratio tests of NT_{Pmin} in SJD.

Table 6.21 Results of likelihood ratio test on NT_{Pmin} in SJD (n=1595)

	deviance	Chisq	ChiDf	Pr(>Chisq)
model-0	6763.1			
model-AGE	6748.9	14.220	1	0.000***
model-SEX	6761.4	1.667	1	0.197
model-EDU	6762.2	0.861	2	0.650
model-STYLE	6748.9	14.140	2	0.001***
model-TONE	6756	7.032	3	0.071
model-MORPHEME	6763.1	0.007	1	0.932
model-DUR	6763	0.105	1	0.747

Only AGE and STYLE are significant main effects selected by the likelihood ratio tests. The significant interactions selected are AGE*TONE and AGE*SEX.

Table 6.22 Output of mixed model on NT_{Pmin} in SJD (n=1595)

Predictor	β	SE	t	p
Intercept	3.939	0.370	10.636	0.000***
AGE (*young*):*old*	0.362	0.108	3.348	0.001**
STYLE *(wordlist)*: *paragraph*	0.489	0.216	2.264	0.025*
minimal pair	−1.925	0.511	−3.769	0.001**
TONE (*T4A*):				
T4B	0.392	0.211	1.856	0.066
T6A	−0.340	0.212	−1.604	0.111
T6B	0.112	0.226	0.495	0.622
AGE*TONE				
*old*T4B*	−0.572	0.274	−2.089	0.037*
*old*T6A*	0.146	0.274	0.533	0.594
*old*T6B*	−0.152	0.289	−0.526	0.599
AGE*SEX				
young:male	−1.113	0.480	−2.320	0.034*
old:male	−0.164	0.481	−0.341	0.737

AGE and STYLE played the same role in the variation of NT_{Pmin} as other regions. Young people use larger NT_{Pmin}. Paragraph and word lists are the contexts for promoting larger NT_{Pmin} among young people. The significant effect of paragraphs again suggests the reduction of articulatory effort in *DelayRising*. Like BSD, the main effect of TONE in SJD is nearly significant in the condition of T4B (p=0.066) while the interaction *old*T4B* is also significant. Young people tend to use larger NT_{Pmin} for T4B morphemes than T4A morphemes while this gap in old people is significantly smaller. Table 6.23 presents this situation using mean scores of NT_{Pmin} split by TONE and AGE. Young people of SJD also borrow PT3 for T4B morphemes.

Table 6.23 Mean NT_{Pmin} in SJD, split by AGE and TONE

TONE \ AGE	***old***	***young***
T4A	2.14	3.59
T4B	1.95	3.99
T6A	1.98	3.29
T6B	2.11	3.73

SEX is responsible for the NT_{Pmin} variation in SJD as well. As can be seen in Table 6.24, old males seem to be the most conservative people in using the earliest P_{min} like BSD. The sex difference is not significant in the old generation but significant in the young generation. Unlike BSD, young males still show more conservativeness than young females. Old females are leading *DelayRising* in BSD but leaders of SJD are young females. It implies that *DelayRising* starts later in SJD than BSD. In Section 6.2.3, there are more discussions to compare *DelayRising* in the four regions.

Table 6.24 Mean NT_{Pmin} in SJD, split by AGE and SEX

SEX \ AGE	*Old*	*young*
female	2.13	4.22
male	1.94	3.10

6.2.1.4. NHD

As discussed in Section 6.1.5, a process of merging /13/ with /113/ for T4A and T6 morphemes is identified in NHD, which is referred to *NHD_Delay/13/.* Though merging /13/ with /113/ belongs to *DelayRising*, this process in NHD is different from that identified in the other three Shànghǎi dialects because *NHD_Delay/13/* causes the loss of a tonal category /13/, i.e. the simplification of tonal inventory in NHD. In the other three Shànghǎi dialects, *DelayRising* only plays a role within the tonal category /13/ or /113/ but does not cause any merging between categories. Given this difference, it is expected that different conditioning effects are responsible for the variation of NT_{Pmin} in NHD. Again, likelihood ratio tests were used to select significant main effects and interactions. Table 6.25 lists the results.

As shown in Table 6.25, AGE and DUR are the significant main effects selected by likelihood ratio tests. Further analyses of interactions are conducted. Since STYLE is hypothesized to play a role in *DelayRising*, though it is not significant as a main effect here, its interactions with other main effects were also tested, but none showed significant effects. The only significant interaction is AGE*SEX. A final model was then built on AGE, DUR and AGE*SEX.

Table 6.25 Results of likelihood ratio tests on NT_{Pmin} in NHD (n=2417)

	deviance	Chisq	ChiDf	Pr(>Chisq)
model-0	4181.1			
model-AGE	4174.1	6.942	1	0.008**
model-SEX	4177.6	3.465	1	0.063
model-EDU	4178.2	2.893	2	0.235
model-STYLE	4180.6	0.461	2	0.794
model-TONE	4178.2	2.890	2	0.236
model-MORPHEME	4180	1.116	1	0.291
model-DUR	4101	80.118	1	0.000***

Table 6.26 Output of mixed model on NT_{Pmin} in NHD (n=2417)

Predictor	β	SE	t	p
Intercept	0.128	0.292	0.436	0.665
AGE (*old*): *young*	1.800	0.320	5.619	0.000***
SEX *(male)*: *female*	0.996	0.311	3.200	0.006**
DUR	0.006	0.001	9.550	0.000***
AGE*SEX				
*young*female*	−1.401	0.446	−3.145	0.006**

Table 6.26 shows that there is a significant effect of DUR: the longer the duration of a morpheme, the later NT_{Pmin} shows up. There are significant effects of AGE, SEX and the interaction SEX*AGE for NT_{Pmin}. The mean scores split by AGE and SEX are presented in Table 6.27. As can be seen in the table, old women and young women do not differ markedly for NT_{Pmin}, but young men have significantly larger NT_{Pmin} than old men. So old men are the most conservative people as they use the smallest NT_{Pmin}; women lead the change in old people. The sex difference is reduced in the young generation as young men have almost the same NT_{Pmin} as women. Results of NHD and BSD together suggest that suburban old women are more active in promoting *DelayRising* while old men are more resistant to it.

Table 6.27 Mean NT_{Pmin} in NHD, split by AGE and SEX

SEX \ AGE	*old*	*young*
female	3.75	3.63
male	2.42	3.45

NHD_Delay/13/ is different from SHUD, BSD and SJD (Sections 6.2.1.1 to 6.2.1.3) in terms of the effect of STYLE. STYLE is not able to explain the variation of NT_{Pmin} in NHD, suggesting *NHD_Delay/13/* is not motivated by saving articulatory effort. Since *NHD_Delay/13/* is not caused by the contact with PTH and SHUD (discussed in the beginning of Section 6.2.1), then the possible motivation for this process is to tonal convergence, the attraction of /113/, a tone with the same pitch height and similar tonal shape with /13/.

6.2.2. Contour loss

The identification of contour loss in each region was done in Section 6.2. Besides the Shànghǎi urban area, young people of all suburban areas are reducing the slope of low rising tone in their own dialects. BSD and SJD have peaking T2 /131/. It is also found that the young generation uses a much smaller degree of convexity to realize this peaking tone. The contour loss of rising contour /13/ or /113/ and peaking contour /131/ are analyzed separately in this section. The same index quantifying the contour loss in WXT2/6 is used for the Shànghǎi dialects, i.e., m_{rising}; $m_{rising}=\frac{ST_{Pmax}-ST_{Pmin}}{NT_{Pmax}-NT_{Pmin}}$. For the peaking tone, it is difficult to use one index to quantify its degree of convexity. Based on the observations in Figures 6.3 and 6.4, the contour loss of peaking tone is reflected by reducing: (1) the slope of the rising part in the first half; and (2) the slope of the falling part in the second half. Hence these two slopes are used together for quantifying the contour loss in peaking tones. The slope of the rising part in the first half is $m_{peaking1}$; $m_{peaking1}=\frac{ST_{Pmax}-ST_{P1}}{NT_{Pmax}-NT_{P1}}$; the slope of the falling part in the second half is $m_{peaking2}$; $m_{peaking2}=\frac{ST_{P10}-ST_{Pmax}}{NT_{P10}-NT_{Pmax}}$.

6.2.2.1. Rising contours

Section 6.1 identified that contour loss of rising contours exist in T4/6 of BSD and T4/6 of SJD, but not in SHUD and NHD. Likelihood ratio tests were applied to BSD and SJD to examine what the significant predictors are for the m_{rising} variation in each dialect. Table 6.28 and Table 6.29 present the results for BSD and SJD.

Table 6.28 Results of likelihood ratio tests on m_{rising} in BSD (n=1439)

	Deviance	Chisq	ChiDf	Pr(>Chisq)
model-0	10774			
model-AGE	10769	4.860	1	0.027**
model-SEX	10774	0.542	1	0.462
model-EDU	10773	1.451	2	0.484
model-PTH	10774	0.199	1	0.655
model-STYLE	10650	123.980	2	0.000***
model-MORPHEME	10774	0.164	1	0.685
model-MC_AB	10772	1.829	3	0.609
model-HEIGHT	10765	9.387	5	0.095
model-DUR	10774	0.231	1	0.631
model-PTH_TONE	10773	1.152	2	0.562

Table 6.29 Results of likelihood ratio tests on m_{rising} in SJD (n=1595)

	Deviance	Chisq	ChiDf	Pr(>Chisq)
model-0	11504			
model-AGE	11493	11.378	1	0.000^{***}
model-SEX	11504	0.372	1	0.542
model-EDU	11501	2.832	2	0.243
model-PTH	NA			
model-STYLE	11470	34.157	2	0.000^{***}
model-MORPHEME	11503	0.830	1	0.362
model-MC_AB	11500	4.525	3	0.210
model-HEIGHT	11495	9.445	6	0.150
model-DUR	11504	0.318	1	0.573
model-PTH_TONE	11501	2.968	2	0.227

Table 6.28 and 6.29 both show that only AGE and STYLE are responsible for the reduction of m_{rising} in BSD and SJD. The significant interaction selected for those two dialects is also the same: AGE*STYLE. Mixed models were built on AGE, STYLE and AGE*STYLE. The output is presented in Table 6.30 for BSD and Table 6.31 for SJD.

Table 6.30 Output of mixed model on m_{rising} in BSD (n=1439)

Predictor	*β*	SE	t	p
Intercept	16.135	1.798	8.975	0.000^{***}
AGE (*young*):*old*	10.252	2.518	4.071	0.001^{**}
STYLE*(wordlist):*				
paragraph	−3.651	1.430	−2.553	0.011^{*}
minimal pairs	−1.092	0.765	−1.427	0.154
AGE*STYLE				
old paragraph*	−5.069	1.820	−2.785	0.005^{**}
old minimal pairs*	−8.156	1.050	−7.764	0.000^{***}

Table 6.31 Output of mixed model on m_{rising} in SJD (n=1595)

Predictor	*β*	SE	t	p
Intercept	16.276	1.879	8.662	0.000^{***}
AGE (*young*):*old*	11.280	2.628	4.293	0.000^{***}
STYLE*(wordlist):*				
paragraph	−4.307	1.407	−3.061	0.002^{**}
minimal pairs	−1.000	0.671	−1.490	0.136
AGE*STYLE				
old paragraph*	−0.169	1.799	−0.094	0.925
old minimal pairs*	−3.345	0.914	−3.662	0.000^{***}

Comparing Table 6.30 and Table 6.31, they have very similar pattern, so discussions for them are combined. To better understand the two tables, mean m_{rising} scores in BSD and SJD are presented in Table 6.32 and Table 6.33, split by AGE and STYLE. Table 6.32 shows that Bǎoshān and Sōngjiāng young people use significantly smaller slope than old people for T4/6 in word list reading. Young people also use a much smaller slope for reading T4/6 morphemes in paragraphs than in word lists. As can be seen in Table 6.32 and Table 6.33, this style difference can also be found in the old generation. The significant effect of *old*paragraph* in BSD shows that the style gap between paragraph reading and word list reading is even bigger in the old generation. This all suggests again that paragraph reading is the context for promoting contour loss. *Old*minimal pair* is significant for both BSD and SJD. Table 6.32 and Table 6.33 also show that the m_{rising} differences between minimal pair and word list reading are not huge among young people in BSD and SJD. But among old people, the m_{rising} differences are considerable. The lower m_{rising} in the reading of minimal pairs might be caused by people's uncertainty or confusions when they are confronted by two or three morphemes having exactly the same segment pronunciations.

Table 6.32 Mean m_{rising} in BSD, split by AGE and STYLE

STYLE \ AGE	***old***	***young***
word list	26.63	16.33
paragraph	17.55	12.56
minimal pair	17.18	15.36

Table 6.33 Mean m_{rising} in SJD, split by AGE and STYLE

STYLE \ AGE	***old***	***young***
word list	27.40	16.19
paragraph	23.75	12.15
minimal pair	23.10	15.20

6.2.2.2. Peaking contours

Since the rising contours of T4/6 in BSD and SJD show a similar pattern, the peaking contours of T2 in BSD and SJD are analyzed together for their contour loss. The index for quantifying contour loss in the peaking contours is the slope between P_1 and P_{max} on the contour, i.e., $m_{peaking1}$, and the slope between P_{max} and P_{10}, i.e., $m_{peaking2}$. Table 6.34 and Table 6.35 give the results of likelihood ratio tests of $m_{peaking1}$ and $m_{peaking2}$:

Table 6.34 Results of likelihood ratio tests on $m_{peaking1}$ (n=1672)

	deviance	Chisq	ChiDf	Pr(>Chisq)
model-0	9693.2			
model-AGE	9653.8	39.4	1	0.000***
model-REGION	9689.4	3.762	1	0.052
model-SEX	9693.1	0.072	1	0.789
model-EDU	9681	12.227	2	0.002**
model-PTH	9692.7	0.465	1	0.496
model-STYLE	9685	8.209	2	0.017*
model-MORPHEME	9693.2	0.003	1	0.957
model-MC_AB	9692	1.133	1	0.287
model-HEIGHT	9683	10.163	5	0.071
model-DUR	9693.1	0.046	1	0.830

Table 6.35 Results of likelihood ratio tests on $m_{peaking2}$ (n=1672)

	deviance	Chisq	ChiDf	Pr(>Chisq)
model-0	11186			
model-AGE	11142	43.2	1	0.000***
model-REGION	11183	2.703	1	0.100
model-SEX	11186	0.227	1	0.634
model-EDU	11164	21.351	2	0.000***
model-PTH	11183	2.759	1	0.097
model-STYLE	11160	26.011	2	0.000***
model-MORPHEME	11185	1.007	1	0.316
model-MC_AB	11186	0.031	1	0.861
model-HEIGHT	11175	10.738	5	0.057
model-DUR	11185	0.869	1	0.351

The same significant effects are selected for $m_{peaking1}$ and $m_{peaking2}$: AGE, EDU and STYLE. AGE and EDU are highly correlated in Bǎoshān and Sōngjiāng. In both areas, speakers who only received primary education are all from the old generation, while college-educated speakers are all from the young generation. When testing EDU in the old people, it is not significant, $\chi^2(1)$= 1.028, p=0.598. This means that the significant effect of EDU depends on the significant effect of AGE. So AGE is put in the final model rather than EDU. Moreover, AGE*REGION is selected as a significant interaction for $m_{peaking1}$ while AGE*STYLE is significant for $m_{peaking1.}$ The final models built for those two variables are shown in Table 6.36 and Table 6.37.

Table 6.36 Output of mixed model on $m_{peaking1}$ (n=1672)

Predictor	***β***	**SE**	**t**	**p**
Intercept	26.331	2.592	10.159	0.000***
AGE (*old*): *young*	−18.211	3.529	−5.161	0.000***
STYLE *(wordlist):*				
paragraph	−1.881	0.647	−2.909	0.050*
minimal pair	1.696	0.668	2.537	0.011*
AGE*REGION				
*old*Sōngjiāng*	15.065	3.526	4.272	0.000***
*young*Sōngjiāng*	3.417	3.525	0.969	0.339

Table 6.37 Output of mixed model on $m_{peaking2}$ (n=1672)

Predictor	***β***	**SE**	**t**	**p**
Intercept	−56.735	2.967	−19.124	0.000***
AGE (*old*): *young*	31.073	4.090	7.597	0.000***
STYLE *(wordlist):*				
paragraph	7.948	3.475	2.287	0.023*
minimal pair	−7.681	1.534	−5.008	0.000***
AGE*STYLE				
young paragraph*	−5.484	4.388	−1.250	0.212
young minimal pair*	5.527	2.055	2.689	0.007**

Table 6.36 and Table 6.37 show that young people have smaller $m_{peaking1}$ and larger $m_{peaking2}$ in reading morphemes of word lists, which both indicate contour loss. The contour loss of peaking tone is stronger in word list reading than in the reading of paragraphs and minimal pairs. Paragraphs and minimal pairs are contexts promoting smaller $m_{peaking1}$ and larger $m_{peaking2}$, i.e., a much flattened peaking contour.

The effect of AGE*REGION on $m_{peaking1}$ shows that the $m_{peaking1}$ of Sōngjiāng old speakers is significantly higher than that of Bǎoshān old speakers, which can be observed in Figures 6.3 and 6.4. Sōngjiāng old speakers keep a higher degree of convexity than Bǎoshān old speakers in realizing peaking tones. As shown in Table 2.3 (Chapter 2), T2 of BSD is transcribed by Chao (1928) as /$^2\underline{31}$/ while T2 of SJD is /$\underline{13}^{b7}$/, showing that those two tones are historically different in $m_{peaking}$.

6.2.3. Summary and discussion

Section 6.2 has investigated the constraints of *DelayRising* and contour loss in Shànghǎi citation tone. The main findings are as follows:

(1) *DelayRising* is not a contact-driven change because there is no difference for NT_{Pmin} between those speaking PTH and those not. Meanwhile, *DelayRising* in SHUD (*SHUD_Delay/13/*), SJD (*BSD_Delay/113/*) and BSD (*SJD_Delay/13/*) are all constrained by STYLE. The turning point of rising tone /13/ or /113/ is more delayed in paragraph reading than in word list reading. It suggests that *DelayRising* in those three regions is motivated by saving articulatory effort, so *DelayRising* in SHUD, SJD and BSD can be said to be internally motivated.

The mechanism of *DelayRising* in NHD (*NHD_Delay/13/*) is different from that in SHUD, SJD and BSD because *NHD_Delay/13/* is not constrained by STYLE. The possible motivation for *NHD_Delay/13/* is to simplify the tonal inventory. This is because *NHD_Delay/13/* refers to a process of merging /13/ of T4A and T6 morphemes with /113/ of T4B and T2 morphemes. The consequence of

NHD_Delay/13/ is not only phonetic but also phonemic, causing the loss of a tonal category /13/.

(2) Turning to other significant constraints, it has been found that SEX is a significant factor for NT_{Pmin} in BSD, SJD and NHD, but not in SHUD. In effect, SEX is a significant predictor in suburban areas but not in urban areas. SEX differences are found in the old generation in three suburban dialects. Old males are the most conservative group showing resistance to *DelayRising*. SEX is significant in the young people of SJD.

TONE and AGE*TONE are also significant in several dialects: SHUD, BSD and SJD. In those dialects, young people all tend to use higher NT_{Pmin} for T4B morphemes, suggesting young people are likely to borrow PT3 for T4B morphemes. Neither TONE nor AGE*TONE is significant for *NHD_Delay/13/* because *NHD_Delay/13/* refers to the *DelayRising* in T4A and T6 morphemes. As discussed at the end of Section 6.1.5, T4B morphemes are also involved in the borrowing of PT3. Thus all the four dialects investigated in Shànghǎi borrow PT3. However, the borrowing of PT3 is implicit borrowing because PT3 and the original tones /13/ or /113/ have very close tonal shapes, so it is difficult to separate in T4B morphemes the *DelayRising* process and the borrowing of PT3 and thus calculate the borrowing rate. Further research is needed to study how to define and analyze implicit borrowings.

(3) Contour loss in citation tone is found in BSD and SJD but not in SHUD and NHD. NHD is not involved in contour loss mainly because it is undergoing tonal merging between /13/ and /113/. SHUD has a different reason, in that its young speakers keep the same steepness of rising T2/4/6 as old people to maintain their identity as Shànghǎi urbanite. Evidence from the questionnaire data can support this argument. First, all of the ten young speakers in urban Shànghǎi thought the Shànghǎi dialect is more pleasant to listen to than PTH[34]. However, in three Shànghǎi suburbs, none of the young speakers thought their local dialect is more pleasant to listen to than PTH. Furthermore, none of the suburban young speakers identified themselves as Shanghainese in the questionnaire. All of them reported that they were either from the Nánhuì, Sōngjiāng or Bǎoshān area rather than Shànghǎi

[34] The full questionnaire is available in Appendix IV, both in Chinese and English. The relevant data used here is based on participants' answers to question "Which one do you prefer to hear between PTH and local dialect?"

city[35].

(4) The contour loss of rising and peaking tones in BSD and SJD is also constrained by STYLE. Morphemes in paragraph reading have more contour loss than in word list reading, suggesting again that contour loss is motivated by saving articulatory effort. However, considering the exception of SHUD, the contour loss of Shànghǎi citation tone is also constrained by REGION. Then it seems that the contour loss of Shànghǎi citation tone is motivated both internally (to save articulatory effort) and externally (urban vs suburban). This issue is further analyzed in Section 6.3 and 6.4 to see if the contour loss of sandhi patterns is similar.

6.3. Identifying the linguistic variables of tone sandhi

As discussed in Section 2.5.2.2, Shànghǎi tone sandhi patterns are relatively simple in the urban area but extremely complex in three suburbs: Nánhuì, Bǎoshān and Sōngjiāng. Previous descriptive results of T2.X, T4.X and T6.X for those three suburban dialects do not support each other. So the first task of this section is to correct or supplement previous transcriptions based on our own data. Unlike Wúxī (right-dominant + left-dominant, c.f. Chan & Ren, 1989) and Shànghǎi urban area (left-dominant, c.f. Zhu, 1999), there are not enough examples of such complex sandhi patterns found in suburban dialects, which makes it hard to decide which pattern is the conservative one. As can be seen in the discussions in Sections 6.3.2 to 6.3.5, our own transcriptions also provide conflicting results, making the whole picture more complex. Our study was not primarily set up to study this relationship. This study presents the data for further reference, in line with Goal I of this study. The data includes the frequency of each sandhi pattern listed in Figures 6.6 to 6.15 and more detailed contingency tables in Appendix V.

This section is structured in the same way as Section 6.1. Section 6.3.1 analyzes the sandhi pattern /55.31/ found in all the four regions, which is possibly borrowed from the Yīn register, like in Wúxī. Sections 6.3.2 to 6.3.5 first present the descriptive results of SHUD, NHD, BSD and SJD and then identify the linguistic variables in those areas. When discussing citation tone, NHD was the last one because the variation in NHD is the most complicated. The order of presenting sandhi patterns is different from that of citation tone because, as discussed in Section

[35] The relevant question is “Where are you from?”.

2.5.2.2, sandhi patterns used in NHD are simpler than in BSD and SJD. To better understand the complexity of Shànghǎi sandhi patterns, we start from the simplest one - SHUD - and then move to the next simplest one in suburban Shànghǎi, NHD, followed by the more complex patterns in BSD and SJD.

6.3.1. Sandhi pattern /55.31/

From the transcription of all the observations, the sandhi pattern /55.31/ is found occasionally in four Shànghǎi dialects. /55.31/ was recorded in SHUD (Xǔ, et al., 1981) but not in three suburban dialects. The researchers found that old people can only pronounce a few words in /55.31/ pattern but young people use both /55.31/ and /22.44/ in words such as 老师 'teacher' and 裤子 'trousers'. So, Xú et al. (1981) predicted the loss of /55.31/ in the young generation. However, it is found that, in Wúxī dialect, young people use more /55.31/, which is the borrowing of a sandhi pattern from the Yīn register. Similarly, /55.31/ is also a sandhi pattern used in the Yīn register of SHUD (T1.X), BSD (T1.X and T5.T3-6), SJD (T1.T3-8 and T5.T1-4) and NHD (T1.T3-6 and T5.T3-6) (Qián, 1992: 639-642). Hence, /55.31/ tokens in Shànghǎi dialects are hypothesized to be borrowings as well, having a different origin than the /55.31/ recorded by Xu & Tang (1981).

First, a contingency table of transcribed /55.31/ split by REGION and AGE is provided in Table 6.38.

Table 6.38 Contingency table of transcribed /55.31/ split by dialect and AGE

dialect	**old**		**young**	
	n	*p*	*n*	*p*
SHUD	3	0.2%	26	1.9%
NHD	8	0.6%	20	1.6%
BSD	27	2.1%	47	3.6%
SJD	26	1.9%	17	1.3%

As can be seen from Table 6.38, the /55.31/ pattern occurs in all four regions, though infrequently, as occurrence is between 0.2% and 3.6%. It can also be seen that young people tend to use more /55.31/ than old people in SHUD, NHD and BSD, but not in SJD. The age difference for using /55.31/ in SJD is not significant (p=0.672). As discussed in Section 6.1.4, Sōngjiāng old speakers show a great deal of variance in choosing tonal shapes for T2, T4 and T6 in citation form, even more than young people. Hence, it is not surprising that old people do not differ from young people in terms of using /55.31/.

A likelihood ratio test was used to test if the differences observed in Table

6.38 are significant. Model-0 was built with only random factors WORD (136 levels) and SPEAKER (80 levels) on the variable /55.31/ (yes vs no). Model-AGE and model-REGION were built and compared with model-0 successively.

Table 6.39 Results of likelihood ratio tests on the occurrence of /55.31/ (n=174) or not (n=10411) in SHD

	deviance	Chisq	Chi Df	Pr(>Chisq)
model-0	1247.3			
model-AGE	1239.6	7.710	1	0.005**
model-REGION	1226.9	20.378	3	0.000***

The results show that both AGE and REGION are significant. Young speakers are more likely to use /55.31/ than old speakers. /55.31/ occurs more in BSD, NHD and SJD than in SHUD. A mixed model containing AGE, REGION and the interaction of AGE and REGION as fixed factors was built and compared with a model with only the main effects AGE and REGION. The chi square is significant (p=0.001). The results of AGE*REGION show that, though BSD and NHD use more /55.31/ than SHUD, the age difference in those using /55.31/ in SHUD is significantly higher than that of BSD (p=0.002) and NHD (p=0.000). This means SHUD is changing faster in borrowing /55.31/ than the suburban dialects.

The linguistic constraints of /55.31/ were also investigated. Like the findings of the Wúxī data, bisyllabic words using /55.31/ mostly use modal voice rather than breathy voice in the σ_1 that comes from the Yáng register. This confirms again that dropping the breathy feature of the nucleus is relevant to bisyllabic words of the Yáng register to switch to /55.31/, a sandhi pattern used in the Yīn register.

6.3.2. Shànghǎi urban dialect (SHUD)

Section 2.5.2.2 reviewed three studies on SHUD tone sandhi and found that: (1) SHUD has a typical left-dominant sandhi system, which tends to extend the citation tone of its first syllable rightward; (2) T2.X, T4.X and T6.X in SHUD have merged, using a sandhi pattern transcribed as /22.44/ or /11.44/; (3) Qián (1992) observed /22.52/ pattern in T2/4/6.T1 words; and (4) Xú, Tāng, & Qián (1981) observed /55.31/ being used in T2/4/6.X for a few words. The last issue has been discussed in Section 6.3.1; this section discusses the other three.

No /22.52/ pattern was observed in our total 2694 valid bisyllabic tokens of

SHUD. Since Qián (1992) did not mention which words were pronounced with /22.52/ in his data, the existence of /22.52/ needs further assessment by using an extended wordlist. The current study can only show that /22.52/ is not used in my dataset. Besides /55.31/ (n=29, P=1.1%), the majority tokens of SHUD T2/4/6.X were transcribed by the author as /22.44/ (n=2552, P=94.7%). 113 tokens of /24.31/ (P=4.2%) were also found in this study, 59 used by old people and 54 used by young people. 80.5% of the /24.31/ occurrences in my dataset are in verb-object structure. It is believed that those /24.31/ occurrences have a "narrowly-used pattern", which applies to bisyllabic words that have a loose internal structure and small pauses between their grammatical components (Xú, Tāng, & Qián, 1981; Zhang, 2007). As stated in Chapter 2, this study is limited to the "broad-used patterns" in SHUD, so those /24.31/ tokens are not analyzed here. Mean F_0 contours of SHUD for the /22.44/ tokens are plotted in Figure 6.5, split by AGE, MC_TONE and MC_AB, to see if there are any ongoing changes.

Figure 6.6 shows that mean F_0 contours of SHUD do not show any obvious differences among different conditions of MC_TONE and MC_AB, which means that T2.X, T4.X and T6.X have completely merged. It can also be seen from Figure 6.6 that young people in SHUD are not undergoing contour loss in sandhi forms, as their grey lines in the figure have almost the same pitch height as the black lines of old people. The only slight difference between the young and the old is in σ_2. Young people have a rising contour while old people have a peaking contour. Then a further examination on σ_2 was done and found old people tend to use a level contour with a falling trend at the end of σ_2 while young people's pitch contour of σ_2 is rising. To realize the rising contour, young people have relative lower onset. Although young people do not have obvious contour loss in the sandhi patterns, they still tend to use the rising tone in σ_2, which can be hypothesized to be due to a desire to reduce pitch differences across syllable boundaries.

Then pitch height of onset of σ_2, i.e. $\sigma_2\_ST_{P1}$, is used to index the differences between old and young people. The result of a likelihood ratio test shows that model-0 and model-AGE built on $\sigma_2\_ST_{P1}$ were significantly different ($\chi^2(1)=8.205$, $p=0.004$), indicating it is an ongoing change. $\sigma_2\_ST_{P1}$, which is the only ongoing change found in the sandhi pattern of SHUD, is further examined in Section 6.4.1.

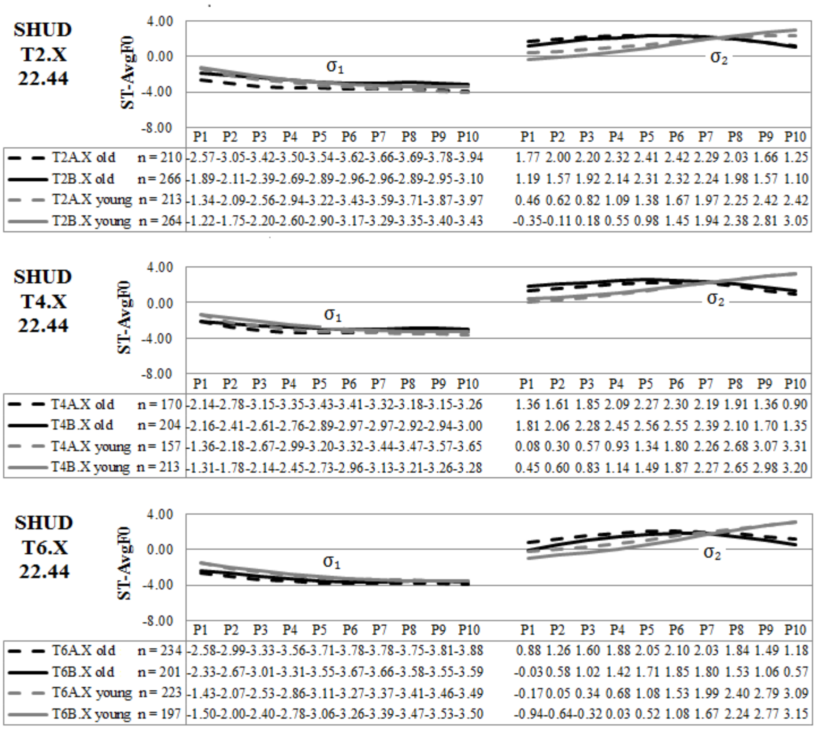


Figure 6.6 Mean F_0 contours of SHUD, split by AGE, MC_TONE and MC_AB

6.3.3. Nánhuì dialect (NHD)

Section 2.5.2.2 reviewed Qián's (1992) transcription on the sandhi pattern of NHH. I first present my own transcription in

Table 6.40. The distinction between my transcription and Qián's (1992) (see Table 2.7) is marked in grey in

Table 6.40. It shows that in some tonal combinations, e.g. T4.T3 and T6.T4, one or two new patterns are added in my transcription. The criteria for adding a new pattern are that the pattern is used by at least one bisyllabic word identified in this category, while this word is used by half or more of the total 20 speakers (n≧10).

Table 6.40 Tone sandhi patterns initialed with T2, T4 and T6 in NHD

<table>
<tr><th></th><th>T1</th><th>T2</th><th>T3</th><th>T4</th><th>T5</th><th>T6</th><th>T7</th><th>T8</th></tr>
<tr><td>T2</td><td colspan="4">22.33</td><td colspan="2">22.34</td><td colspan="2">22.3</td></tr>
<tr><td>T4</td><td colspan="2" rowspan="2">22.42</td><td>24.31
22.42</td><td>24.31
22.24</td><td>22.24</td><td>22.24
22.42</td><td colspan="2" rowspan="2">22.4</td></tr>
<tr><td>T6</td><td>24.31
22.42</td><td>24.31
22.42
22.24</td><td>22.24
22.42
24.31</td><td>22.24
22.42</td></tr>
</table>

Qián (1992) identified two sandhi patterns for T2.X: /22.33/ (/22.3/) and /22.35/, in which /22.35/ is used for T2.T5-6 and /22.33/ (/22.3/) is used for the rest of T2.X. As can be seen from
Table 6.40, my transcription is mostly the same as Qián (1992), except the tone value of /22.35/. First, 35 tokens of 13 bisyllabic words were found to use either /24.31/ or /22.24/, which are two unexpected patterns in T2.X. But none of those 13 bisyllabic words has more than 10 tokens of /24.31/ or /22.24/, so I agree with Qián (1992) that /22.33/ (/22.3/) and /22.35/ are the dominant patterns for T2.X.

Second, I transcribed /22.35/ as /22.34/. Figure 6.7 presents the mean F_0 contours of T2.X in NHD. I use /4/ rather than /5/ for the P_{max} of σ_2 for two reasons. First, the highest point of mean F_0 contours for /22.34/ pattern (P_7 in the old speakers, M=2.99 ST) is only 2.25 ST higher than the P_{min} of σ_2 (P_1, M=0.74 ST). One point difference (/3/ vs /4/) is sufficient to describe the gap of 2.25 ST. Second, the P_{max} of /55.31/ pattern distinguished in NHD (see Section 6.3.1) has its P_{max} over 6 ST in old people, which is transcribed as /5/, so pitch height around 3 ST should be transcribed as scale /4/. Figure 6.7 illustrates that, given the frequency of each pattern, old and young people do not differ greatly in the use of the sandhi patterns. Regarding the acoustic realization of /22.33/, young people's σ_1 is around 1 ST higher than old people's while young people's σ_2 is around 0.5 ST lower than old people's. Together this causes the reduction of pitch differences across syllable boundaries. Regarding the acoustic realization of /22.34/, young people's σ_1 is around 1 ST higher than old people's while young people's P_{max} of σ_2 is around 1 ST lower than old people's, which not only indicates the reduction of pitch differences across syllable boundaries but also the contour loss of σ_2.

Turning to T4/6.X, there are three sandhi patterns identified: /22.52/ (/22.5/), /24.31/ and /22.24/ in Qián's (1992) transcription. Figure 6.8 and Figure 6.9 present the mean F_0 contours of T4.X and T6.X in NHD based on the data of current study. I transcribed /22.52/ (/22.5/) as /22.42/ (/22.4/) with the assistance of acoustic data. As can be seen in Figure 6.8 and Figure 6.9, the P_{max} of σ_2 is around 4 ST in old people, lower than the P_{max} of /55.31/ whose P_{max} is over 6 ST.

Then I transcribed σ_2 as /42/ or /<u>42</u>/ (in checked syllable). Qián (1992) found that /22.42/ only occurred in T4/6.T1−2. However, in my transcription, /22.42/ also occurs in T4.T3, T4.T5 and T6.T3−6. Bisyllabic words using /22.42/ in those newly added combinations are listed in Table 6.41. The other two patterns not recorded in

Qián (1992) but seen in my transcription are the /22.24/ of T6.T4 and /24.31/ of T6.T5. Bisyllabic words using those patterns are also presented in Table 6.41.

.

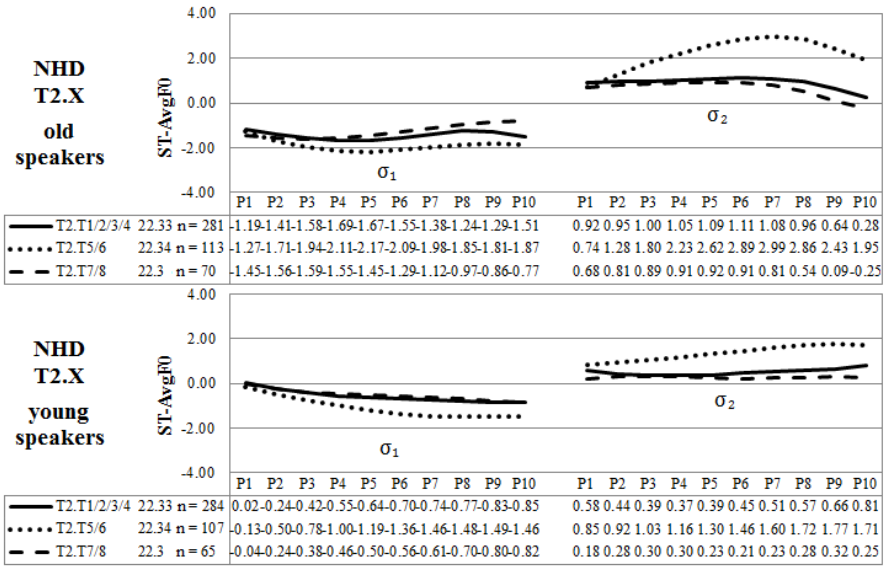
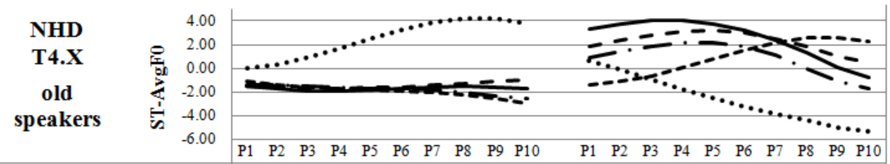


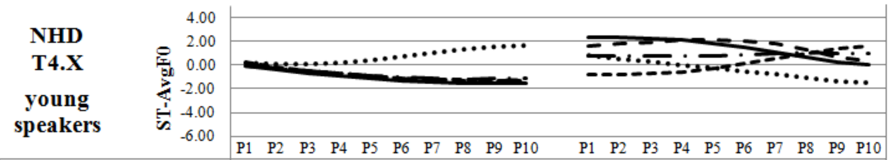
Figure 6.7 Mean F_0 contours of T2.X in NHD, split by AGE and sandhi patterns

Table 6.41 Bisyllabic words in NHD using sandhi patterns different to Qián (1992)

tonal combination	sandhi pattern	bisyllabic words	token n.
T4.T3	/22.42/	某种 'some kind'	n=11
T4.T5	/22.42/	鳝片 'eel slice'	n=10
T6.T3	/22.42/	事体 'affair; matter'	n=17
		面孔 'face'	n=10
		字典 'dictionary'	n=19
T6.T4	/22.42/	谜语 'maze'	n=18
		闹猛 'bustling with noise and excitement'	n=17
	/22.24/	导演 'directors of film, drama, etc.'	n=12
		护士 'nurse'	n=14
		运道 'fortune'	n=13
T6.T5	/22.42/	卖相 'appearance'	n=14
		夜快 'at night fall'	n=15
		浪费 'waste'	n=10
		硬劲 'manage to do something with difficulty'	n=12
		大蒜 'garlic'	n=13
	/24.31/	万岁 'long live'	n=12
T6.T6	/22.42/	夜饭 'supper'	n=20
		烂饭 'rice cooked with too much water'	n=15

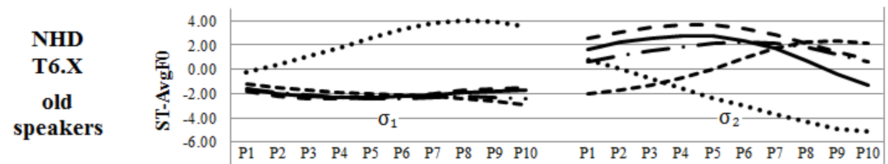


			σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
T2.T1/2	22.42	n=128	-1.50	-1.75	-1.89	-1.92	-1.84	-1.70	-1.57	-1.53	-1.56	-1.69	3.30	3.74	4.01	4.00	3.74	3.22	2.36	1.28	0.10	-0.72
T2.T3/4/5/6	24.31	n=66	0.00	0.32	0.92	1.73	2.52	3.28	3.88	4.22	4.17	3.82	0.68	-0.09	-0.91	-1.72	-2.48	-3.19	-3.81	-4.42	-4.99	-5.36
T2.T7/8	22.42	n=55	-1.50	-1.65	-1.71	-1.71	-1.66	-1.56	-1.43	-1.31	-1.13	-0.97	1.87	2.34	2.79	3.08	3.21	3.02	2.52	1.87	1.02	0.43
T2.T3/4/5/6	22.24	n=121	-1.09	-1.40	-1.56	-1.67	-1.80	-1.93	-2.07	-2.25	-2.58	-3.02	-1.42	-1.12	-0.62	0.04	0.79	1.53	2.17	2.58	2.60	2.22
T2.T3/4/5/6	22.42	n=32	-1.39	-1.44	-1.51	-1.58	-1.67	-1.75	-1.84	-2.02	-2.29	-2.54	0.92	1.43	1.88	2.16	2.14	1.81	1.06	0.01	-1.12	-1.72

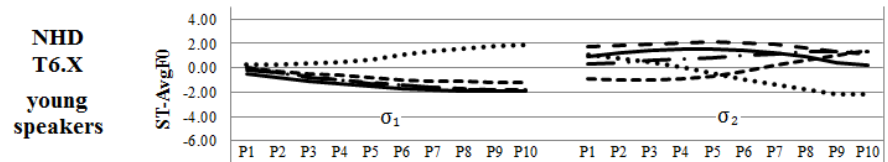


			σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
T2.T1/2	22.42	n=124	-0.04	-0.35	-0.64	-0.90	-1.13	-1.32	-1.47	-1.55	-1.56	-1.51	2.38	2.33	2.25	2.11	1.87	1.54	1.13	0.68	0.25	0.05
T2.T3/4/5/6	24.31	n=63	0.22	0.07	0.06	0.22	0.45	0.70	1.04	1.38	1.59	1.68	0.87	0.53	0.25	0.00	-0.26	-0.49	-0.78	-1.09	-1.38	-1.48
T2.T7/8	22.42	n=55	-0.06	-0.30	-0.49	-0.72	-0.95	-1.13	-1.27	-1.36	-1.32	-1.25	1.62	1.77	1.94	2.12	2.16	2.08	1.77	1.31	0.72	0.33
T2.T3/4/5/6	22.24	n=96	0.22	-0.20	-0.50	-0.70	-0.91	-1.07	-1.18	-1.25	-1.27	-1.31	-0.75	-0.81	-0.72	-0.55	-0.26	0.12	0.53	0.98	1.39	1.62
T2.T3/4/5/6	22.42	n=53	0.16	-0.28	-0.53	-0.73	-0.92	-1.04	-1.15	-1.17	-1.16	-1.08	0.79	0.78	0.82	0.79	0.81	0.92	1.03	1.07	0.99	0.93

Figure 6.8 Mean F_0 contours of T4.X in NHD, split by AGE and sandhi patterns



			σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
T2.T1/2	22.42	n = 118	-1.65	-2.00	-2.20	-2.32	-2.33	-2.23	-2.07	-1.89	-1.76	-1.71	1.66	2.20	2.57	2.78	2.72	2.40	1.74	0.75	-0.38	-1.25
T2.T3/4/5/6	24.31	n = 63	-0.22	0.37	1.09	1.84	2.61	3.30	3.81	4.06	3.96	3.56	0.83	0.11	-0.76	-1.58	-2.33	-3.02	-3.69	-4.29	-4.87	-5.14
T2.T7/8	22.42	n = 60	-1.84	-2.16	-2.41	-2.46	-2.35	-2.17	-1.97	-1.75	-1.58	-1.50	2.54	3.05	3.46	3.65	3.63	3.38	2.86	2.19	1.41	0.85
T2.T3/4/5/6	22.24	n = 81	-1.21	-1.47	-1.68	-1.87	-2.01	-2.13	-2.25	-2.38	-2.58	-2.88	-2.00	-1.68	-1.29	-0.70	0.05	0.90	1.73	2.27	2.40	2.17
T2.T3/4/5/6	22.42	n = 108	-1.60	-1.92	-2.15	-2.31	-2.38	-2.38	-2.32	-2.25	-2.30	-2.40	0.65	1.15	1.55	1.90	2.14	2.24	2.16	1.88	1.26	0.67



			σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
T2.T1/2	22.42	n = 102	-0.45	-0.75	-1.07	-1.34	-1.55	-1.70	-1.81	-1.89	-1.92	-1.89	0.95	1.23	1.41	1.52	1.55	1.46	1.26	0.90	0.45	0.25
T2.T3/4/5/6	24.31	n = 47	0.29	0.28	0.33	0.49	0.72	1.05	1.37	1.62	1.81	1.87	1.06	0.83	0.47	0.05	-0.43	-0.90	-1.32	-1.78	-2.11	-2.17
T2.T7/8	22.42	n = 51	0.01	-0.36	-0.72	-1.00	-1.23	-1.44	-1.56	-1.72	-1.84	-1.80	1.74	1.86	1.98	2.08	2.11	2.03	1.90	1.68	1.33	1.13
T2.T3/4/5/6	22.24	n = 74	0.01	-0.24	-0.46	-0.63	-0.78	-0.95	-1.07	-1.13	-1.18	-1.21	-0.93	-1.00	-1.00	-0.87	-0.67	-0.30	0.19	0.64	1.08	1.41
T2.T3/4/5/6	22.42	n = 109	-0.14	-0.44	-0.75	-1.04	-1.26	-1.44	-1.63	-1.77	-1.81	-1.79	0.31	0.44	0.58	0.74	0.87	1.02	1.18	1.30	1.33	1.34

Figure 6.9 Mean F_0 contours of T6.X in NHD, split by AGE and sandhi patterns

Comparing the pattern choices of old and young people, the frequency of each sandhi pattern listed in figures above show that young people do not differ widely from old people. The contingency table displaying the frequency of sandhi pattern used by each bisyllabic word and split by AGE is presented in Appendix V for reference. This study focuses on the variation of each sandhi pattern's acoustic realizations between old people and young people. As can be observed from Figure 6.7 to Figure 6.9, young people of NHD have contour loss for each sandhi pattern. This is reflected in: (1) the σ_1 of /22.24/ and /22.42/ both having higher pitch contours among young people; (2) the σ_2 contour of /22.33/ having a lower pitch contour among young people, and (3) the σ_1 of /24.31/ and the σ_2 of /22.42/, /24.31/ and /22.35/ all having smaller steepness among young people either in the rising contour or falling contour.

6.3.4. Bǎoshān dialect (BSD)

Section 2.5.2.2 reviewed the transcriptions of Qián's (1992) and county annals (Zhū, 1992) on the sandhi patterns of BSD. My transcription is closer to Qián's (1992) as this study and Qián (1992) both collected data in Bǎoshān Luódiàn. I first present my own transcription in Table 6.42. The distinction between my transcription and Qián's (1992) (Table 2.8) is marked in grey in Table 6.42. Table 6.42 shows that in some tonal combinations, e.g. T2.T1-5, a new pattern is added in my transcription. The criteria for adding the new pattern are the same as with NHD: the new pattern should be used by at least one bisyllabic word identified in this category and that word is used by at least half of the total 20 speakers (n≧10).

Table 6.42 Tone sandhi patterns initialed with T2, T4 and T6 in BSD

<table>
<tr><th></th><th></th><th>T1</th><th>T2</th><th>T3</th><th>T4</th><th>T5</th><th>T6</th><th>T7</th><th>T8</th></tr>
<tr><td>T2</td><td>113</td><td colspan="2">22.42
24.31</td><td>24.31
22.42</td><td colspan="2">24.31
22.24</td><td>24.31
22.24</td><td colspan="2">24.3
23.4?</td></tr>
<tr><td>T4</td><td>323</td><td colspan="2">22.42</td><td colspan="2">22.42
22.24?
24.31</td><td>22.24
22.42</td><td>22.23
24.31
22.42</td><td colspan="2">24.3
22.4</td></tr>
<tr><td>T6</td><td>35</td><td>22.42
24.31</td><td>22.42
22.24?</td><td>24.31
22.24?</td><td>24.31
22.24
22.42</td><td>24.31
22.24?</td><td>24.31
22.24</td><td colspan="2">24.3</td></tr>
</table>

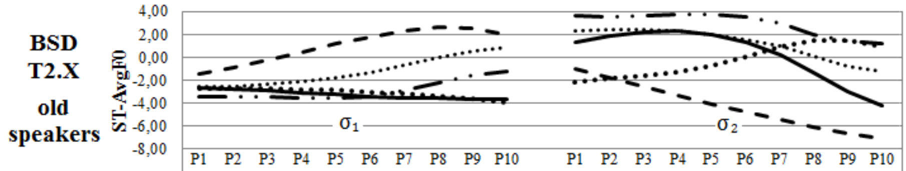
There are three sandhi patterns identified for T2.X: /22.52/ (/22.4/, /23.4/), /24.31/ (/24.3/) and /22.23/ in Qián's (1992) transcription. Figure 6.10 to Figure 6.12 below present the mean F_0 contours of T2.X, T4.X and T6.X in BSD based on the data of the current study. I transcribed /22.52/ as /22.42/ with the assistance of acoustic data. As seen in Figure 6.10, the P_{max} of σ_2 is around 2 ST in old people, lower than the P_{max} of /55.31/ whose P_{max} is over 6 ST. Then I transcribed σ_2 as /42/. As can be seen from Table 6.42 and Table 6.43, this pattern (Qián's /22.52/ or my /22.42/) is also found in T4.T1-6, T6.T1-2 and T6.T4. As illustrated in Figure 6.11 and Figure 6.12, their P_{max} of σ_2 is also around 2 ST in old people, so I transcribed all the /22.52/ as /22.42/. I also transcribed Qián's /22.23/ as /22.24/ because I found the P_{max} of this pattern is around 1.5 ST in old people while its σ_1 is around −3 ST. If σ_1 is transcribed as /22/, considering the pitch difference of 4.5 ST, the P_{max} of σ_2 should be transcribed as /4/ rather than Qián's /3/.

Qián (1992) found that /22.42/ occurred in T2.T1-2, T4.T1-5 and T6.T1−2. However, in my transcription, /22.42/ is also found in T2.T3 and T6.T4. Table 6.43 lists the bisyllabic words using sandhi patterns different to Qián (1992). /24.31/ is newly added in T2.T1-2 and T6.T1 as well. However, /22.24/ identified by Qián (1992) in T4.T3-4, T6.T2-3 and T6.T5 was not found in my dataset, so they are followed by a question mark in Table 6.42. In addition, Qián (1992) documented /33.52/, a pattern only occurring in T4.T6 of BSD. This level tone plus high falling tone is also used by one word of T4.T6 in my dataset: 近视 'myopic' (n=15). However, I transcribed it as /22.42/, not /33.52/.

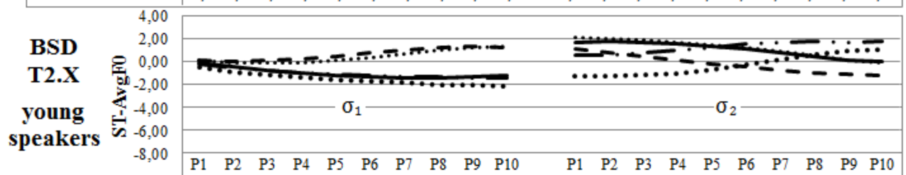
Table 6.43 Bisyllabic words in BSD using sandhi patterns different to Qián (1992)

tonal combination	sandhi pattern	bisyllabic words	token n.
T2.T1	/24.31/	曹操 'Cao Cao, a famous historical figure in China'	n=11
T2.T2	/24.31/	名堂 'variety'	n=12
T2.T3	/22.42/	难板 'seldom'	n=15
T6.T1	/24.31/	便当 'convenient'	n=17
T6.T4	/22.42/	具有 'possess'	n=20

Like NHD, the contingency table displaying the frequency of the sandhi patterns used for each bisyllabic word and split by AGE is presented in Appendix V for reference. Comparing pattern choices of old and young people, the frequency of each sandhi pattern listed in Figures 6.10 to 6.12 shows that young people do not differ markedly from old people. However, at the lexical level there are striking differences: on the one hand, young people may use /24.31/ for a word mainly pronounced as /22.42/ by old people, and on the other hand, use /22.42/ for another word mainly pronounced as /24.31/ by old people, As discussed above, it is hard to decide on the basis of the current dataset which one is the conservative sandhi form. This obstructs further analysis of the variation patterns.

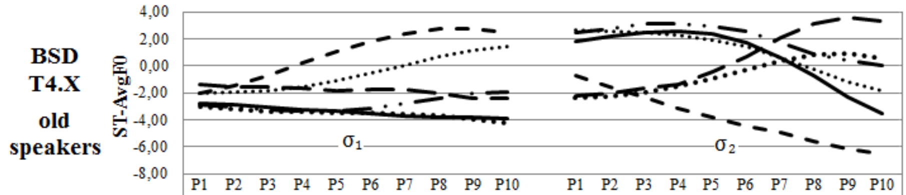


Pattern	n	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
22.42	n = 208	-2,6	-2,7	-2,8	-3,0	-3,2	-3,4	-3,5	-3,5	-3,6	-3,7	1,38	1,91	2,26	2,32	2,01	1,31	0,20	-1,2	-3,0	-4,2
22.24	n = 24	-2,6	-2,7	-2,7	-2,8	-2,8	-3,0	-3,1	-3,3	-3,5	-3,8	-2,1	-1,8	-1,6	-1,2	-0,7	0,09	0,89	1,45	1,46	0,95
24.31	n = 178	-1,4	-0,8	-0,2	0,47	1,17	1,81	2,35	2,65	2,51	2,03	-1,0	-1,7	-2,5	-3,3	-4,0	-4,7	-5,4	-6,0	-6,6	-7,0
24.31	n = 54	-2,5	-2,5	-2,3	-2,0	-1,7	-1,2	-0,6	0,02	0,57	0,89	2,38	2,46	2,44	2,32	2,02	1,60	0,99	0,17	-0,7	-1,2
22.4	n = 8	-3,4	-3,4	-3,4	-3,4	-3,5	-3,4	-2,8	-2,2	-1,5	-1,2	3,62	3,59	3,65	3,78	3,75	3,51	2,95	2,05	1,46	1,28

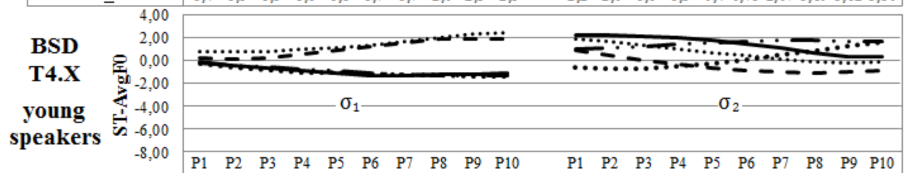


Pattern	n	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
22.42	n = 213	-0,1	-0,4	-0,7	-0,9	-1,1	-1,3	-1,4	-1,3	-1,3	-1,1	1,68	1,73	1,70	1,58	1,38	1,11	0,78	0,43	0,09	-0,0
22.24	n = 18	-0,4	-0,9	-1,1	-1,3	-1,5	-1,7	-1,8	-1,9	-2,0	-2,1	-1,2	-1,2	-1,2	-1,0	-0,7	-0,2	0,17	0,64	0,93	1,10
24.31	n = 170	0,12	0,06	0,11	0,25	0,49	0,76	1,03	1,25	1,32	1,24	1,16	0,81	0,45	0,11	-0,1	-0,4	-0,7	-0,9	-1,1	-1,2
24.31	n = 27	0,11	-0,1	-0,1	-0,0	0,10	0,37	0,70	0,97	1,22	1,38	2,10	2,02	1,86	1,66	1,43	1,17	0,89	0,55	0,16	-0,0
22.4	n = 36	-0,2	-0,5	-0,7	-0,9	-1,1	-1,2	-1,3	-1,3	-1,4	-1,4	0,54	0,61	0,81	1,05	1,28	1,55	1,70	1,75	1,67	1,74

Figure 6.10 Mean F_0 contours of T2.X in BSD, split by AGE and sandhi patterns

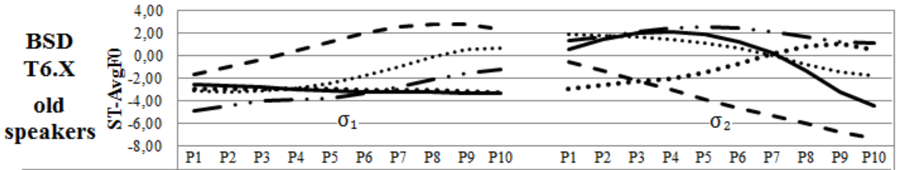


Pattern	n	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
22.42	n = 160	-2,7	-2,9	-3,0	-3,2	-3,3	-3,5	-3,6	-3,7	-3,8	-3,9	1,80	2,22	2,52	2,61	2,37	1,72	0,65	-0,7	-2,3	-3,5
22.24	n = 88	-2,9	-3,2	-3,3	-3,4	-3,4	-3,5	-3,6	-3,7	-3,9	-4,2	-2,3	-2,2	-1,9	-1,5	-0,9	-0,2	0,41	0,88	0,92	0,53
24.31	n = 87	-2,0	-1,4	-0,6	0,25	1,10	1,83	2,38	2,73	2,78	2,44	-0,7	-1,5	-2,3	-3,1	-3,8	-4,4	-4,9	-5,5	-6,1	-6,5
24.31	n = 46	-2,0	-1,9	-1,7	-1,5	-1,1	-0,5	0,03	0,66	1,14	1,41	2,67	2,62	2,50	2,26	1,92	1,42	0,65	-0,1	-1,1	-1,8
22.4	n = 5	-2,8	-3,0	-3,2	-3,3	-3,3	-3,1	-2,8	-2,3	-1,9	-1,9	2,48	2,73	3,11	3,18	2,98	2,55	1,77	0,92	0,40	0,03
22.24	n = 5	-1,4	-1,5	-1,5	-1,6	-1,8	-1,7	-1,7	-2,0	-2,3	-2,3	-2,2	-2,0	-1,6	-1,3	-0,4	0,75	2,07	3,19	3,62	3,30

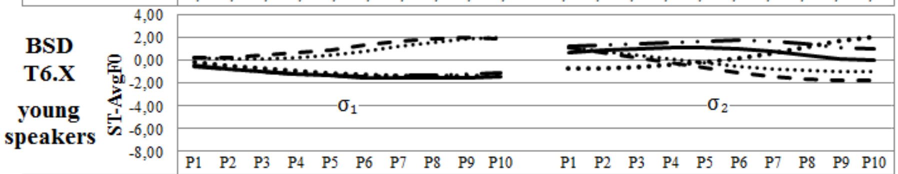


Pattern	n	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
22.42	n = 140	-0,1	-0,4	-0,7	-0,9	-1,1	-1,2	-1,2	-1,2	-1,2	-1,1	2,22	2,17	2,10	1,95	1,75	1,47	1,11	0,72	0,37	0,33
22.24	n = 86	-0,3	-0,6	-0,8	-0,9	-1,0	-1,1	-1,2	-1,3	-1,3	-1,4	-0,5	-0,7	-0,6	-0,5	-0,2	0,06	0,45	0,86	1,23	1,61
24.31	n = 101	0,25	0,16	0,27	0,52	0,85	1,22	1,57	1,84	1,94	1,91	0,90	0,44	0,01	-0,3	-0,6	-0,8	-1,0	-1,0	-1,0	-0,8
24.31	n = 19	0,77	0,73	0,81	0,95	1,12	1,34	1,65	2,05	2,32	2,41	1,92	1,69	1,35	0,98	0,68	0,44	0,16	-0,1	-0,2	-0,0
22.4	n = 30	-0,1	-0,3	-0,5	-0,7	-0,9	-1,1	-1,2	-1,3	-1,3	-1,3	1,05	1,13	1,27	1,39	1,55	1,69	1,78	1,80	1,70	1,62

Figure 6.11 Mean F_0 contours of T4.X in BSD, split by AGE and sandhi patterns



Pattern	n	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
22.42	n = 114	-2,5	-2,6	-2,7	-2,9	-3,0	-3,2	-3,2	-3,2	-3,3	-3,3	0,54	1,43	1,96	2,11	1,89	1,26	0,19	-1,3	-3,1	-4,4
22.24	n = 41	-2,8	-2,9	-2,8	-2,9	-2,9	-3,0	-3,0	-3,0	-3,1	-3,3	-2,9	-2,5	-2,3	-2,0	-1,4	-0,7	0,14	0,89	1,10	0,53
24.31	n = 210	-1,6	-1,0	-0,2	0,50	1,26	1,95	2,49	2,80	2,72	2,30	-0,5	-1,3	-2,1	-3,0	-3,8	-4,6	-5,2	-6,0	-6,7	-7,3
24.31	n = 48	-3,1	-3,2	-3,1	-2,8	-2,4	-1,7	-1,0	-0,1	0,53	0,71	1,93	1,82	1,70	1,46	1,09	0,62	0,02	-0,7	-1,4	-1,8
22.4	n = 11	-4,9	-4,4	-3,9	-3,8	-3,7	-3,3	-2,7	-2,1	-1,5	-1,1	1,36	1,71	2,16	2,48	2,54	2,40	2,07	1,62	1,22	1,12



Pattern	n	σ1 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	σ2 P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
22.42	n = 140	-0,3	-0,7	-0,9	-1,2	-1,3	-1,5	-1,5	-1,5	-1,4	-1,4	0,65	0,85	1,03	1,13	1,10	0,99	0,77	0,48	0,16	-0,0
22.24	n = 56	-0,1	-0,4	-0,6	-0,9	-1,1	-1,3	-1,4	-1,3	-1,3	-1,3	-0,6	-0,6	-0,6	-0,4	-0,1	0,20	0,66	1,19	1,70	2,09
24.31	n = 155	0,24	0,26	0,40	0,63	0,94	1,28	1,61	1,86	1,94	1,92	1,11	0,68	0,24	-0,2	-0,7	-1,1	-1,4	-1,6	-1,8	-1,8
24.31	n = 20	0,24	0,14	0,12	0,21	0,40	0,76	1,19	1,54	1,85	1,98	1,07	0,77	0,45	0,15	-0,1	-0,5	-0,7	-0,8	-0,9	-0,9
22.4	n = 31	-0,5	-0,8	-1,0	-1,1	-1,3	-1,4	-1,3	-1,3	-1,2	-1,1	1,26	1,35	1,47	1,60	1,69	1,73	1,62	1,43	1,14	0,98

Figure 6.12 Mean F_0 contours of T6.X in BSD, split by AGE and sandhi patterns

Thus this study turns to an examination of the variation of each sandhi pattern's acoustic realizations between old people and young people. As can be observed from Figure 6.10 to Figure 6.12, young people of NHD have contour loss for /22.42/, /22.24/ and /24.31/. Like NHD, the contour loss is reflected in: (1) the σ_1s of /22.24/ and /22.42/ (/22.4/) both having higher pitch contour among young people; and (2) the σ_1 of /24.31/, the σ_2s of /22.24/, /22.42/ and /24.31/ all having smaller steepness among young people either in the rising contour or falling contour.

6.3.5. Sōngjiāng dialect (SJD)

Section 2.5.2.2 also reviewed the transcriptions of Qián's (1992) and county annals (Zhū, 1992) on the sandhi patterns of SJD. Like BSD, my transcriptions also show closer similarity to Qián (1992) than the county annals. First, Table 6.44 presents the transcription of author. The patterns identified in my transcription but not in Qián's (1992) (Table 2.9) are marked in grey in Table 6.42. The criteria for adding new patterns are the same as with NHD and BSD. The new pattern should be used by at least one bisyllabic word identified in this category and be used in that word by at least half of the 20 speakers in SJD (n ≧ 10). The words using sandhi pattern different to Qián's (1992) are listed in Table 6.45. The patterns identified in Qián's (1992) transcription but not in mine are followed by a question mark in Table 6.44.

Table 6.44 Tone sandhi patterns initialed with T2, T4 and T6 in SJD

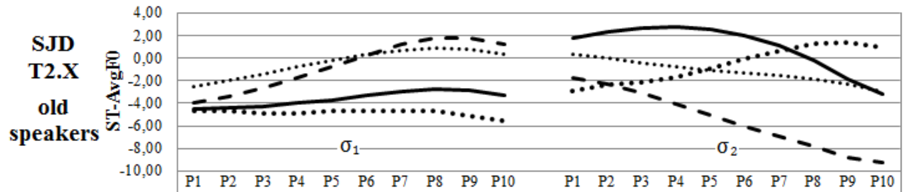
<table>
<tr><th></th><th></th><th>T1</th><th>T2</th><th>T3</th><th>T4</th><th>T5</th><th>T6</th><th>T7</th><th>T8</th></tr>
<tr><td>T2</td><td>²31</td><td>22.42</td><td>22.42</td><td>22.42
24.31(a few) ?</td><td>22.42
24.31?
22.24</td><td colspan="2">24.31</td><td colspan="2">24.31</td></tr>
<tr><td>T4</td><td rowspan="2">113</td><td colspan="2">24.31</td><td>24.31
22.44</td><td>24.31
22.44?
22.24</td><td>22.44
22.24?</td><td>22.24</td><td>24.31</td><td>24.31
22.4</td></tr>
<tr><td>T6</td><td>22.42
22.24
23.44?</td><td colspan="2">24.31
22.24
22.44</td><td>24.31
22.24
22.44</td><td>22.24
22.44</td><td>22.24
22.44?</td><td colspan="2">24.31</td></tr>
</table>

Table 6.45 Bisyllabic words in SJD using sandhi patterns different to Qián (1992)

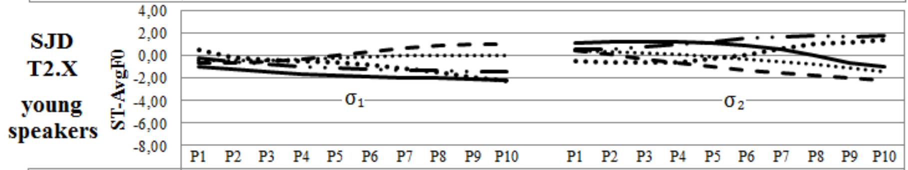
tonal combination	sandhi pattern	bisyllabic words	token n.
T2.T4	/22.24/	渠道 ‘channel’	n=15
T4.T5	/22.44/	负数 ‘negative number’	n=18
		鳝片 ‘eel slice’	n=19
		上劲 ‘enthusiatic about doing something’	n=13
		上照 ‘photogenic’	n=12
		有劲 ‘interesting’	n=20
T6.T1	/22.24/	便当 ‘convienient’	n=20
T6.T2	/24.31/	旧年 ‘last year’	n=11
		烂糊 ‘pulpy’	n=15
	/22.24/	蛋黄 ‘yolk’	n=18
		弄堂 ‘alley’	n=16
		寿头 ‘a fool’	n=14
		用场 ‘function’	n=19
T6.T3	/24.31/	事体 ‘affair; matter’	n=21
		具体 ‘specific’	n=12
		面孔 ‘face’	n=19
	/22.24/	效果 ‘effect’	n=20
		字典 ‘dictionary’	n=11
T6.T4	/24.31/	号码 ‘number’	n=19
		具有 ‘possess’	n=10
		谜语 ‘maze’	n=12
		闹猛 ‘bustling with noise and excitement’	n=18

As in NHD and BSD, I do not agree with Qian’s transcriptions of /22.52/ and /22.23/, and I transcribed them as /22.42/ and /22.24/. The reasons are almost the same as those presented in NHD and BSD. As can be observed from Figure 6.13 to Figure 6.15, the P_{max}s of /22.44/ and /22.24/ are both around 2 ST in old people, which are close to the P_{max} of /24.31/ occurring at the end of the σ_1 of /24.31/ but roughly 2 ST higher than the onset of σ_2 /3/. At the same time, considering that the P_{max} of /55.31/ is around 6 ST, the P_{max}s of /22.44/ and /22.24/ are both transcribed as /4/ rather than /5/. Looking at Table 6.45, the sandhi patterns used by T2.X, T4.X and T6.X in SJD are shown to be very complicated. Three patterns /22.44/, /22.24/ and /24.31/ show up in the 24 tonal combinations without any obvious rules.

Therefore this study has to turn to an examination of the variation of each sandhi pattern's acoustic realizations between old people and young people. As can be observed from Figure 6.13 and Figure 6.15, young people of SJD also have contour loss for /22.42/, /22.44/, /22.24/ and /24.31/. The manifestations of contour loss for /22.42/, /22.24/ and /24.31/ are exactly the same as those found in BSD, namely: (1) the σ_1 of /22.24/ and /22.42/ (/22.4/) both have higher pitch contour among young people; and (2) the σ_1 of /24.31/ and the σ_2s of /22.24/, /22.42/ and /24.31/ all have smaller steepness among young people either in the rising contour or falling contour. In addition, the contour loss for /22.44/ does not occur in SHUD, NHD and BSD. As can be observed in Figure 6.14 and Figure 6.15, young people use a higher pitch contour to realize the σ_1 of /22.44/ but a lower pitch contour to realize the σ_2 of /22.44/.

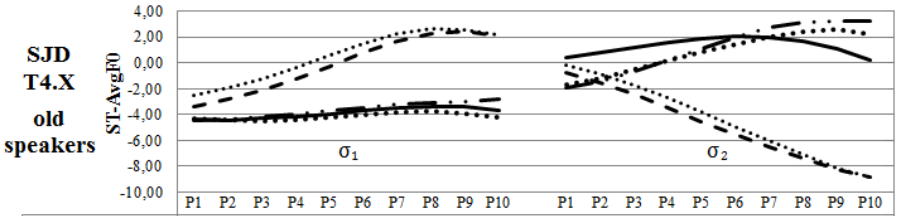


		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
22.42	n = 280	-4,4	-4,3	-4,2	-4,0	-3,6	-3,3	-2,9	-2,7	-2,8	-3,2	1,82	2,37	2,67	2,76	2,55	1,99	1,08	-0,2	-1,8	-3,1
22.24	n = 14	-4,6	-4,7	-4,9	-4,8	-4,7	-4,6	-4,7	-4,6	-5,0	-5,5	-2,8	-2,4	-2,0	-1,6	-0,9	-0,0	0,65	1,29	1,43	0,99
24.31	n = 121	-3,9	-3,4	-2,6	-1,7	-0,7	0,27	1,23	1,82	1,78	1,19	-1,7	-2,3	-3,0	-4,0	-5,0	-6,0	-6,9	-7,8	-8,7	-9,2
24.31	n = 61	-2,5	-1,9	-1,4	-0,7	-0,1	0,34	0,73	0,90	0,77	0,36	0,31	-0,0	-0,3	-0,7	-1,0	-1,3	-1,5	-1,9	-2,3	-2,9

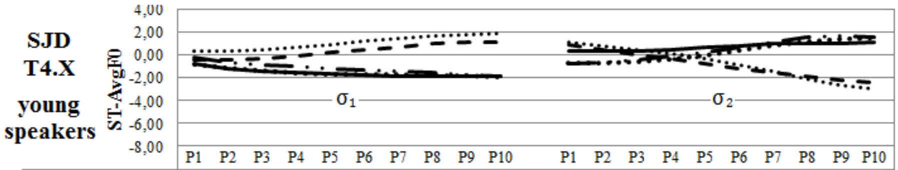


		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
22.42	n = 307	-0,99	-1,25	-1,45	-1,63	-1,77	-1,88	-1,96	-2,01	-2,10	-2,15	1,08	1,18	1,26	1,25	1,14	0,91	0,53	-0,01	-0,66	-0,94
22.24	n = 5	0,55	-0,11	-0,34	-0,49	-0,66	-0,86	-1,13	-1,50	-1,97	-2,23	-0,49	-0,55	-0,65	-0,57	-0,30	0,03	0,61	1,07	1,14	1,44
24.31	n = 99	-0,60	-0,63	-0,53	-0,31	-0,02	0,33	0,65	0,91	1,04	1,00	0,44	0,10	-0,30	-0,69	-1,02	-1,31	-1,56	-1,80	-2,03	-2,17
24.31	n = 67	-0,42	-0,45	-0,39	-0,32	-0,23	-0,11	-0,02	0,00	-0,01	-0,01	0,46	0,34	0,22	0,07	-0,11	-0,30	-0,51	-0,74	-1,11	-1,40
22.4	n = 36	-0,25	-0,56	-0,76	-0,96	-1,10	-1,23	-1,34	-1,35	-1,40	-1,40	0,54	0,61	0,81	1,05	1,28	1,55	1,70	1,75	1,67	1,74

Figure 6.13 Mean F_0 contours of T2.X in SJD, split by AGE and sandhi patterns

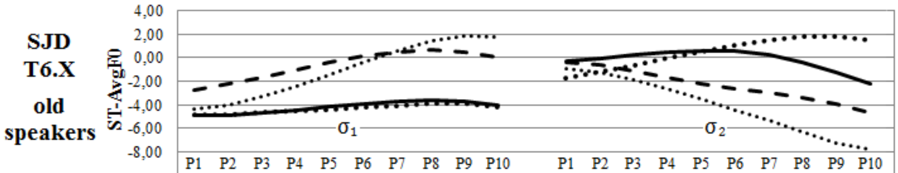


Pattern	n	σ1 P1	σ1 P2	σ1 P3	σ1 P4	σ1 P5	σ1 P6	σ1 P7	σ1 P8	σ1 P9	σ1 P10	σ2 P1	σ2 P2	σ2 P3	σ2 P4	σ2 P5	σ2 P6	σ2 P7	σ2 P8	σ2 P9	σ2 P10
22.44	n = 118	-4,3	-4,4	-4,3	-4,1	-3,9	-3,7	-3,5	-3,3	-3,3	-3,6	0,42	0,78	1,19	1,57	1,87	2,04	2,00	1,72	1,08	0,21
22.24	n = 85	-4,2	-4,4	-4,5	-4,3	-4,1	-3,9	-3,8	-3,7	-3,9	-4,2	-1,7	-1,1	-0,5	0,15	0,82	1,46	2,04	2,45	2,56	2,25
24.31	n = 148	-3,3	-2,8	-2,1	-1,2	-0,2	0,80	1,68	2,30	2,45	2,08	-0,7	-1,5	-2,4	-3,4	-4,5	-5,5	-6,4	-7,3	-8,2	-8,8
24.31	n = 45	-2,5	-1,9	-1,2	-0,3	0,59	1,51	2,22	2,63	2,56	2,21	-0,2	-0,8	-1,7	-2,7	-3,7	-4,9	-6,0	-7,1	-8,1	-8,8
22.24	n = 10	-4,4	-4,3	-4,1	-3,9	-3,7	-3,4	-3,2	-3,0	-3,0	-2,7	-1,9	-1,4	-0,6	0,24	1,11	2,00	2,72	3,16	3,24	3,19

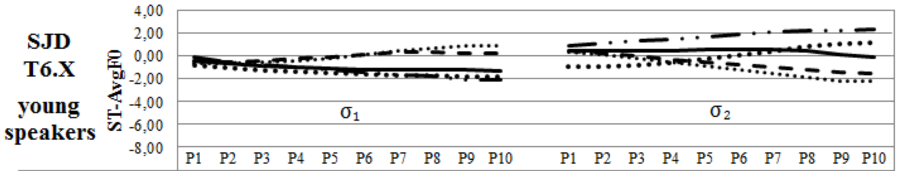


Pattern	n	σ1 P1	σ1 P2	σ1 P3	σ1 P4	σ1 P5	σ1 P6	σ1 P7	σ1 P8	σ1 P9	σ1 P10	σ2 P1	σ2 P2	σ2 P3	σ2 P4	σ2 P5	σ2 P6	σ2 P7	σ2 P8	σ2 P9	σ2 P10
22.44	n = 140	-0,7	-1,1	-1,3	-1,5	-1,6	-1,7	-1,8	-1,8	-1,8	-1,9	0,36	0,31	0,36	0,48	0,65	0,82	0,97	1,05	1,02	1,15
22.24	n = 114	-0,8	-1,2	-1,4	-1,5	-1,6	-1,7	-1,7	-1,7	-1,8	-1,9	-0,7	-0,7	-0,5	-0,3	-0,0	0,39	0,81	1,17	1,37	1,52
24.31	n = 127	-0,4	-0,4	-0,3	-0,1	0,17	0,46	0,72	0,95	1,06	1,09	0,84	0,43	0,03	-0,3	-0,7	-1,1	-1,5	-1,9	-2,2	-2,4
24.31	n = 46	0,34	0,36	0,45	0,64	0,94	1,23	1,48	1,68	1,82	1,88	1,11	0,83	0,48	0,10	-0,3	-0,9	-1,4	-2,0	-2,6	-2,9
22.24	n = 11	-0,2	-0,6	-0,9	-1,0	-1,1	-1,3	-1,4	-1,5	-1,7	-1,8	-0,8	-0,6	-0,3	-0,0	0,28	0,71	1,16	1,55	1,68	1,60

Figure 6.14 Mean F_0 contours of T4.X in SJD, split by AGE and sandhi patterns



Pattern	n	σ1 P1	σ1 P2	σ1 P3	σ1 P4	σ1 P5	σ1 P6	σ1 P7	σ1 P8	σ1 P9	σ1 P10	σ2 P1	σ2 P2	σ2 P3	σ2 P4	σ2 P5	σ2 P6	σ2 P7	σ2 P8	σ2 P9	σ2 P10
22.44	n = 82	-4,9	-4,9	-4,6	-4,4	-4,1	-3,9	-3,7	-3,5	-3,6	-4,0	-0,3	-0,0	0,21	0,44	0,59	0,53	0,25	-0,3	-1,2	-2,2
22.24	n = 174	-4,8	-4,8	-4,6	-4,5	-4,3	-4,2	-4,0	-3,8	-3,9	-4,1	-1,7	-1,2	-0,6	-0,0	0,49	1,03	1,51	1,82	1,81	1,46
24.31	n = 117	-2,8	-2,2	-1,6	-1,0	-0,4	0,11	0,48	0,63	0,46	0,03	-0,3	-0,5	-1,0	-1,6	-2,1	-2,6	-3,0	-3,3	-3,9	-4,6
24.31	n = 58	-4,3	-4,0	-3,3	-2,4	-1,4	-0,4	0,56	1,40	1,84	1,75	-0,8	-1,2	-1,9	-2,6	-3,5	-4,4	-5,3	-6,2	-7,2	-7,8



Pattern	n	σ1 P1	σ1 P2	σ1 P3	σ1 P4	σ1 P5	σ1 P6	σ1 P7	σ1 P8	σ1 P9	σ1 P10	σ2 P1	σ2 P2	σ2 P3	σ2 P4	σ2 P5	σ2 P6	σ2 P7	σ2 P8	σ2 P9	σ2 P10
22.44	n = 104	-0,48	-0,70	-0,85	-0,97	-1,08	-1,17	-1,20	-1,19	-1,22	-1,27	0,44	0,42	0,43	0,47	0,51	0,53	0,52	0,42	0,14	-0,06
22.24	n = 175	-0,80	-1,05	-1,22	-1,38	-1,52	-1,64	-1,69	-1,73	-1,81	-1,84	-0,96	-0,90	-0,77	-0,59	-0,32	0,02	0,42	0,79	1,02	1,17
24.31	n = 95	-0,44	-0,51	-0,44	-0,27	-0,06	0,14	0,29	0,35	0,27	0,20	0,45	0,26	0,00	-0,29	-0,57	-0,81	-1,04	-1,26	-1,46	-1,51
24.31	n = 43	-0,68	-0,68	-0,58	-0,41	-0,17	0,10	0,39	0,67	0,85	0,89	0,37	0,09	-0,20	-0,52	-0,85	-1,22	-1,59	-1,92	-2,20	-2,22
22.4	n = 12	-0,09	-0,50	-0,73	-0,99	-1,26	-1,45	-1,62	-1,82	-2,05	-2,10	0,92	1,15	1,33	1,49	1,66	1,84	2,07	2,18	2,21	2,33

Figure 6.15 Mean F_0 contours of T6.X in SJD, split by AGE and sandhi patterns

To sum up Section 6.3, there is only one ongoing change identified in SHUD. Young people of SHUD tend to use a rising tone in σ_2 for realizing /22.44/. To quantify this ongoing change, the pitch height of onset of σ_2, i.e. $\sigma_2\_ST_{P1}$, is used as an index. Like the citation form, no obvious contour loss has been identified in the sandhi forms of SHUD. However, the sandhi patterns of NHD, BSD and SJD are all undergoing contour loss. The contour loss identified in NHD is reflected in sandhi patterns /22.33/, /22.35/, /22.24/, /22.42/ and /24.31/; the contour loss identified in BSD is reflected in sandhi patterns /22.42/, /22.24/ and /24.31/; the contour loss identified in SJD is reflected in sandhi patterns /22.42/, /22.24/, /24.31/ and /22.44/. Three patterns - /22.24/, /22.42/ and /24.31/ - occur in all three suburban dialects.

In order to make comparisons among three suburban dialects, Section 6.4 focuses on the contour loss of those three sandhi patterns besides the investigation of $\sigma_2\_ST_{P1}$ in SHUD. As with the Wúxī dialect, five indexes were used to quantify the contour loss of /22.24/, /22.42/ and /24.31/. Young people raise the overall pitch height in the σ_1 of /22.24/ and /22.42/, which is quantified by *MeanST_(22).XX*. "MeanST" is calculated for the level tone to get a value that represents the pitch height of the whole contour. "22.XX" indicates that the contour loss is found in either /22.24/ or /22.42/. The parentheses of "(22).XX" indicate the syllable involved in variation, i.e., the σ_1 of /22.24/ or /22.42/. The contour loss of σ_1 and σ_2 in /24.31/ are denoted as *SLOPE_(24).31* and *SLOPE_24.(31)*; the contour loss of σ_2 in /22.24/ is denoted as *SLOPE_22.(24)*; and the contour loss of σ_2 in /22.42/ is denoted as *SLOPE_22.(42)*. "SLOPE" is calculated using different algorithms for rising contours (σ_2 in /22.24/ and σ_1 in /24.31/) and falling contours (σ_2 in /22.42/ and /24.31/). The slope between P_1 and P_{max} is calculated for the steepness of rising contours while the slope between P_{max} and P_{10} is calculated for the steepness of falling contours. The slope between P_{max} and P_1 or P_{10} is calculated to present the steepness of rising or falling contours because, irrespective of being a rising tone or a falling tone, P_{max} is the target for tone realization (Zhū 2010: 277).

In the following Section 6.4, *MeanST_(22).XX*, *SLOPE_(24).31*, *SLOPE_24.(31)*, *SLOPE_22.(24)* and *SLOPE_22.(42)* are investigated by means of likelihood ratio tests, and mixed models for each suburban dialect and presented in Sections 6.4.2 to 6.4.4. Section 6.4 starts with analysis of $\sigma_2\_ST_{P1}$ in SHUD.

6.4. Analyzing the linguistic variables of sandhi tone

6.4.1. Shànghǎi urban dialect (SHUD)

As can be observed in Figure 6.1, old and young people of SHUD differ in the tone shape of σ_2。Young people tend to use a rising tone in σ_2 which causes a low onset, i.e. $\sigma_2\_ST_{P1}$. The conditioning effect of $\sigma_2\_ST_{P1}$ is investigated in this section to better understand this ongoing change. First, the results of likelihood ratio tests between the models with each fixed factor in design, and the null model, are presented in Table 6.46.

Table 6.46 Results of likelihood ratio tests on $\sigma_2\_ST_{P1}$ in SHUD (n=2552)

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	10145			
model-AGE	10137	8.205	1	0.004**
model-SEX	10145	0.183	1	0.669
model-EDU	10144	0.402	1	0.526
model-PTH	10144	0.530	1	0.467
model-STYLE	10145	0.152	1	0.697
model-LEXICAL ATTESTEDNESS	10145	0.008	1	0.930
model-σ_1_MC_TONE	10141	4.277	2	0.118
model-σ_1_PTH_TONE	10143	2.042	2	0.360
model-σ_1_MC_AB	10144	0.769	1	0.381
model-σ_1_HEIGHT	10142	2.845	5	0.724
model-σ_1_DUR	10139	6.133	1	0.013*
model-σ_2_MC_REGISTER	9914.4	230.440	1	0.000***
model-σ_2_PTH_TONE	10110	35.069	3	0.000***
model-σ_2_HEIGHT	10143	1.786	5	0.878
model-σ_2_DUR	10085	60.215	1	0.000***
model-σ_2_CODA	10145	0.069	1	0.792

The exploratory results found that $\sigma_2\_ST_{P1}$ differ in σ_2_MC_TONE. The onset of σ_2 that is from T1, T3, T5 and T7 is generally higher than that of σ_2 that is from T2, T4, T6 and T8. σ_2_MC_TONE can be collapsed to two conditions, the Yīn register and the Yáng register. This new variable is called σ_2_MC_REGISTER. A likelihood ratio test on σ_2_MC_REGISTER was conducted and the results shown in Table 6.46 suggest that it is the most powerful conditioning factor according to the reduced deviance it causes (230.440). Meanwhile, the effect of σ_2_PTH_TONE is correlated to σ_2_MC_REGISTER as PT1 is mostly developed from MC_T1, the tone from high register, while PT4 is mostly developed from MC_T4 and MC_T6, the tone from low register. So σ_2_PTH_TONE is not entered in the final model. Besides, σ_2_MC_REGISTER and σ_2_PTH_TONE, AGE, σ_1_DUR and σ_2_DUR are found to be significant fixed factors in the likelihood ratio tests. AGE*σ_2_MC_REGISTER is found to be a significant interaction.

Table 6.47 Output of mixed model on *$σ_2$_ST1* in SHUD (n=2552)

predictor	*β*	SE	t	*p*
Intercept	0.361	0.390	0.925	0.359
AGE (*old*): *young*	−1.207	0.408	−2.956	0.008^{***}
$σ_2$_MC_REGISTER *(low register)*:				
high register	4.001	0.183	21.839	0.000^{***}
$σ_1$_DUR	−0.001	0.001	−1.388	0.165
$σ_2$_DUR	−0.006	0.001	−7.034	0.000^{***}
AGE*$σ_2$_MC_REGISTER				
young: high register	−0.244	0.125	−1.958	0.050^{*}

As can be seen from Table 6.47, AGE, $σ_2$_MC_REGISTER, $σ_2$_DUR and AGE*$σ_2$_MC_REGISTER are significant in the final model built on $σ_2$_ST1. As expected, young people are more likely to have lower onset while tones from the Yīn register tend to have a higher onset. The longer duration $σ_2$ has, the lower onset it gets. Table 6.48 presents the mean *$σ_2$_ST1* split by AGE and $σ_2$_MC_REGISTER to understand their interactions. The interaction of AGE*$σ_2$_MC_REGISTER shows that young people have smaller onset differences (M=−1.074) between the high and low registers than those of old people (M=1.416).

To explore the relationship between low register and low onset, as discussed in Section 4.1.2, "voiced" obstruents from low register are accompanied by breathiness in isolation but become purely voiced only in unstressed intervocalic positions (Cao & Maddieson, 1992). So $σ_2$ from the Yáng register is believed to have a voiced initial, which should be responsible for the low onset. Young people shrink the difference between the high and low register. The tentative explanation for this is that young people may devoice the voiced initial of $σ_2$ due to the influence of Mandarin.

Table 6.48 Mean *$σ_2$_ST1* in SHUD, split by AGE and $σ_2$_MC_REGISTER

σ2_MC_REGISTER \ AGE	*old*	*young*
high register	3.29	1.88
low register	−1.16	−2.24

6.4.2. Nánhuì dialect (NHD)

Five indexes - *MeanST_(22).XX*, *SLOPE_(24).31*, *SLOPE_24.(31)*, *SLOPE_22.(24)* and *SLOPE_22.(42)* - are investigated for the contour loss of NHD. Those results are presented in Section 2.1.1.6.4.2.1 to Section 2.1.1.6.4.2.5.

6.4.2.1. *MeanST_(22).XX*

MeanST_(22).XX represents the contour loss found in the σ_1 of either /22.24/ or /22.42/, i.e., young people raising the overall pitch height in the σ_1 of /22.24/ and /22.42/. First, likelihood ratio tests were conducted on *MeanST_(22).XX* (see Table 6.49). The variables tested are the five social factors and 11 linguistic factors described in Section 3.4.7. As shown in Table 6.49, AGE, STYLE, σ_1_HEIGHT and σ_2_HEIGHT are significant. Then interactions were investigated but none were found to be significant. So AGE, STYLE, σ_1_HEIGHT and σ_2_HEIGHT were entered into a final model to see their conditioning effect on *MeanST_(22).XX*.

Table 6.49 Results of likelihood ratio tests on *MeanST_(22).XX*

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	7666.5			
model-AGE	7661.9	4.610	1	0.032*
model-SEX	7666.5	0.000	1	0.996
model-EDU	7665.6	0.971	2	0.615
model-PTH	7666.5	0.020	1	0.888
model-STYLE	7651.7	14.845	1	0.000***
model-LEXICAL ATTESTEDNESS	7663.3	3.240	1	0.072
model-σ_1_MC_TONE	7661.1	5.388	2	0.068
model-σ_1_PTH_TONE	7661.9	4.591	2	0.101
model-σ_1_MC_AB	7663.9	2.666	1	0.103
model-σ_1_HEIGHT	7627.4	39.106	5	0.000***
model-σ_1_DUR	7666.5	0.021	1	0.886
model-σ_2_MC_TONE	7654.4	12.161	7	0.095
model-σ_2_PTH_TONE	7660.4	6.111	3	0.106
model-σ_2_HEIGHT	7649.7	16.881	5	0.005**
model-σ_2_DUR	7666.4	0.172	1	0.678
model-σ_2_CODA	7665.9	0.615	1	0.433

Table 6.50 Output of mixed model on *MeanST_(22).XX* in NHD

Predictor	***β***	**SE**	**t**	**p**
Intercept	−0.656	0.356	−1.845	0.073*
AGE (*old*): *young*	0.906	0.418	2.167	0.044*
STYLE *(wordlist):paragraph*	0.585	0.151	3.873	0.000***
σ_1_HEIGHT *(close):*				
near close	−0.259	0.162	−1.605	0.111
close mid	−0.333	0.279	−1.193	0.236
mid	0.098	0.254	0.385	0.701
open mid	−0.500	0.131	−3.829	0.000***
open	−0.790	0.126	−6.292	0.000***
σ_2_HEIGHT *(close):*				
near close	−0.064	0.171	−0.376	0.707
close mid	0.030	0.155	0.193	0.847
mid	0.126	0.166	0.758	0.450
open mid	−0.547	0.138	−3.958	0.000***
open	−0.259	0.135	−1.919	0.048*

The output in Table 6.50 shows that, as expected, young people use higher mean pitch to realize the σ_1 of either /22.24/ or /22.42/. STYLE is also responsible. Once again, paragraph is the context promoting contour loss as its β is positive. σ_1_HEIGHT and σ_2_HEIGHT are significant. If a morpheme has an open mid or open vowel in its σ_1 or σ_2, the pitch height of σ_1 and σ_2 both tend to be lower than the σ_1 or σ_2 in a morpheme having a close vowel. This shows the intrinsic effect of vowel height: close vowels have a higher F_0 than open vowels (Hombert, et al., 1979; Lehiste, 1970).

6.4.2.2. SLOPE_(24).31

Young people tend to pronounce the σ_1 of /24.31/ with a less steep slope. *SLOPE_(24).31* is used to quantify this process. Likelihood ratio tests were conducted on the same predictors: five social factors and 11 linguistic factors. The results are listed in Table 6.51.

Table 6.51 Results of likelihood ratio tests on *SLOPE_(24).31* in NHD (n=213)

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	1682.4			
model-AGE	1666.6	15.765	1	0.000***
model-SEX	1682.1	0.279	1	0.597
model-EDU	1676.7	5.653	2	0.059
model-PTH	1682.3	0.012	1	0.912
model-STYLE	1682.4	0.006	1	0.938
model-LEXICAL ATTESTEDNESS	1681.4	0.996	1	0.318
model-σ_1_MC_TONE	1681.3	1.095	2	0.578
model-σ_1_PTH_TONE	1680.7	1.637	2	0.441
model-σ_1_MC_AB	1682	0.398	1	0.528
model-σ_1_HEIGHT	1679.6	2.754	4	0.600
model-σ_1_DUR	1681.6	0.770	1	0.380
model-σ_2_MC_TONE	1681.7	0.709	1	0.400
model-σ_2_PTH_TONE	1681.4	0.934	1	0.334
model-σ_2_HEIGHT	1677	5.380	3	0.146
model-σ_2_DUR	1681	1.326	1	0.250

The results in Table 6.51 show that the variation of *SLOPE_(24).31* is only constrained by AGE. The slope of σ_1 is smaller for young people (M=15.54) than for old people (M=37.24). Apart from this, it seems no any other factor in our design is responsible for this variation. Then the analyses of the σ_2 of /24.31/ pattern are continued.

6.4.2.3. SLOPE_24.(31)

The slope between P_{max} and P_{10} is used to quantify the contour loss of the σ_2 of /24.31/, which is denoted as *SLOPE_24.(31)*. Likelihood ratio tests were conducted on *SLOPE_24.(31)* .

Table 6.52 Results of likelihood ratio tests on *SLOPE_24.(31)* in NHD (n=213)

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	2035.9			
model-AGE	2030.3	5.588	1	0.018*
model-SEX	2033.1	2.822	1	0.093
model-EDU	2033.5	2.375	2	0.305
model-PTH	2034.1	1.837	1	0.175
model-STYLE	2030	5.865	1	0.015*
model-LEXICAL ATTESTEDNESS	2035.9	0.064	1	0.801
model-σ_1_MC_TONE	2032.1	3.863	2	0.145
model-σ_1_PTH_TONE	2032.1	3.863	2	0.145
model-σ_1_MC_AB	2032.9	3.026	1	0.082
model-σ_1_HEIGHT	2029.1	6.778	4	0.148
model-σ_1_DUR	2035	0.907	1	0.341
model-σ_2_MC_TONE	2035.7	0.178	1	0.673
model-σ_2_PTH_TONE	2033.7	2.245	1	0.134
model-σ_2_HEIGHT	2033	2.966	3	0.397
model-σ_2_DUR	2008.4	27.503	1	0.000***

In Table 6.52, AGE, STYLE and σ_2_DUR are found to be significant but no interaction is significant. So only those three factors were put in the final mixed model.

Table 6.53 Output of mixed model on *SLOPE_24.(31)* in NHD (n=213)

predictor	***β***	**SE**	**t**	***p***
Intercept	−59.277	6.084	−9.743	0.000***
AGE (*young*):*old*	−19.602	5.331	−3.677	0.003**
STYLE *(wordlist):paragraph*	17.987	8.656	2.078	0.048*
σ_2_DUR	0.173	0.031	5.591	0.000***

Table 6.53 shows the effect of σ_2_DUR. The longer the duration of a morpheme, the steeper slope σ_2 has. STYLE is responsible for the variation of *SLOPE_24.(31)* as well. Contour loss is more likely to be found in paragraph reading as σ_2's falling contour has a steeper slope in such a speaking style.

Taking a look at *SLOPE_(24).31* and *SLOPE_24.(31)* together, the contour loss of σ_1 and σ_2 is connected, as the likelihood ratio test finds that *SLOPE_(24).31* is constrained by the pitch height of the onset of σ_2. The higher F_0 σ_2 starts, the steeper slope σ_1 has: $\chi^2(1)=357.75$, $p=0.000^{***}$. The contour loss of σ_1 is only constrained by

AGE but the STYLE effect is found to be responsible for the contour loss of σ_2. It suggests that the STYLE effect directly affects the contour loss of σ_2, but indirectly affects the contour loss of σ_1.

6.4.2.4. *SLOPE_22.(24)*

The slope between P_1 and P_{max} is used to quantify the contour loss of the σ_2 of /22.24/, which is denoted as *SLOPE_22.(24)*. Likelihood ratio tests were conducted on *SLOPE_22.(24)*, and the results listed in Table 6.54.

Table 6.54 Results of likelihood ratio tests on *SLOPE_22.(24)* in NHD (n=390)

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	2922.8			
model-AGE	2910.4	12.451	1	0.000***
model-SEX	2922.3	0.482	1	0.487
model-EDU	2916.1	6.746	2	0.034*
model-PTH	2922.8	0.015	1	0.902
model-STYLE	2922.8	0.008	1	0.928
model-LEXICAL ATTESTEDNESS	2921.1	1.759	1	0.185
model-σ_1_MC_TONE	2918.3	4.471	2	0.107
model-σ_1_PTH_TONE	2920.6	2.180	2	0.336
model-σ_1_MC_AB	2922.8	0.010	1	0.919
model-σ_1_HEIGHT	2917	5.829	5	0.323
model-σ_1_DUR	2919.1	3.701	1	0.054
model-σ_2_MC_TONE	2921.9	0.881	3	0.830
model-σ_2_PTH_TONE	2921.9	0.894	1	0.345
model-σ_2_HEIGHT	2914.0	8.809	5	0.117
model-σ_2_DUR	2911	11.793	1	0.001***

Table 6.54 shows that AGE, EDU and σ_2_DUR are significant. As expected, young people pronounce the rising contour of the σ_2 of /22.24/ with a smaller slope. Longer duration of σ_2 is responsible for a less steep rising σ_2. The effect of EDU depends on the effect of AGE. Young people generally have received higher education than old people. But the effect of EDU is not significant in the dataset of old people alone or in the dataset of young people alone (old people: $\chi^2(1)=0.286$, $p=0.593$; young people: $\chi^2(1)=1.840$, $p=0.175$). STYLE is not significant, and neither is STYLE*AGE. The σ_1 of /22.24/ is constrained by STYLE but σ_2 is not. Then the connection of σ_1 and σ_2 is examined. *SLOPE_22.(24)* is constrained by the pitch height of σ_1's offset. The lower F_0 of σ_1's offset anticipates the steeper slope of σ_2.

In sum, the effect of STYLE is responsible for the contour loss of σ_1 but not for that of σ_2. The contour loss of σ_1 and σ_2 is connected. This situation is similar to the contour loss found in /24.31/, where the effect of STYLE is responsible for the contour loss of σ_2 but not for that of σ_1, which is referred to as "STYLE's unilateral effect". In the next section, STYLE's unilateral effect is examined in the sandhi pattern /22.42/ of NHD.

6.4.2.5. SLOPE_22.(42)

The slope between P_{max} and P_{10} is used to quantify the contour loss of the σ_2 of /22.42/, which is denoted as *SLOPE_22.(42)*. Likelihood ratio tests were conducted on *SLOPE_22.(42)*, and the results are listed in Table 6.55.

Table 6.55 Results of likelihood ratio tests on *SLOPE_22.(42)* in NHD (n=792)

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	7085.1			
model-AGE	7079.8	5.303	1	0.021**
model-SEX	7081.8	3.273	1	0.070
model-EDU	7083.4	1.756	2	0.416
model-PTH	7084.2	0.909	1	0.341
model-STYLE	7082.9	2.221	1	0.136
model-LEXICAL ATTESTEDNESS	7084.9	0.232	1	0.630
model-σ_1_MC_TONE	7081.9	3.179	1	0.075
model-σ_1_PTH_TONE	7081.1	4.034	2	0.133
model-σ_1_MC_AB	7084.9	0.169	1	0.681
model-σ_1_HEIGHT	7078.6	6.523	5	0.259
model-σ_1_DUR	7084.3	0.805	1	0.370
model-σ_2_MC_TONE	7064.1	21.033	5	0.001***
model-σ_2_PTH_TONE	7067.1	18.032	3	0.000***
model-σ_2_HEIGHT	7079.8	0.317	5	0.378
model-σ_2_DUR	7080.3	4.767	1	0.029**
model-σ_2_MC_TONE (collapsed)	7065.4	19.699	1	0.000***

Neither STYLE nor STYLE*AGE is significant. This means that /22.42/ is similar to /22.24/ and /24.31/ in that the STYLE effect is only found in one syllable, i.e. STYLE's unilateral effect. STYLE is thus responsible for the σ_1 of /22.42/ and /22.24/ as well as the σ_2 of /24.31/. Which syllable is constrained by STYLE might be related to which syllable is more marked. The σ_1s of /22.42/ and /22.24/ are both level tones, less marked than their σ_2 – a rising /24/ or falling /42/. The σ_2 of /24.31/ is a falling tone, less marked that its σ_1 – a rising /24/. It seems that the less marked syllable in a sandhi pattern is more likely to be constrained by STYLE.

Turning to the other significant factors in Table 6.55, AGE, σ_2_DUR, σ_2_MC_TONE and σ_2_PTH_TONE are significant. σ_2_MC_TONE and σ_2_PTH_TONE are correlated since PTH_TONE is developed from MC_TONE. Their corresponding relations are shown in Table 3.15 (Chapter 3).

Further investigations into *SLOPE_22.(42)* predicted by those two variables show that *SLOPE_22.(42)* is intrinsically constrained by σ_2_MC_TONE because if σ_2 belongs to MC_T3-T6, *SLOPE_22.(42)* is around -20, while if σ_2 belongs to MC_T1 or T2, *SLOPE_22.(42)* is around −35. Regarding σ_2_PTH_TONE, syllables of PT1 have greater *SLOPE_22.(42)* than syllables from the other three PTH tonal categories because PT1 develops from MC_T1. But the effect MC_T2 cannot be traced in the predictor σ_2_PTH_TONE because MC_T2 and MC_T3 together developed into PT2 and the effect of MC_T2 was reduced in PT2.

To fully illustrate the effect of MC_T1 and MC_T2, σ_2_MC_TONE was chosen for the final model rather than σ_2_PTH_TONE. To show the effect of σ_2_MC_TONE in a better way, a new predictor σ_2_MC_TONE (collapsed) was added to build the mixed model. It collapses MC_T1-2 into one condition and collapses MC_T3 to MC_T6 into another condition. σ_2_MC_TONE (collapsed) does not interact with AGE. So σ_2_MC_TONE (collapsed) and AGE were used to build the final mixed model.

Table 6.56 Output of mixed model on *SLOPE_22.(42)* in NHD (n=792)

Predictor	***β***	**SE**	**t**	***p***
Intercept	-46.490	5.208	−8.926	0.000***
AGE (*old*): *young*	13.301	5.684	2.340	0.031*
σ_2_MC_TONE(collapsed) *(T1-2): T3-6*	13.368	2.696	4.959	0.000***
σ_2_DUR	0.033	0.015	2.197	0.028*

The results of Table 6.56 confirm that young people have smaller *SLOPE_22.(42)* than old people, while both young and old people show smaller *SLOPE_22.(42)* in bisyllabic words whose σ_2 is from MC_T3-T6. σ_2_MC_TONE's significant effect suggests that NHD is probably similar to the Wúxī dialect, using some substitution rules determined by its σ_2. σ_2_DUR shows that the longer duration σ_2 has, the steeper slope it has.

6.4.3. Bǎoshān dialect (BSD)

6.4.3.1. MeanST_(22).XX

Like the analysis of *MeanST_(22).XX* in NHD, the first step is to conduct likelihood ratio tests on *MeanST_(22).XX.* The variables tested are the five social factors and 11 linguistic factors introduced in Section 3.4.7.

Table 6.57 Results of likelihood ratio tests on *MeanST_(22).XX* in BSD (n=1486)

	deviance	Chisq	ChiDf	Pr(>Chisq)
model-0	5296.9			
model-AGE	5278.3	18.638	1	0.000***
model-SEX	5296.9	0.041	1	0.840
model-EDU	5293.1	3.877	2	0.144
model-PTH	5295.4	1.505	1	0.220
model-STYLE	5296.1	0.893	1	0.345
model-LEXICAL ATTESTEDNESS	5295.2	1.760	1	0.185
model-σ_1_MC_TONE	5294.5	2.443	2	0.295
model-σ_1_PTH_TONE	5293.7	3.273	2	0.195
model-σ_1_MC_AB	5294	2.914	1	0.088
model-σ_1_HEIGHT	5259.4	37.521	5	0.000***
model-σ_1_DUR	5293.9	3.097	1	0.078
model-σ_2_MC_ REGISTER	5286.5	10.441	7	0.165
model-σ_2_PTH_TONE	5290.9	6.007	3	0.111
model-σ_2_HEIGHT	5280.8	16.194	5	0.006**
model-σ_2_DUR	5296.4	0.516	1	0.473
model-σ_2_CODA	5294.2	2.709	1	0.100

As can be seen from Table 6.57, AGE, σ_1_HEIGHT and σ_2_HEIGHT are significant. The interaction AGE*STYLE is found to be significant as well. So AGE, σ_1_HEIGHT, σ_2_HEIGHT and AGE*STYLE were entered into the final model together to see their conditioning effects on the variation of *MeanST_(22).XX* in BSD.

Table 6.58 Output of mixed model on *MeanST_(22).XX* in BSD (n=1486)

predictor	β	SE	t	P
Intercept	−2.286	0.342	−6.678	0.000***
AGE (*old*): *young*	1.390	0.428	3.250	0.003**
STYLE *(paragraph): wordlist*	−0.450	0.213	−2.111	0.036*
σ_1_HEIGHT *(close):*				
near close	−0.192	0.191	−1.005	0.317
close mid	0.524	0.303	1.729	0.087
mid	−0.407	0.311	−1.310	0.193
open mid	−0.509	0.151	−3.364	0.001**
open	−0.839	0.148	−5.672	0.000***
σ_2_HEIGHT *(close):*				
near close	−0.196	0.203	−0.967	0.336
close mid	0.005	0.183	0.029	0.977
mid	−0.126	0.175	−0.724	0.471
open mid	0.554	0.193	2.867	0.005**
open	−0.338	0.160	−2.107	0.038*
AGE*STYLE *young*wordlist*	0.739	0.212	3.486	0.001***

Table 6.58 shows that young people use higher mean pitch to realize the σ_1 of either /22.24/ or /22.42/. STYLE is also responsible, and again paragraph is the context for

promoting contour loss as its β is positive. The effects of σ_1_HEIGHT and σ_2_HEIGHT are significant, as was the case for *MeanST_(22).XX* in NHD. If a morpheme has an open mid or open vowel in its σ_1 or σ_2, the pitch height of σ_1 and σ_2 both tend to be lower than the σ_1 or σ_2 in a morpheme having close vowel. It confirms again the intrinsic effect of vowel height (Hombert, et al., 1979; Lehiste, 1970).

Regarding the effect of AGE*STYLE, *MeanST_(22).XX* split by AGE and STYLE is presented in Table 6.59. It shows that although young people still show smaller *MeanST_(22).XX* in the paragraph reading, they have a much smaller slope difference between the two styles than old people.

Table 6.59 *MeanST_(22).XX* in BSD, split by AGE and STYLE

STYLE \ AGE	*old*	*young*
wordlist	−3.16	−1.19
paragraph	−2.61	−1.01

6.4.3.2. *SLOPE_(24).31*

Young people of BSD tend to pronounce the σ_1 of /24.31/ with a less steep slope. *SLOPE_(24).31* is used to quantify this process. Likelihood ratio tests were conducted on the same predictors: five social factors and 11 linguistic factors. The results are listed in Table 6.60.

Table 6.60 Results of likelihood ratio tests on *SLOPE_(24).31* in BSD (n=1040)

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	7758.2			
model-AGE	7739.3	18.905	1	0.000***
model-SEX	7754.7	3.430	1	0.064
model-EDU	7756.1	2.114	2	0.348
model-PTH	7753.3	4.848	1	0.028*
model-STYLE	7752.3	5.883	1	0.015*
model-LEXICAL ATTESTEDNESS	7757	1.195	1	0.274
model-σ_1_MC_TONE	7754.3	3.890	2	0.143
model-σ_1_PTH_TONE	7754	4.171	2	0.124
model-σ_1_MC_AB	7757.7	0.439	1	0.508
model-σ_1_HEIGHT	7755.5	2.662	5	0.752
model-σ_1_DUR	7748.9	9.255	1	0.002**
model-σ_2_MC_TONE	7742.3	15.855	7	0.026*
model-σ_2_PTH_TONE	7751.9	6.286	3	0.100
model-σ_2_HEIGHT	7751.7	6.507	5	0.260
model-σ_2_DUR	7755.6	2.596	1	0.107
model-σ_2_CODA	7741.9	16.255	1	0.000***

Results in Table 6.60 show that the variation of *SLOPE_(24).31* is constrained by

AGE, PTH, STYLE, σ_1_DUR, σ_2_MC_TONE, σ_2_PTH_TONE and σ_2_CODA. To test if the effect of PTH is determined by AGE, model-PTH was built on the data of the old speakers, as the BSD participants who cannot speak PTH are all old people. The results show that model-PTH does not differ from model-0 in old people: $\chi^2(1)=1.967$, $p=0.161$. So the significant effect of PTH in Table 6.60 depends on the effect of AGE. Consequently, PTH was not entered in the final model.

σ_2_MC_TONE and σ_2_CODA are intrinsically correlated. σ_2_CODA was entered into the final model because the output of model-σ_2_MC_TONE shows that syllables with MC_T7 & MC_T8 (σ_2_CODA) have smaller *SLOPE_(24).31* than intercept MC_T1. The interaction between AGE and STYLE is also significant. So, finally, AGE, STYLE, σ_1_DUR, σ_2_CODA and AGE*STYLE were entered into the final model.

Table 6.61 Output of mixed model on *SLOPE_(24).31* in BSD (n=1040)

Predictor	***β***	**SE**	**t**	***p***
Intercept	28.160	3.780	7.449	0.000***
AGE (*old*): *young*	−18.438	4.077	−4.523	0.000***
STYLE *(paragraph): wordlist*	6.387	1.789	3.570	0.001***
σ_1_DUR	−0.031	0.009	−3.450	0.001***
σ_2_CODA *(checked syllable)*:				
smooth syllable	6.853	1.372	4.994	0.000***
AGE*STYLE				
young:wordlist	−3.183	1.793	−1.776	0.076

Table 6.61 is the output of the final mixed model. Young people pronounce the σ_1 of /24.31/ in a steeper slope, showing contour loss. Morphemes in paragraph reading and in checked syllables (MC_T7-8) also tend to be involved in contour loss. σ_1_DUR shows that the longer duration σ_1 has, the more contour loss is involved. AGE*STYLE shows again that young people have smaller slope differences between the two styles than old people. Table 6.62 shows this interaction by the mean scores of *SLOPE_(24).31* in BSD, split by AGE and STYLE.

Table 6.62 Mean *SLOPE_(24).31* in BSD, split by AGE and STYLE

STYLE \ AGE	*old*	*young*
wordlist	34.49	14.99
paragraph	28.54	10.93

6.4.3.3. SLOPE_24.(31)

The slope between P_{max} and P_{10} is used to quantify the contour loss of the σ_2 of /24.31/, which is denoted as *SLOPE_24.(31)*. Likelihood ratio tests were conducted on *SLOPE_24.(31)* of BSD and are shown in Table 6.63.

AGE, STYLE and σ_2_DUR are found to be significant. These are the same factors found in *SLOPE_24.(31)* of NHD. Besides those three main effects, interaction AGE*STYLE is also significant. These four factors were entered in the final mixed model.

Table 6.63 Results of likelihood ratio tests on *SLOPE_24.(31)* in BSD (n=1040)

	deviance	Chisq	ChiDf	Pr(>Chisq)
model-0	10080			
model-AGE	10067	12.963	1	0.000***
model-SEX	10079	0.996	1	0.318
model-EDU	10079	0.980	2	0.613
model-PTH	10078	2.027	1	0.155
model-STYLE	10058	21.556	1	0.000***
model-LEXICAL ATTESTEDNESS	10080	0.132	1	0.716
model-σ_1_MC_TONE	10080	0.421	2	0.810
model-σ_1_PTH_TONE	10080	0.269	2	0.874
model-σ_1_MC_AB	10079	0.699	1	0.403
model-σ_1_HEIGHT	10072	8.446	5	0.133
model-σ_1_DUR	10078	2.053	1	0.152
model-σ_2_MC_TONE	10067	12.583	7	0.083
model-σ_2_PTH_TONE	10074	6.214	3	0.102
model-σ_2_HEIGHT	10070	10.332	5	0.066
model-σ_2_DUR	9939.7	140.290	1	0.000***
model-σ_2_CODA	10077	2.718	1	0.099

Table 6.64 Output of mixed model on *SLOPE_24.(31)* in BSD (n=1040)

predictor	*β*	SE	t	*p*
Intercept	−68.130	6.473	−10.526	0.000***
AGE (*old*): *young*	13.510	6.285	2.149	0.035*
STYLE *(paragraph): wordlist*	−19.120	3.832	−4.989	0.000***
σ_2_DUR	0.193	0.015	12.736	0.000***
AGE*STYLE *young:wordlist*	19.160	5.326	3.596	0.000***

Table 6.64 shows the effect of σ_2_DUR. The longer the duration of a morpheme, the steeper slope σ_2 has. STYLE is responsible for the variation of *SLOPE_24.(31)* as well. Contour loss is more likely to be found in paragraph reading as the second falling contour has a steeper slope in paragraph readings. The interaction AGE*STYLE also shows that young people have smaller slope differences between the two styles than old people. Table 6.65 shows this interaction more directly using the mean scores of *SLOPE_24.(31)* in BSD, split by AGE and STYLE.

Table 6.65 Mean *SLOPE_24.(31)* in BSD, split by AGE and STYLE

STYLE \ AGE	*old*	*young*
wordlist	−57.85	-29.68
paragraph	−38.87	-16.79

6.4.3.4. SLOPE_22.(24)

SLOPE_22.(24) - the slope between P_1 and P_{max} for the σ_2 of /22.24/ - is used to quantify this syllable's contour loss. Likelihood ratio tests were conducted on *SLOPE_22.(24)* and the results listed in Table 6.66.

Table 6.66 Results of likelihood ratio tests on *SLOPE_22.(24)* in BSD (n=280)

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	2032.6			
model-AGE	2022	10.640	1	0.001**
model-SEX	2032.6	0.055	1	0.814
model-EDU	2028.7	3.960	2	0.138
model-PTH	2029.6	3.057	1	0.080
model-STYLE	2026.2	6.460	1	0.011*
model-LEXICAL ATTESTEDNESS	2032.3	0.359	1	0.549
model-σ_1_MC_TONE	2031.2	1.389	2	0.499
model-σ_1_PTH_TONE	2030.7	1.916	2	0.384
model-σ_1_MC_AB	2030.6	2.063	1	0.151
model-σ_1_HEIGHT	2029.2	3.398	5	0.639
model-σ_1_DUR	2032.2	0.387	1	0.534
model-σ_2_MC_TONE	2027	5.677	7	0.578
model-σ_2_PTH_TONE	2030.6	2.068	3	0.558
model-σ_2_HEIGHT	2027.2	5.433	5	0.365
model-σ_2_DUR	2028.7	3.953	1	0.047*
model-σ_2_CODA	2032.5	0.127	1	0.721

Table 6.66, AGE, STYLE and σ_2_DUR are significant. AGE*STYLE is the only significant interaction. Then these four factors were put in the final mixed model.

Table 6.67 Output of mixed model on *SLOPE_22.(24)* in BSD (n=280)

predictor	***β***	**SE**	**t**	***p***
Intercept	30.122	3.233	9.316	0.000***
AGE (*old*): *young*	−9.652	2.507	−3.850	0.001**
STYLE *(wordlist): paragraph*	−8.838	2.843	−3.109	0.003**
σ_2 DUR	−0.026	0.011	−2.317	0.021*
AGE*STYLE *young:paragraph*	6.281	3.365	1.867	0.063

Table 6.67 shows that young people pronounce the rising contour of the σ_2 of /22.24/ with a smaller slope. Significant σ_2_DUR indicates that longer duration of σ_2 is

responsible for a less steep rising σ_2. The effect of STYLE shows up again. Paragraph reading is the context of promoting contour loss as its *β* is negative, indicating a less steep slope is more likely to occur in paragraph reading. The interaction *young*paragraph* is nearly significant in the final model (p=0.063). This shows that young people have smaller slope differences between the two styles than old people because the difference (*β*) in old people is −8.838 while in young people it is −2.557 (β=−8.838+6.281). Table 6.68 presents this interaction, with the mean scores of *SLOPE_22.(24)* in BSD, split by AGE and STYLE.

Table 6.68 Mean *SLOPE_22.(24)* in BSD, split by AGE and STYLE

STYLE \ AGE	*old*	*young*
wordlist	23.51	15.52
paragraph	18.26	13.47

6.4.3.5. *SLOPE_22.(42)*

SLOPE_22.(42) - the slope between P_{max} and P_{10} for the σ_2 of /22.42/ - is used to quantify the contour loss of this falling tone. Likelihood ratio tests were conducted on *SLOPE_22.(42)*, and the results are listed in

Table 6.69.

Table 6.69 Results of likelihood ratio tests on *SLOPE_22.(42)* in BSD (n=1206)

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	9245.1			
model-AGE	9224.7	20.359	1	0.000***
model-SEX	9244.7	0.399	1	0.527
model-EDU	9237.6	7.541	2	0.023**
model-PTH	9237.9	7.215	1	0.007**
model-STYLE	9222.4	22.699	1	0.000***
model-LEXICAL ATTESTEDNESS	9243	2.135	1	0.144
model-σ_1_MC_TONE	9243.2	1.858	2	0.395
model-σ_1_PTH_TONE	9243.2	1.934	2	0.380
model-σ_1_MC_AB	9245	0.067	1	0.796
model-σ_1_HEIGHT	9238.2	6.883	5	0.230
model-σ_1_DUR	9243.2	1.909	1	0.167
model-σ_2_MC_TONE	9214.3	30.808	5	0.000***
model-σ_2_PTH_TONE	9217.2	27.869	3	0.000***
model-σ_2_HEIGHT	9243.3	1.826	5	0.873
model-σ_2_DUR	9236	9.120	1	0.003**
model-σ_2_MC_TONE (collapsed)	9224.5	20.556	1	0.000***

Besides AGE and STYLE, which are the two most frequently significant predictors for contour loss, EDU, PTH, σ_2_MC_TONE and σ_2_PTH_TONE also show significant effects. Thus the effects of EDU and PTH on *SLOPE_22.(42)* were tested to see whether they are independent of AGE. Further likelihood ratio tests showed that the effect of EDU is not significant in the dataset of old people alone or in the dataset of young people alone (old people: $\chi^2(2)=0.908$, p=0.635; young people: $\chi^2(2)=1.337$, p=0.513). The effect of PTH is also not significant in the dataset of old people ($\chi^2(1)=2.511$, $p=0.113$). So the significant effects of EDU and PTH on *SLOPE_22.(42)* are mainly caused by the significant effects of AGE. Consequently, EDU and PTH were not entered into the final model.

σ_2_MC_TONE and σ_2_PTH_TONE are also correlated. Further investigations into *SLOPE_22.(42)* predicted by those two variables show that *SLOPE_22.(42)* is intrinsically constrained by σ_2_MC_TONE: if σ_2 belongs to MC_T3-T6, *SLOPE_22.(42)* is around −30; if σ_2 belongs to MC_T1-2, *SLOPE_22.(42)* is around -40. This split is exactly the same as the split of MC_TONE found in *SLOPE_22.(42)* in NHD (see Section 6.4.2.5). Therefore, a new predictor σ_2_MC_TONE (collapsed) was added to build the mixed model. It collapses MC_T1-2 into one condition and collapses MC_T3-6 into another condition. σ_2_MC_TONE (collapsed) also interacts with AGE. AGE*STLYE is significant here. Thus AGE, STYLE, σ_2_MC_TONE (collapsed), AGE*σ_2_MC_TONE (collapsed) and AGE*STYLE were used to build the final mixed model.

Table 6.70 Output of mixed model on *SLOPE_22.(42)* in BSD (n=1206)

Predictor	***β***	**SE**	**t**	***p***
Intercept	−71.398	4.934	−14.471	0.000***
AGE (*old*): *young*	38.883	5.747	6.766	0.000***
STYLE *(wordlist): paragraph*	24.890	3.462	7.190	0.000***
σ_2_MC_TONE (collapsed) *(T1-2):T3-6*	16.087	2.652	6.066	0.000***
σ_2_DUR	0.044	0.014	3.187	0.002***
AGE*STYLE *young:paragraph*	-19.018	3.640	−5.224	0.000**
AGE*σ_2_MC_TONE (collapsed) *young:T3-6*	−11.423	3.025	−3.776	0.000***

Table 6.70 shows that young people use less steep falling tones for the σ_2 of /22.42/. STYLE is responsible as well. Contour loss is more likely to be found in paragraph reading as the second falling contour has a steeper slope in paragraph reading. Predictor σ_2_MC_TONE (collapsed) plays the same role in *SLOPE_22.(42)* in NHD. It also shows that contour loss is more likely to be found in bisyllabic words whose σ_2 is from MC_T3-T6. As in NHD, σ_2_DUR also shows that the longer duration σ_2 has, the steeper slope it has.

Regarding two significant interactions - AGE*σ_2_MC_TONE (collapsed) and AGE*STYLE - they both show that young people have smaller slope differences than old people between the two styles and between the two conditions of σ_2_MC_TONE (collapsed). To understand these two interactions better, Table 6.71 and Table 6.72 list the mean scores of *SLOPE_22.(42)* in BSD, split by AGE and σ_2_MC_TONE (collapsed) and by AGE and STYLE.

Table 6.71 Mean *SLOPE_22.(42)* in BSD, split by AGE and σ_2_MC_TONE (collapsed).

σ2_MC_TONE(collapsed) \ AGE	*old*	*young*
T1-2	−59.70	−24.13
T3-6	-49.73	-16.76

Table 6.72 Mean *SLOPE_22.(42)* in BSD, split by AGE and STYLE

STYLE \ AGE	*old*	*young*
wordlist	−60.76	−23.51
paragraph	−34.89	−14.52

6.4.4. Sōngjiāng dialect (SJD)

6.4.4.1. MeanST_(22).XX

Like the analyses of *MeanST_(22).XX* in NHD and BSD, likelihood ratio tests were conducted on *MeanST_(22).XX*. The results are listed in Table 6.73

Table 6.73 Results of likelihood ratio tests on *MeanST_(22).XX* in SJD (n=1655)

	deviance	**Chisq**	**ChiDf**	**Pr(>Chisq)**
model-0	5724.9			
model-AGE	5703.9	20.970	1	0.000***
model-SEX	5724.7	0.255	1	0.614
model-EDU	5713.9	11.042	2	0.004**
model-STYLE	5724.8	3.843	1	0.049*
model-LEXICAL ATTESTEDNESS	5724.7	0.257	1	0.612
model-σ_1_MC_TONE	5724.8	0.109	2	0.947
model-σ_1_PTH_TONE	5724.2	0.664	2	0.718
model-σ_1_MC_AB	5723.8	1.143	1	0.285
model-σ_1_HEIGHT	5700.7	24.236	5	0.000***
model-σ_1_DUR	5723.7	1.167	1	0.280
model-σ_2_MC_ REGISTER	5712	12.898	7	0.075
model-σ_2_PTH_TONE	5718.6	6.306	3	0.098
model-σ_2_HEIGHT	5714	10.946	5	0.052
model-σ_2_DUR	5724	0.937	1	0.333
model-σ_2_CODA	5721.4	3.467	1	0.063

Table 6.73 shows that AGE, EDU, STYLE and σ_1_HEIGHT are significant.

σ_2_HEIGHT is very close to significance. σ_1_HEIGHT and σ_2_HEIGHT are both found to be significant in *MeanST_(22).XX* of NHD and BSD. The results in Table 6.73 are in line with the findings in NHD and BSD. The significant effect of EDU is again found to depend on AGE, because the effect of EDU is not significant in either the dataset of old people or in the dataset of young people (old people: $\chi^2(2)=3.639$, p=0.162; young people: $\chi^2(1)=2.854$, p=0.091). No interaction was found to be significant. So AGE, STYLE and σ_1_HEIGHT were entered into the final model to see their conditioning effects on the variation of *MeanST_(22).XX* in SJD.

Table 6.74 Output of mixed model on *MeanST_(22).XX* in SJD (n=1655)

predictor	*β*	SE	t	p
Intercept	−3.950	0.385	−10.267	0.000***
AGE (*old*): *young*	2.977	0.439	6.776	0.000***
STYLE *(wordlist): paragraph*	0.527	0.264	1.994	0.048*
σ_1_HEIGHT *(close):*				
near close	0.037	0.248	0.150	0.881
close mid	0.034	0.419	0.081	0.936
mid	−0.354	0.374	−0.946	0.346
open mid	−0.493	0.204	−2.415	0.017*
open	−0.968	0.204	−4.749	0.000***

Results in Table 6.74 show that young people use higher mean pitch to realize the σ_1 of either /22.24/ or /22.42/. STYLE is also responsible, with paragraph being the context for promoting contour loss as its β is positive. The effect of σ_1_HEIGHT is the same as was found for *MeanST_(22).XX* in NHD and BSD. If a morpheme has an open mid or open vowel in its σ_1, the pitch height of σ_1 tends to be lower than the σ_1 in a morpheme with a close vowel. It confirms again the intrinsic effect of vowel height (Hombert, et al., 1979; Lehiste, 1970).

6.4.4.2. SLOPE_(24).31

Young people in SJD tend to pronounce the σ_1 of /24.31/ with a less steep slope. *SLOPE_(24).31* is used to quantify this process. Likelihood ratio tests were conducted, and the results are shown in Table 6.75. Results show that the variation of *SLOPE_(24).31* in SJD is constrained by AGE, EDU, STYLE, σ_1_DUR and σ_2_DUR. To test if the effect of EDU is determined by AGE, model-EDU was built for the data of the old and young speakers separately. The results show that model-EDU does not differ from model-0 in old people: $\chi^2(2)=1.191$, $p=0.551$. Neither does Model-EDU does differ from model-0 in young people: $\chi^2(1)=1.110$, $p=0.292$. So EDU was not entered into the final model. Besides the significant main effects in

Table 6.60, the interaction between AGE and STYLE is also significant. So, finally, AGE, STYLE, σ_1_DUR, σ_2_CODA and AGE*STYLE were entered into the final model.

Table 6.75 Results of likelihood ratio tests on *SLOPE_(24).31* in SJD (n=971)

	deviance	Chisq	ChiDf	Pr(>Chisq)
model-0	7452.7			
model-AGE	7428.8	23.869	1	0.000***
model-SEX	7452	0.624	1	0.430
model-EDU	7443.7	8.913	2	0.012*
model-STYLE	7448.6	4.054	1	0.044*
model-LEXICAL ATTESTEDNESS	7452.6	0.038	1	0.845
model-σ_1_MC_TONE	7446.8	5.817	2	0.055
model-σ_1_PTH_TONE	7447.9	4.775	2	0.092
model-σ_1_MC_AB	7451.7	0.964	1	0.326
model-σ_1_HEIGHT	7443.8	8.850	5	0.115
model-σ_1_DUR	7417.3	35.319	1	0.000***
model-σ_2_MC_TONE	7443.1	9.557	7	0.215
model-σ_2_PTH_TONE	7451.4	1.220	3	0.748
model-σ_2_HEIGHT	7445.2	7.409	5	0.192
model-σ_2_DUR	7414.6	38.019	1	0.000***
model-σ_2_CODA	7452.6	0.074	1	0.785

Table 6.76 Output of mixed model on *SLOPE_(24).31* in SJD (n=971)

Predictor	***β***	**SE**	**t**	***p***
Intercept	50.701	4.436	11.429	0.000***
AGE (*old*): *young*	−21.650	4.223	−5.126	0.000***
STYLE *(wordlist): paragraph*	7.043	2.973	2.369	0.020*
σ_1_DUR	−0.052	0.011	−4.756	0.000***
σ_2_DUR	−0.045	0.008	−5.498	0.000***
AGE*STYLE *young:paragraph*	−6.770	2.449	−2.764	0.006**

Table 6.76 is the output of the final mixed model. As expected, young people pronounce the σ_1 of /24.31/ with a steeper slope, showing contour loss. σ_1_DUR and σ_2_DUR show that the longer the durations of σ_1 and σ_2, the more likely that the σ_1 of /24.31/ is involved in contour loss. AGE*STYLE again shows that young people have smaller slope differences between two styles than old people. Table 6.77 shows this interaction more directly with the mean scores of *SLOPE_(24).31* in BSD, split by AGE and STYLE.

Table 6.77 Mean *SLOPE_(24).31* in SJD, split by AGE and STYLE

STYLE \ AGE	*old*	*young*
wordlist	41.42	15.20
paragraph	33.01	10.66

6.4.4.3. *SLOPE_24.(31)*

To quantify the contour loss of the σ_2 of /24.31/, the slope between P_{max} and P_{10} was calculated, which is denoted as *SLOPE_24.(31)*. Likelihood ratio tests were conducted and presented in Table 6.78.

Table 6.78 Results of likelihood ratio tests on *SLOPE_24.(31)* in SJD (n=971)

	deviance	Chisq	ChiDf	Pr(>Chisq)
model-0	8848.2			
model-AGE	8830.3	17.838	1	0.000***
model-SEX	8848.2	0.004	1	0.951
model-EDU	8836.8	11.351	2	0.003**
model-STYLE	8843.5	4.714	1	0.030*
model-LEXICAL ATTESTEDNESS	8844.9	3.302	1	0.069
model-σ_1_MC_TONE	8843.3	4.848	2	0.089
model-σ_1_PTH_TONE	8844.6	3.536	2	0.171
model-σ_1_MC_AB	8848.2	0.010	1	0.921
model-σ_1_HEIGHT	8839.9	8.234	5	0.144
model-σ_1_DUR	8844.8	3.380	1	0.066
model-σ_2_MC_TONE	8840.7	7.510	7	0.378
model-σ_2_PTH_TONE	8842	6.184	3	0.103
model-σ_2_HEIGHT	8839.5	8.665	5	0.123
model-σ_2_DUR	8729.5	118.700	1	0.000***
model-σ_2_CODA	8847.6	0.619	1	0.432

Table 6.78 shows that AGE, EDU, STYLE and σ_2_DUR are significant. First, the effects of EDU on *SLOPE_24.(31)* were tested to see if they are independent of AGE. Further likelihood ratio tests showed that the effect of EDU is not significant in the dataset of old and young people separately (old people: $\chi^2(2)=3.448$, $p=0.178$; young people: $\chi^2(1)=0.502$, $p=0.479$). So EDU was not entered into the final model. In addition, no significant interaction was found. So AGE, STYLE and σ_2_DUR were entered into the final mixed model.

Table 6.79 Output of mixed model on *SLOPE_24.(31)* in SJD (n=971)

Predictor	*β*	SE	t	*p*
Intercept	−78.147	7.029	−11.118	0.000***
AGE (*old*): *young*	24.766	5.808	4.264	0.000***
STYLE *(paragraph): wordlist*	−12.658	6.307	−2.007	0.048*
σ_2_DUR	0.200	0.017	0.743	0.000***

Table 6.79 is the output of the final model. Young people have higher slopes for the σ_2 of /22.42/. σ_2_DUR shows that the longer the duration of σ_2, the more likely that the σ_2 of /24.31/ is less steep. Since σ_2_DUR is also the significant predictor for *SLOPE_(24).31*, it shows the correlation of *SLOPE_(24).31* and *SLOPE_24.(31)*.

The effect of STYLE shows that contour loss is greater in paragraph reading than in word list reading.

6.4.4.4. *SLOPE_22.(24)*

SLOPE_22.(24) - the slope between P_1 and P_{max} for the $σ_2$ of /22.24/ - is used to quantify this syllable's contour loss. Likelihood ratio tests were conducted on *SLOPE_22.(24)* and the results listed in Table 6.80.

Table 6.80 Results of likelihood ratio tests on *SLOPE_22.(24)* in SJD (n=587)

	deviance	Chisq	ChiDf	Pr(>Chisq)
model-0	4300			
model-AGE	4287.5	12.556	1	0.000***
model-SEX	4299.8	0.182	1	0.669
model-EDU	4296.4	3.663	2	0.160
model-STYLE	4299.2	0.815	1	0.367
model-LEXICAL ATTESTEDNESS	4299.7	0.317	1	0.574
model-$σ_1$_MC_TONE	4296.4	3.580	2	0.167
model-$σ_1$_PTH_TONE	4295.8	4.216	2	0.122
model-$σ_1$_MC_AB	4299.7	0.296	1	0.586
model-$σ_1$_HEIGHT	4293	6.992	5	0.221
model-$σ_1$_DUR	4297.1	2.900	1	0.089
model-$σ_2$_MC_TONE	4291.8	8.188	5	0.146
model-$σ_2$_PTH_TONE	4295.7	4.296	3	0.231
model-$σ_2$_HEIGHT	4292.1	7.918	5	0.161
model-$σ_2$_DUR	4291	9.020	1	0.003**

AGE and $σ_2$_DUR are found to be significant but no interaction is significant, so only those two factors were put in the final mixed model. No significant effects are found for STYLE or AGE*STYLE. The contour loss of $σ_1$ in /22.24/ is constrained by STYLE but $σ_2$ is not. The "STYLE's unilateral effect" shows up again. *SLOPE_22.(24)* is also constrained by *MeanST_(22).24.*

Table 6.81 Output of mixed model on *SLOPE_22.(24)* in SJD (n=587)

Predictor	***β***	**SE**	**t**	***p***
Intercept	35.409	2.939	12.048	0.000***
AGE (*old*): *young*	−10.621	2.439	−4.354	0.000***
$σ_2$ DUR	−0.035	0.011	−3.292	0.001**

Table 6.81 is the output of final mixed model built on *SLOPE_22.(24)* in SJD. Young people show contour loss. $σ_2$_DUR also shows that the longer the duration of $σ_2$, the more likely that the $σ_2$ of /24.31/ is less steep.

6.4.4.5. *SLOPE_22.(42)*

SLOPE_22.(42) - the slope between P_{max} and P_{10} on the $σ_2$ of /22.42/ - is used to quantify the contour loss of this falling tone. Likelihood ratio tests were conducted on *SLOPE_22.(42)* and the results listed in Table 6.82.

Table 6.82 Results of likelihood ratio tests on *SLOPE_22.(42)* in SJD (n=471)

	deviance	Chisq	ChiDf	Pr(>Chisq)
model-0	5134.3			
model-AGE	5114.6	19.698	1	0.000***
model-SEX	5133.8	0.537	1	0.464
model-EDU	5114.1	20.212	2	0.000***
model-STYLE	5129.6	4.689	1	0.030*
model-LEXICAL ATTESTEDNESS	5133.8	0.474	1	0.491
model-σ_1_MC_AB	5132.9	1.414	1	0.235
model-σ_1_HEIGHT	5128.6	5.700	5	0.337
model-σ_1_DUR	5134.2	0.087	1	0.768
model-σ_2_MC_TONE	5121.5	12.857	7	0.076
model-σ_2_PTH_TONE	5129.9	4.380	3	0.223
model-σ_2_HEIGHT	5129.2	5.075	5	0.407
model-σ_2_DUR	5134	0.332	1	0.564
model-σ_2_CODA	5127.2	7.082	1	0.008**

Table 6.82 shows that AGE, EDU, STYLE and σ_2_CODA are significant. Model-EDU does not differ from model-0 in old people ($\chi^2(2)=2.105$, $p=0.349$) or young people ($\chi^2(1)=2.008$, $p=0.157$). So the significant effect of EDU is correlated with AGE and was not entered into the final model. Besides three main effects, AGE*STYLE reveals its significant effect again.

Table 6.83 Output of mixed model on *SLOPE_22.(42)* in SJD (n=471)

Predictor	*β*	SE	t	*p*
Intercept	−35.796	7.649	−4.680	0.000***
AGE (*old*): *young*	21.656	6.622	3.270	0.002**
STYLE *(paragraph): wordlist*	−13.638	4.974	−2.742	0.008**
σ_2_CODA *(checked syllable): smooth syllable*	−14.230	6.212	−2.291	0.024*
AGE*STYLE *young:wordlist*	10.127	4.461	2.270	0.024*

Table 6.83 shows that young people use less steep falling tones for the σ_2 of /22.42/. STYLE has a significant effect as well. Contour loss is more likely to be found in paragraph reading as the second falling contour has a steeper slope in paragraph reading. The significant *young:wordlist* shows that young people have smaller slope differences than old people between the two styles. Table 6.84 lists the mean scores of *SLOPE_22.(42)* in SJD, split by AGE and STYLE to illustrate this interaction.

Table 6.84 Mean *SLOPE_22.(42)* in SJD, split by AGE and STYLE

STYLE \ AGE	*old*	*young*
wordlist	−63.96	-31.935
paragraph	−47.86	−25.837

σ_2_CODA shows that contour loss is more likely to be found in bisyllabic words whose σ_2 is from MC_T7 and MC_T8, i.e., the checked syllable. This is because *SLOPE_22.(42)* for a checked syllable is higher than it is for a smooth syllable. This is exactly the same as with *SLOPE_22.(44)* for the Wúxī dialect (Section 5.4.1). As discussed in Chapter 5, the realization of σ_2's falling contour is in a conflicting context since σ_1 ends in /2/ while σ_2 starts in /4/. In the checked syllable, the duration of which is much shorter than that of smooth syllable, it is even harder to fully realize the high falling tone for the σ_2. So the checked syllable of /22.42/ tends to reduce the steepness of its falling trend to increase the naturalness and to minimize articulatory effort.

6.4.5. Summary and discussion

Section 6.4 examines five shared ongoing changes of sandhi patterns in suburban Shànghǎi by mixed models: *MeanST_(22).XX*, *SLOPE_(24).31, SLOPE_24.(31), SLOPE_22.(24)* and *SLOPE_22.(42)*. In sum, this section first compares the significant constraints found in three suburban dialects.

Table 6.85 discloses the following information:

(1) Besides *SLOPE_(24).31, SLOPE_22.(24)* and *SLOPE_22.(42)* of NHD, and *SLOPE_22.(24)* of SJD, the other 11 dependent variables are all constrained by STYLE, either as a main effect and/or in interaction with AGE. The effects of STYLE in these variables all point in the same direction: in paragraph reading contour loss occurs more heavily than in word list reading. This confirms that contour loss is mainly motivated by the reduction of articulatory effort in the less formal style.

Four variables are not constrained by STYLE, and these are indicated with a grey background in Table 6.85. As noted above, they are showing "STYLE's unilateral effect". STYLE is responsible for the σ_1 of /22.42/ and /22.24/ and the σ_2 of /24.31/ in NHD, but not responsible for the σ_2 of /22.42/ and /22.24/ and the σ_1 of /24.31/. Similarly, STYLE is responsible for the σ_1 of /22.24/ in SJD but not responsible for its σ_2. It means that in those four dependent variables only one of the syllables is constrained by STYLE. According to Yip (2001), contour tones are more marked than level tones and rising tones are more marked than falling tones. The σ_1 of /22.42/ and /22.24/ are both level tones, less marked than their σ_2 - a rising /24/ or a falling /42/. The σ_2 of /24.31/ is a falling tone, less marked that its σ_1 - a rising /24/. It seems that the less marked syllable in a sandhi pattern is more likely to be constrained by STYLE and affected by the economy of effort. Less marked syllables are pronounced with less articulatory effort.

Table 6.85 Significant constraints for the contour loss of sandhi patterns in Shànghăi

	NHD	BSD	SJD
MeanST_(22).XX	AGE STYLE σ_1_HEIGHT σ_2_HEIGHT	AGE STYLE σ_1_HEIGHT σ_2_HEIGHT AGE*STYLE	AGE STYLE σ_1_HEIGHT
SLOPE_(24).31	AGE	AGE STYLE σ_1_DUR σ_2_CODA AGE*STYLE	AGE STYLE σ_1_DUR σ_2_DUR AGE*STYLE
SLOPE_24.(31)	AGE STYLE σ_2_DUR	AGE STYLE σ_2_DUR AGE*STYLE	AGE STYLE σ_2_DUR
SLOPE_22.(24)	AGE σ_2_DUR	AGE STYLE σ_2_DUR AGE*STYLE	AGE σ_2_DUR
SLOPE_22.(42)	AGE σ_2_DUR	AGE STYLE σ_2_DUR σ_2_MC_TONE (collapsed) AGE*STYLE AGE*σ_2_MC_TONE (collapsed)	AGE STYLE σ_2_ CODA AGE*STYLE

(2) σ_1_HEIGHT and σ_2_HEIGHT are found to be significant for *MeanST_(22).XX* in NHD and BSD while σ_1_HEIGHT is significant for *MeanST_(22).XX* in SJD. This confirms once more the intrinsic vowel effect (Hombert, et al., 1979; Lehiste, 1970) and also suggests that σ_1_HEIGHT or σ_2_HEIGHT are possible predictors for the dependent variable associated with F_0 rather than slope or pitch shapes. These findings are in line with the findings for the Wúxī dialect.

(3) σ_1_DUR is significant for *SLOPE_(24).31* in BSD and SJD. σ_2_DUR is significant for *SLOPE_24.(31), SLOPE_22.(24)* and *SLOPE_22.(42)* in NHD and BSD, as well as for *SLOPE_24.(31)* and *SLOPE_22.(24)* in SJD. The effects of

STYLE for those variables are all the same: the longer duration one syllable has, the more likely this syllable has contour loss. This finding is not in line with the near ceiling performance hypothesis (Y. Xu & Liu, 2012), which claims that speech is produced near an overall performance ceiling in terms of articulatory effort. In accordance with near ceiling performance hypothesis, the reduction of suprasegmental components is likely due to time pressure rather than due to economy of effort. However, our results appear to contradict the near ceiling performance hypothesis because longer duration is found to be a context of promoting contour loss rather than preventing it in the present study.

(4) σ_2_CODA is found to be significant for *SLOPE_22.(42)* in SJD but not in NHD and BSD. σ_2_CODA shows that contour loss is more likely to be found in bisyllabic words whose σ_2 is from MC_T7 and MC_T8, i.e., the checked syllable in SJD. In other words, the falling trend in σ_2 is lost if it is from MC_T7 and MC_T8 in SJD. σ_2_CODA is not a predictor for *SLOPE_22.(42)* of NHD and BSD because those two dialects have already lost the falling trend in σ_2 and only have the /22.4/ pattern in checked syllables, rather than /22.<u>42</u>/ patterns. So SJD, NHD and BSD show coherence towards the contour loss of /22.42/. It is difficult to fully realize σ_2's falling contour in checked syllables in a conflicting context like /22.42/. So the checked syllables of /22.42/ are more likely to undergo contour loss to increase the naturalness and to save articulatory effort.

6.5. Summary and discussion for Chapter 6

With reference to Goal I, two ongoing changes were identified in the citation tone: *DelayRising* and *Contour loss*, which was described in Section 6.1 and analyzed in Section 6.2. Regarding tone sandhi, contour loss was also identified in suburban dialects and was analyzed in Section 6.4.

DelayRising is mainly internally motivated rather than a contact-driven change as there is no difference for NT_{Pmin} between those speaking PTH and those not. The significant effects of STYLE on *DelayRising* suggest that *DelayRising* is an attention-induced change in SHUD, SJD and BSD. It is more likely to occur in the less formal speaking style, i.e., paragraph reading in this study. However, *DelayRising* in NHD is motivated by tonal convergence. The tone of T4A and T6 morphemes (/13/) is attracted by the tone of T2 and T4B morphemes (/113/) mainly because they have similar tonal shapes rather than speakers' tendency to save articulatory effort. So in NHD, STYLE is not significant.

Contour loss is internally motivated as well since STYLE plays the same role in this ongoing change. *Contour loss* is not found in urban but only in suburban areas. It means *contour loss* is also constrained by REGION. Urban young people are quite resistant to contour loss because they want to maintain their identity of Shànghǎi urbanite. In effect, urban and suburban young speakers show divergence towards the contour loss in Shànghǎi. This is not only reflected in the contour loss of citation tone, but also in the contour loss of tone sandhi. As analyzed in Section 6.3, the Shànghǎi urban area does not show any contour loss among young people, but all three suburban dialects are undergoing contour loss to some extent. In addition, urban and suburban young speakers also show divergence towards the /55.31/ pattern, which is a sandhi pattern borrowed from Yīn register and a register producing a higher pitch than the Yáng register. In Chapter 5, it is found that there are associations between externally triggered *WXfalling* with suburban lead in the Wúxī data. Here in the Shànghǎi data, this kind of suburban lead shows in a more obvious way that the suburban dialects are undergoing contour loss while the urban dialect is not. Furthermore, the contour loss in the Shànghǎi suburban dialects is also triggered externally. Suburban young people tend to have acoustic reduction in realizing tonal contours because they have linguistic insecurity and inferiority regarding the dialect they use. With reference to Goal II, the contour loss found in the Shànghǎi data is externally triggered; meanwhile the contour loss is internally constrained by the principle of economic effort, which further reinforces its transmission in suburban young people.

Besides contour loss in tone sandhi, another ongoing change found in the sandhi patterns is that old and young people of SHUD differ in the tone shape of σ_2. This is only found in the urban dialect but not in suburban dialects. Urban young people tend to use a rising tone in σ_2 for /22.44/ pattern and consequently a low onset, i.e. $\sigma_2\_ST_{P1}$. The analysis found that this change is probably caused by the devoicing of σ_2's initial, which is a characteristic of Mandarin. In sum, both Shàngh ǎi urban and suburban dialects are undergoing changes. The suburban areas are undergoing rather drastic changes, such as the reduction of tonal categories, borrowing sandhi patterns and reducing steepness of rising tone and convexity of peaking tone. However, the variation found in the urban area is small. In regard to *DelayRising* or the lowering of $\sigma_2\_ST_{P1}$, urban young speakers seem to make subconscious efforts to maintain the urban accent and urban identity.

Chapter 7. Discussion and conclusion

The present dissertation is a sociophonetic study on tonal variation in the Wúxī and Shànghǎi dialects. This chapter summarizes and discusses the results found in the previous chapters. In Chapter 1, two general goals and two methodological issues of this study were formulated. Section 7.1 discusses the methodological issues since the findings on the domain of tone and tone normalization are the basis for the acoustic analysis of the whole study. Section 7.2 and Section 7.3 discuss and evaluate Goal I of this study. Section 7.2 presents the tonal variation identified in the citation tone. Linguistic variables found in the Wúxī and Shànghǎi dialects are discussed together. Section 7.3 offers a description and discussion of the variation of tone sandhi in the Wúxī and Shànghǎi dialects. In Section 7.4, the significant internal, external and extra-linguistic factors found in the present study are discussed in terms of Goal II. Section 7.5 discusses the methodological and theoretical implications for the fields of Chinese dialectology, urban language studies and sociolinguistics, and offers several prospects for future research.

7.1. Methodological findings

7.1.1. Issue I revisited: the domain of tone in variationist research

In tone languages, the tonal features are restricted to the voiced part of a syllable. However, not all voiced parts of the syllable are tone-bearing. Therefore, the concept "domain of tone" needs to be defined as the part of the syllable from which the F_0 should be extracted for tonal analyses. In order to make valid measurements of F_0, the domain of tone in the Wú dialects was defined before performing acoustic analyses.

Three voiced portions - voiced onset (C), the prenuclear onglide (G) and the final nasal consonant (X) - were investigated in Section 4.1 to see if they constitute the domain of tone for Wú. The principal idea was first to define a Base Set of CV syllables initialed with voiceless obstruent whose domain of tone is undoubtedly the nucleus (V). Then F_0 contours and durational values from syllables containing one of the tested portions - voiced onset, the prenuclear onglide (G) and the final nasal consonant (X) – are compared quantitatively with F_0 contours from the same syllable without those portions to see which one fits the Base Set better. The Root

Mean Square Deviation (RMSD) was used to calculate the distance between two contours. A smaller RMSD indicates a better overall fit.

In light of evidence from tonal shape and duration, it was found that the initial voiced consonant does not carry too much crucial tone information, but the final nasal consonant does. Regarding the prenuclear onglide, the data from Shànghǎi (Sōngjiāng) suggested that it does not make any difference whether the onglide is included or excluded, but the Wúxī data supported the inclusion of the onglide in the domain of tone. It is difficult to mark the boundary between onglide and nucleus clearly, so the onglide was measured as part of the domain of tone to make the criteria more practical. In sum, the domain of tone in the Wú dialects includes three portions: the (optional) prenuclear onglide (G), the nucleus (V) and the (optional) final nasal consonant (X).

7.1.2. Issue II revisited: tone normalization

In order to make accurate statements about between-speaker differences of tone, the fundamental frequency associated with linguistic tone needs to be normalized. Previous tone normalization procedures mainly served for categorizing tones, but did not aim to preserve sociolinguistic variation. Therefore, it is necessary to re-evaluate the effectiveness of tone normalization procedures from the perspective of variationist studies. Following the sociophonetic studies of vowel variation (Adank, 2003; Adank, et al., 2004; van der Harst, 2011), three criteria were used to evaluate a tone normalization procedure: (1) preserves phonemic variation, (2) minimizes anatomical variation, and (3) preserves sociolinguistic variation. The current study compared sixteen normalization procedures by linear discriminant analysis (LDA).

The results show that ST-x_H (a semitone transformation relative to the mean of speaker-constant P_{max} in Hz), ST-x_{max} (a semitone transformation relative to the P_{max} in Hz of each speaker's data) and ST-AvgF_0 (a semitone transformation relative to each speaker's average pitch in Hz) are top three normalization procedures. Considering cross-linguistic comparison and the stability of x_H (the mean of speaker-constant P_{max} in Hz), x_{max} (the P_{max} value in Hz of each speaker's data) and AvgF_0 (each speaker's average pitch in Hz), ST-AvgF_0 was found to be the best tone normalization procedure and hence, this was used to normalize all the F_0 values in this study.

7.2. Goal I revisited: findings on the variation of citation tone

The first general goal of this study is to present a description of tonal variation in Wú Chinese: the urban and suburban dialects spoken in Wúxī and Shànghǎi, including citation tones on monosyllables and sandhi tones on bisyllables. This section summarizes the finding on the variation of citation tone.

Four ongoing changes were identified in the citation tone in the Wúxī urban and suburban dialects: *WXfalling*, *WXlevel*, *WXT4* and *WXT2/6*. Two ongoing changes were identified in the citation tone in the Shànghǎi dialects: (1) *DelayRising* in the Shànghǎi urban dialect (SHUD), Nánhuì dialect (NHD), Bǎoshān dialect (BSD) and Sōngjiāng dialect (SJD); and (2) *contour loss* of BSD and SJD. In this section, the same type of linguistic variables identified in the Wúxī and Shànghǎi dialects will be discussed together.

WXT4 belongs to Zhuó Shǎng Guī Qù (T4A > T6). It causes the complete merger of three lax tones in the Yáng register (T2, T4 and T6) in the Wúxī dialects, where T2 and T6 have already merged. *DelayRising* found in NHD (*NHD_Delay/13/*) also results in the merger of T2, T4 and T6. *WXT4* and *DelayRising* in NHD are discussed together in Section 7.2.1 as they cause similar mergers. *WXT2/6*, the contour loss process found in the Wúxī dialect, is discussed in Section 7.2.2 together with the contour loss found in BSD and SJD. Next, *DelayRising* found in SHUD (*SHUD_Delay/13/)*, BSD (*BSD_Delay/113/)* and SJD (*SJD_Delay/13/*) is discussed in Section 7.2.3. Finally, the tonal borrowing *WXfalling* and *WXlevel* is discussed in Section 7.2.4.

7.2.1. Merger of lax tones in the Yáng register

Two merger processes of lax tones in the Yáng register are identified in the Wúxī and Shànghǎi dialects: *WXT4* and *DelayRising* in NHD.

WXT4 is a variable identified in the T4 morphemes of the Wúxī dialects. This variable has three variants: peaking tone, low rising tone and low dipping tone. To decide which one is the conservative form and which one is innovative, the distribution of the three variants split by AGE and REGION was investigated. The acoustic investigation of each variant also helps to identify the trajectory of change: first the peaking tone shortens its falling part after P_{max} to become a direct rising tone, and then the low rising tone raises its onset and lengthens its falling part before P_{min} to become the low dipping tone, i.e., peaking tone > low rising tone > low dipping tone. Considering the trajectory of change, *WXT4* proceeds in a phonetically gradual manner and is therefore a "regular sound change" (Labov 1994: 543), suggesting *WXT4* is internally motivated in its early stages (peaking tone > low rising tone). Meanwhile, the conditioning effects of STYLE on "low rising tone > low dipping tone" suggest that it is both internally and externally motivated. On the one hand, paragraph reading promotes greater NT_{Pmin} (the normalized time of F_0 minimum) of low rising tone or dipping tone in the data of old people, suggesting "low rising tone > low dipping tone" is associated with the reduction of speech effort internally; on the other hand, young people have much higher NT_{Pmin} in the reading of word lists and minimal pairs, suggesting their borrowing of PT3 externally.

DelayRising in NHD refers to a process of merging T4A and T6 morphemes with T2 and T4B morphemes in NHD. The conservative form of T4A and T6 morphemes is /13/ while that of T2 and T4B is /113/. The merger is completed by moving P_{min} of /13/ backwards in order to approach /113/. *DelayRising* in NHD is a typical process of tonal convergence as /13/ and /113/ have the same pitch height as well as close pitch shapes. At the same time, *DelayRising* is also constrained by SEX, in which old males are using the smallest NT_{Pmin}. The third principle of Labov's gender paradox focuses on men's conservativeness, stating that: "In linguistic change from below, women use higher frequencies of innovative forms than men do" (Labov, 2001: 292). This principle would imply that *DelayRising* is a change from below, which is a linguistic change that operates within the (linguistic) system, below the level of individuals' conscious awareness (ibid: 279). Apart from AGE, SEX, DUR and AGE*SEX, *DelayRising* in NHD is not constrained by other factors. So it is concluded that *DelayRising* is internally motivated.

Turning to the other Shànghǎi dialects, SHUD has finished the merger of T2, T4 and T6; and BSD and SJD have finished the merger of T4 and T6. As a result of the merger, the Shànghǎi urban and suburban dialects have the same outcome: tone /113/. Regards the speed of the changes, SHUD first finished the merger of T2, T4 and T6, and then NHD followed. BSD and SJD turn out to be the slowest in terms of the merger among T2, T4 and T6, because the young people of BSD and SJD can still keep a distinction between T2 (a peaking tone) and T4/6 (a low rising tone).

In sum, in the Shànghǎi dialects, the merger of lax tones in the Yáng register suggests the convergence of suburban dialects towards urban dialects. This in turn suggests that although *DelayRising* in NHD is evidenced to be an internal change when judged by its own data, it is constrained by REGION externally when linking it with other Shànghǎi urban and suburban dialects.

Similarly, the Wúxī urban and suburban dialects also show the same outcome: tone /213/. Suburban speakers use more conservative forms than urban speakers for *WXT4*, suggesting the same convergence.

Juxtaposing the findings in Wúxī and Shànghǎi, the mergers of lax tones in the Yáng register in those dialects all are changes that started in the urban area and spread from the urban area to the suburbs. So those changes are both internally and externally motivated.

7.2.2. Contour loss

Contour loss refers to the acoustic reduction of tonal shapes in this study. To be specific, it can be a reduction of steepness in rising or falling tones and a reduction of convexity in peaking tones. Contour loss in citation tone is evidenced by the ongoing changes identified in T2/6 of the Wúxī dialects, and T2 and T4/6 of the Bǎoshān and Sōngjiāng dialects. The low rising T2/6 morphemes used in Wúxī urban and suburban areas are undergoing the reduction of steepness. A rising contour of T2/6 with a less steep slope is favored by young people. The peaking T2 morphemes of BSD and SJD are undergoing the reduction of convexity while the rising T4/6 morphemes of BSD and SJD are also undergoing the reduction of steepness like WXT2/6.

Strictly speaking, contour loss is also observed in Wúxī T4 morphemes and Nánhuì T4A & T6 morphemes. Wúxī T4 morphemes are undergoing a merger with T2/6, so they are not investigated for contour loss. Likewise, T4A and T6 of NHD are undergoing *DelayRising*, and so they are not investigated for contour loss either. In contrast, T2 and T4B morphemes of NHD are not undergoing any contour loss. This is perhaps because they are the merging target of T4A and T6 morphemes and hence are more stable.

Our data also demonstrates that the Shànghǎi urban dialect is not involved in any significant contour loss in citation tone. This suggests that contour loss is motivated by urbanization and regional contact. Due to the increased contact with urban speakers, suburban young people lack confidence in the use of their own dialect. This was borne out in the interviews, where young speakers expressed feelings of linguistic insecurity and inferiority regarding the dialect they use.

Therefore, they tend to discard the tonal features of their own dialects. In comparison, urban young people, especially those from the Shànghǎi urban area, show less linguistic insecurity (or more "urban pride") based on the questionair (see Section 6.2.3) and hence are more resistant to contour loss. This shows the divergence of suburban dialects and urban dialects.

Contour loss is also constrained by STYLE. In paragraph reading, rising tones or peaking tones reveal more acoustic reduction in steepness or convexity. Speakers tend to pay less attention to their speech in paragraph reading than in word list reading, and therefore are more economic in articulatory effort. So contour loss is internally motivated as a result of economic effort. Meanwhile, as stated above, contour loss is also externally motivated due to urbanization and regional contact. To conclude, contour loss of citation tone is an ongoing change that is both internally and externally motivated.

7.2.3. DelayRising

The processes of young people moving P_{min} backward in the rising tone /13/ or /113/ are collectively referred to as *DelayRising*. *DelayRising* is identified and analyzed in SHUD, SJD, BSD and NHD. *DelayRising* of NHD has been discussed in Section 7.2.1 so it is not repeated here. Strictly speaking, T4 morphemes in the Wúxī dialects are also involved in *DelayRising* in the step of changing from the direct rising tone into the delayed rising tone (step 4 in Figure 5.5).

DelayRising is not a change induced by contact with PTH because there is no difference of NT_{Pmin} between those speaking PTH and those not. *DelayRising* is similar to *contour loss* in terms of internal constraint STYLE. The paragraph reading promotes more delayed rising tones than word list reading, indicating that *DelayRising* is induced by reducing articulatory effort. *DelayRising* is also externally constrained by SEX. Old males are the most conservative group showing resistance to *DelayRising* in BSD while both old and young males are resistant to *DelayRising* in SJD. As mentioned above, the third principle of Labov's gender paradox is that "women use higher frequencies of innovative forms than men do" during linguistic change from below (Labov, 2001: 292). This principle would also imply that *DelayRising* is a change from below, which is a linguistic change that operates within the (linguistic) system, below the level of individuals' conscious awareness (ibid: 279). Moreover, *DelayRising* has never previously been documented in the Shànghǎi dialects. It is also hard to detecte by ear in our data but proceeds in a phonetically gradual manner. These features of *DelayRising*, including a change from below, acoustic-perceptual non-salience and "regular sound change" (Labov 1994: 542), all lead to the conclusion that *DelayRising* is an internally-induced change.

7.2.4. Borrowing PTH tone

WXfalling and *WXlevel* refer to the small number of high falling and mid level tones in the Wúxī dialect which are interpreted as the borrowings of PTH falling tone (PT4) and neutral tone. The same high falling tokens and mid level tokens are also identified in the Shànghǎi dialects, but no age differences were found in those dialects, so they were not defined as ongoing changes.

The mixed model shows that Wúxī suburban young people use significantly more falling tokens than old people and urban young people. Suburban young speakers do not differ from urban speakers in their proficiency of local dialects and they also do not have more intensive contact with PTH, so the only possible explanation for suburban young people's tone borrowing is that they have greater linguistic insecurity about their local dialect. Almost every suburban young speaker can distinguish the accent differences between their dialect and urban dialect. A majority of them think their local dialect is less pleasant to hear and less useful than PTH as shown by the results of their interview (see Section 5.2.1), while fewer urban young speakers have such attitudes.

7.3. Goal I revisited: findings on the variation of tone sandhi

Returing to the second part of Goal I, this section summarizes the ongoing changes of sandhi patterns. Firstly, contour loss of sandhi patterns identified in the Wúxī and Shànghǎi dialects are discussed in Section 7.3.1. Then, in Section 7.3.2, a sandhi pattern /55.31/ identified in the Wúxī and Shànghǎi dialects is discussed.

7.3.1. Contour loss

The contour loss of sandhi pattern is analyzed in the unit of the syllable and thus has the same reflections as citation tone: a reduction of steepness in rising or falling tones and a reduction of convexity in peaking tones. Contour loss occurs in every sandhi pattern used in the Wúxī dialects and Shànghǎi suburban dialects. Yet, for the sake of comparison, only /22.44/, /22.24/ and /24.31/ were investigated in the Wúxī dialects, and only /22.42/, /22.24/ and /24.31/ were investigated in the Shànghǎi suburban dialects. To quantify the contour loss found in those patterns, different indexes were used according to tonal shapes. For the level tone like the σ_1 of /22.24/ and /22.44/, the F_0 values of ten points abstracted from this level contour were averaged to get a mean score that represents the pitch height of the entire contour. They were called "*MeanST_(22).24*" and "*MeanST_(22).44*". For the rising tone like the σ_2 of /22.24/, the slope between P_1 and P_{max} was calculated to quantify the steepness of rising contour. For the falling tone like the σ_2 of /22.42/, the slope between P_{max} and P_{10} was calculated as an index for the steepness of falling contours. Table 6.86 gives a summary of contour loss investigated in different dialects.

Table 6.86 Contour loss of sandhi patterns investigated in Wúxī and Shànghǎi

Wúxī urban & suburban dialects	NHD, BSD & SJD
MeanST_(22).24	*MeanST_(22).XX*
MeanST_(22).44	*SLOPE_(24).31*
SLOPE_24.(31)	*SLOPE_24.(31)*
SLOPE_22.(24)	*SLOPE_22.(24)*
SLOPE_22.(42)	*SLOPE_22.(42)*

Comparing the results of Wúxī and Shànghǎi, one shared characteristic is that they are all constrained by STYLE, in which the paragraph reading shows more acoustic reduction of rising tones or peaking tones than the word list reading. This is in line with the findings for the contour loss identified in citation tone. It can be concluded that the tonal reductions, no matter whether in the citation tone or tone sandhi, are internally associated with the reduction of speech effort.

Turning to the differences, Wúxī urban and suburban dialects do not have a different changing rate towards contour loss while Shànghǎi dialects do. No contour loss was identified in the Shànghǎi urban area. Like the contour loss of citation tone, Shànghǎi suburban dialects show divergence towards the Shànghǎi urban dialect. Thus, the contour loss of sandhi patterns in Shànghǎi is both internally and externally motivated. However, the effect of the external factor (REGION) was not found in the Wúxī data, suggesting the contour loss of citation in the Wúxī dialects is only internally motivated.

7.3.2. Borrowing of /55.31/ pattern

A small number of /55.31/ patterns occur in both the Wúxī dialects and Shànghǎi dialects but mostly used by young people. So the occurrence of /55.31/ is an ongoing change. /55.31/ of T2.X, T4.X and T6.X in those dialects occurs due to the borrowing of the sandhi pattern from the Yīn register, whose nucleus of σ_1 is in the modal voice. In our dataset, all the T2.X, T4.X and T6.X phrases using /55.31/ patterns lost their breathy voice and are in modal voice for their first syllables. So this borrowing is probably caused by dropping the breathy voice in T2.X, T4.X and T6.X. The borrowing of /55.31/ pattern is indirectly influenced by the modal voice of PTH and thus the outcome is to borrow a pattern within the tonal inventory of Wú rather than a pattern of PTH. In comparison with the borrowings in citation tone (Section 7.2.4), the borrowing of high falling tone or mid level tone in citation system are directly from PTH tone.

7.4. Goal II revisited: constraints on tonal variation and change

The second general goal of this study is to elicit the internal, external and extra-linguistic constraints as well as their interaction within the variationist paradigm. Section 7.4.1 will summarize the significant internal factors identified and Section 7.4.2 is for the significant external factors.

7.4.1. Internal factors

This section intends to present and discuss the internal factors for tonal variations identified in this study.

STYLE (paragraph reading & word list reading)

Chapter 1 summarized a model regarding the relationship between speaking style and vowel reduction. It is repeated here: speaking style is associated with vowel space because different styles are tied to the amount of attention paid to speech. It is assumed that less attention is paid to informal speech and therefore it is produced with less articulatory effort. Consequently, reduced attention yields reduced vowel spaces, i.e., more vowel reduction. Since this model has shown its validity in the study of vowel reduction, the first task of this study was to identify whether STYLE is associated with tonal reduction and whether this model is valid for tonal research. In this study, word list reading was used to represent the formal style and paragraph reading was a less formal style.

As noted above, the acoustic reduction of tonal shape identified in this study is called contour loss. Contour loss was found in both the citation tone and tone sandhi. Most contour loss identified in this study is constrained by STYLE (paragraph reading vs word list reading). Monosyllabic morphemes or bisyllabic phrases in paragraph reading are always more likely to have contour loss than those in word list reading. Reading paragraphs is a less monitored style than reading word list. Thus these findings make clear that STYLE is a significant explanatory factor of tonal reduction. Informal speaking style is intrinsically associated with paying less attention and less articulatory effort. So it is concluded that contour loss is attention-induced and constrained by economy of effort.

MC_TONE and MC_AB

In order to investigate Zhuó Shǎng Guī Qù ongoing in the Wúxī dialect, MC_TONE and MC_AB were used as predictors. Meanwhile, these two predictors were tested in the other tonal variables and found to be significant not only in *WXT4*, but also in *DelayRising* of NHD (*NHD_Delay/13/*). In contrast, the conditioning effects of MC tonal category were not significant in *contour loss* identified in all the six dialects, *DelayRising* of SHUD, BSD and SJD and the borrowings in citation tone and tone sandhi. It seems that some tonal variations are in accordance with the MC tonal categories but some are not. What are the differences of *WXT4* and *NHD_Delay/13/* from the others? As summarized above, *WXT4* and *NHD_Delay/13/* are both phonemic mergers emerged within the Wú dialects while the other processes of *DelayRising* and contour loss are phonetic change.

Regarding phonemic merger, the tonal convergence is hypothesized as one of the internal motivations. In Chapter 1, the assumption is proposed that tones from different MC categories are more likely to merge if they are similar in tonal shape. As mentioned above, *NHD_Delay/13/* provides the best evidence for accepting this assumption since /13/ and /113/ have the same pitch height as well as close pitch shapes.

In contrast, *WXT4* does not seem to demonstrate tonal convergence, which is the tendency for a tonal category to change towards another category due to their tonetic similarity. The most advanced form of *WXT4* is a low dipping tone, quite distinct from the most conservative form, which is a peaking tone. Although the conservative variant of *WXT4* appears very different from T6, the merger between them is still triggered and proceeds in a phonetically gradual manner. As shown in Figure 5.5, the peaking variant took four steps to become a dipping variant. The distinction between each two neighboring steps is subtle. This might be an explanation for why the Wúxī dialects can maintain T4 category so long.

To sum up, two internal motivations for tonal variation are identified in this study: the principle of economic effort and tonal convergence. The principle of economic effort causes the phonetic reduction of tonal shapes while the tonal convergence causes the phonemic reduction. In conclusion, they both cause the simplification of tonal systems, which is the general development tendency of the Wú dialects.

7.4.2. External factors and extra-linguistic factors

As discussed in Chapter 1, LEXICAL ATTESTEDNESS, PTH_TONE, REGION and PTH are the expected external factors. This section first discusses their effects. Besides those three factors, our data shows that the "minimal pair" condition of STYLE and MORPHEME can also predict contact-driven changes, e.g., tonal borrowing. So the discussion of STYLE (minimal pair) and MORPHEME follows. Finally, two extra-linguistic factors - SEX and EDU - are discussed.

LEXICAL ATTESTEDNESS

LEXICAL ATTESTEDNESS refers to whether a phrase occurs only in the Wú dialects or both the Wú dialects and PTH (*Wú lexical item* versus *PTH & Wú lexical item)*. LEXICAL ATTESTEDNESS is found to be significant in predicting the split of /24.31/ and /22.44/ for Type II T2.X in the Wúxī dialects. Young people, who are more influenced by PTH, tend to use a rising tone for the σ_1 of *PTH & Wú lexical items* and then adopt /24.31/ for the whole phrase. This shows the conditioning effects of LEXICAL ATTESTEDNESS on the tone borrowing of PT2. In addition, for the contour loss in the σ_2 of /22.24/, *Wú* lexicons have much higher slopes in the σ_2 of /22.24/ than *PTH* and *Wú* lexicons, indicating conservative features. Those two findings together support the notion that the shared PTH and Wú lexical items are more likely to occur with the advanced form in the external tonal variation while Wú lexical items are more resistant towards advanced forms.

PTH_TONE

The conditioning effects of PTH_TONE are the most important evidence to demonstrate the borrowing of PTH tone in an ongoing change. PTH_TONE is found to be significant for three variables:

(1) *WXfalling*: falling tokens are more likely to be found in the morphemes having a corresponding tone of PT4.
(2) *WXT2/6*: regarding the contour loss of WXT2/6, the contour loss of T2 morphemes among young people is significantly slower than T6 morphemes. The force to keep the rising steepness of T2 morphemes comes from young people's borrowing of PT2. T2 morphemes mostly use PT2, a mid rising tone /35/ in PTH. Young people tend to borrow PT2 to

pronounce T2 morphemes and consequently cause its rising slope to become steeper than the slope of T6. Thus the contour loss of T2 morphemes is slower than T6 morphemes.

(3) *DelayRising*: The borrowing of PT3 in the dipping tone is embedded in the *DelayRising* of SHUD, BSD and SJD. T4B morphemes are found to have greater NT_{Pmin} than others. The corresponding PTH tone of T4B is PT3, a dipping tone /214/. It is believed that the large NT_{Pmin} is caused by the borrowing of PT3.

It is relatively easy to identify the borrowing of PT4 because the borrowing tone differs markedly from the original tone. This kind of borrowing is called explicit borrowing. However, the borrowing of PT2 and PT3 are quite difficult to detect since the borrowed tone almost have the same tonal shapes as the original tones. This is called implicit borrowing. Explicit borrowing can be identified on the basis of phonetic transcription, while the identification of implicit borrowing relies more on acoustic analyses. The importance of acoustic analyses is comfirmed in this case.

REGION

An urban-suburban contrast has been found in *WXfalling*, *WXT4*, and *SLOPE_22.(44)* in the Wúxī dialects while in the Shànghǎi dialects, this contrast is reflected in the borrowing of /55.31/ and contour loss of citation tones and sandhi tones. *WXT4* and *SLOPE_22.(44)* causes convergence between urban and suburban areas while the others all indicate divergence. The effects of REGION have been detailed in Section 7.3.2 and 7.3.3. Generally speaking, the divergence of urban and suburban areas is caused by the linguistic insecurity of suburban young people. Such divergence is wider in Shànghǎi than in Wúxī, which is the greatest difference of those two cities. In other words, the "urban pride" in Shànghǎi is greater than that of Wúxī, which might be a particular psychological complexion existed in a metropolis like Shànghǎi and Běijīng. It reminds us that the dialect loss in the context of urbanization might be greater in the suburban areas than the urban area.

PTH

Whether a participant can speak PTH or not is expected to influence externally motivated variation like tone borrowing and contour loss. However, our data shows that PTH is not significant in any external ongoing changes. This is perhaps because the data of whether a participant can speak PTH or not was obtained by self-reports, which might influence the validity of data. It has been noted that women tend to over-report the use of the standard language (Trudgill, 2000: 61-80).

MORPHEME

MORPHEME is an external factor for predicting the variation of citation tone. It has two conditions: bound morpheme and free morpheme. MORPHEME is found to be significant in predicting the following variables:

(1) *WXfalling*: the falling tokens borrowed from PT4 are more likely to occur in bound morphemes than in free morphemes.

(2) *WXT4*: bound morphemes are more likely to use advanced variants, i.e., rising tone or dipping tone.

(3) *T4/6* in SJD and BSD: the majority of T4/6 morphemes in SJD and BSD are rising tone. However, a small number of peaking tokens, mainly used by T2 morphemes, are also found in T4/6. Though they are not ongoing changes, bound morphemes are significant for predicting the utterances of peaking tokens.

Generally speaking, bound morphemes in the isolated form are more likely to be involved in ongoing changes. As bound morphemes cannot occur on their own as independent words, speakers tend to borrow either PTH tone or a tone used in other MC tonal category to pronounce bound morphemes which appear alone.

STYLE (minimal pair)

Apart from word list reading and paragraph reading, another condition of STYLE of eliciting citation data is the reading of minimal pairs. Generally speaking, the task of reading minimal pairs in a sociolinguistic interview is used to test whether the speakers are able to distinguish the two pairs in production. However, how this task works in tonal variation research was unclear before conducting this study. Minimal pairs as a condition of STYLE is significant in predicting *WXlevel*: morphemes in the reading of minimal pairs are much less likely to be realized as level tones than in the reading of word lists.

The role of minimal pairs is also found in the variation of NT_{Pmin} in the dipping *WXT4*: young speakers have much higher NT_{Pmin} in the reading of minimal pairs and word lists than in paragraph reading. It seems that the reading of minimal pairs is the context for promoting tonal borrowings. The borrowing is more likely to occur in minimal pairs because in this speech style, two or three morphemes having exactly the same segment pronunciations appear at the same time. The speaking style of minimal pairs is known to increase speakers' linguistic insecurity and can cause hypercorrections. In this case, they are more likely to borrow PTH tone or the tone of other morphemes in the same pair if they cannot distinguish the tone contrast.

Now turning to non-linguistic factors, the conditioning effects of AGE, SEX and level of education (EDU) are summarized as follows:

AGE

AGE (old generation vs. young generation) is used as an indication of ongoing changes in tonal systems. In a likelihood ratio test, if a mixed model analysis with AGE as the only fixed factor and the random factors yields a better fit to the data than the model containing only random factors, then AGE is defined as a significant main effect in likelihood ratio test, which serves as a prerequisite for recognizing the tonal variable.

SEX

SEX differences are found to be responsible for the NT_{Pmax} of peaking WXT4, where it interacts with REGION. Urban females have much later peaks than males, so urban females are more advanced, while suburban females do not differ from suburban males significantly. SEX is also responsible for *DelayRising* of the Shànghǎi suburban dialects, where suburban (old) males are the conservative group.

According to the third principal of Labov's gender paradox (Labov, 2001: 292), the conditioning effects of SEX suggest *DelayRising* and the NT_{Pmax} of peaking *WXT4* are all changes below the level of awareness. As argued above, they are both internally motivated. As stated in Chapter 3, SEX is investigated in this study to see if it is a social constraint in Chinese speech community. Our data shows that SEX is a significant social constraint in the speech communities of Wúxī and Shànghǎi. Since SEX is responsible for the NT_{Pmax} of peaking WXT4 and *DelayRising*, it can be also concluded that SEX mainly plays a role in the internally motivated variations in Wúxī and Shànghǎi.

EDU

EDU is not a significant explanatory factor in any ongoing changes identified in this study. EDU is intrinsically correlated with AGE in this study because old people did not have the opportunity to receive formal or good education. Thus the majority of old participants had a low level of EDU, while all the young participants received compulsory education for at least nine years. EDU turns out to be significant in many likelihood ratio tests but not significant when being tested separately in the data of old people and young people. This suggests that the significant effects of EDU depend on the effects of AGE in our dataset. The further sociolinguistic research should also take into account the relationships between AGE and EDU in China.

7.5. Conclusion and further research

Sections 7.1 to 7.4 provided a summary of the findings reported in the preceding chapters and have returned to the two methodological issues and the two general goals proposed in Chapter 1. To place these findings in a broader perspective, the final section of this dissertation will first disscuss the methodological and theoretical implications for the fields of Chinese dialectology, "Urban Language Survey" (ULS) and sociolinguistics and offer several prospects for future research.

7.5.1. Methodological and theoretical implications

First, this study identified four types of variation in citation tones in the speech communities of Wúxī and Shànghǎi. These are (i) the merger of lax tones in the Yáng register, (ii) contour loss, (iii) *DelayRising* and (iv) the borrowing of PTH tone. Two types of variation in sandhi tone were identified: (i) contour loss in sandhi patterns and (ii) the borrowing of the /55.31/ pattern from the Yīn register.

Among these types of variation, the merger of lax tones in the Yáng register and tone borrowings had been discussed before in Chinese dialectology, whereas contour loss in citation tones, tone sandhi and *DelayRising* are new findings emerging from this study. As discussed above, contour loss and *DelayRising* are both phonetically gradual change in progress, which are difficult to detect using only auditory judgements. The fine-grained acoustic analyses and comparisons adopted by this study can assist future researchers to distinguish those subtle differences in phonetic realization and to further identify these types of variation.

Regarding the merger of lax tones in the Yáng register, especially *WXT4*, this study has provided a model for interpreting the findings of previous dialectology studies from the perspective of language variation and change. For example, various transcriptions of *WXT4* in previous dialectology studies can now be interpreted as having captured different steps in an ongoing chain change. Such findings in dialectology studies are the base for formulating our hypothesis that *WXT4* involves four steps in changing from a peaking variant to a dipping variant. Furthermore, this study has also benefitted from the standard research tool of Chinese dialectology –

the Dialect Survey Word List – in order to trace the MC tonal category of a morpheme before designing the word lists for eliciting citation tones and tone sandhi accordingly. These methodological contributions from the current study suggest that the methods and findings of previous Chinese dialectology should be taken into account in the sociophonetic studies on Chinese dialects. Conversely, the previous findings of Chinese dialectology could gain new implications when being interprated in the framework of sociolingusitics. Chinese dialectologists and sociophoneticians need to collaborate in studying the linguistic consequences of urbanization in China.

Second, after the identification of tonal variables, the present study used the methods of likelihood ratio testing and mixed modelling to identify the significant factors for each tonal variable. The merger of lax tones in the Yáng register and *DelayRising* were found to represent internally triggered variation while contour loss and tone borrowings in both citation tone and tone sandhi were found to be represent externally triggered variation. Suburban dialects are more likely to show a convergence towards urban dialects in the internally triggered variation while showing divergence in the externally triggered variation. In other words, the suburban dialects in the Wúxī and Shànghǎi areas are undergoing bidirectional variation towards both urban variety and standard Mandarin - PTH.

As summarized in Section 7.2, the merger of lax tones in the Yáng register as found in both suburban Wúxī (*WXT4*) and Shànghǎi (*DelayRising* in NHD) suggest that suburban areas are converging towards an urban variety. However, suburban young speakers in Shànghǎi are also changing at a faster pace in discarding the tonal features of their own dialects than urban young speakers, in particular in terms of the contour loss of both citation tone and tone sandhi (Section 7.2.2 and Section 7.3.1). Similarly, in Wúxī, suburban young speakers borrow more PTH falling tones (PT4) and neutral tones than urban young speakers (Section 7.2.4). These developments suggest that suburban dialects are less resistant to PTH influence than urban dialects because the suburban young generation is less determined to preserve their dialect than the urban young generation. Based on the questionnaire (Section 6.2.3), it is possible to hypothesize that suburban speakers will be undergoing psychological accommodation in the context of urbanization, if they cannot build their new identity as members of the urban community, they tend to abandon their own dialects in order to weaken their suburban identity and to build a new and broader identity as being “Chinese” by using PTH.

The finding of “bidirectional variation” undergoing in Chinese suburban areas is not only descriptive but also carry theoretical implications for the “Urban Language Survey” (ULS). The current study was designed in the framework of ULS, but not solely focused on the urban dialects. As reviewed in Chapter 1, ULS originated in China in the context of large-scale urbanization in order to examine the linguistic consequences of widespread urbanization. It adopts Labovian theory and consequently focuses on linguistic realities in urban areas. However, in order to trace the impact of urbanization on language change and variation, the linguistic realities of suburban areas are also crucial. The urban and suburban speech communities need to be investigated as coherent units in order to answer what urbanization brings to the Chinese dialects.

When ULS expands to surbuban areas, speech samples that are more representive demographically should be included. Once this is achieved, we have to deal with data sets that are hundreds of times larger than before. Yet “big data” is no longer an obstacle for expanding the research scope in ULS. As shown by this study, present day methods of acoustic analysis not only provide us with quantitative data but also with a (semi-) automatic processing of data accompanying technological innovations.

In addition to its methodological and theoretical implications for Chinese dialectology and ULS, this study has also contributed towards the variationist study of lexical tone.

This study revealed the constraining effects of three internal factors: STYLE (paragraph reading & word list reading), MC_TONE and MC_AB; five external factors: LEXICAL ATTESTEDNESS, PTH_TONE, REGION, MORPHEME and STYLE (minimal pair); as well as the two non-linguistic factors of AGE and SEX in tonal variation. Tonal variation is a rather underexplored area in the entire field of variationist research and is still in need of a stock of standard exploratory factors for the analysis of tonal variation. Research results from other fields such as variationist studies (e.g., SEX, EDU, STYLE, and REGION), phonetics (vowel HEIGHT) and Chinese dialectology (e.g., MC_TONE, MC_AB, PTH_TONE) were first used to propose the most probable exploratory factors to impact tonal variation before being tested in our data. This method of adopting concepts and findings from other fields is the necessary road to explore a new area – tonal variation and change.

Besides SEX, the other significant factors mentioned include several not discussed before in variationist studies, i.e., MC_TONE, MC_AB, PTH_TONE, LEXICAL ATTESTEDNESS and MORPHEME, REGION (urban vs suburban). These factors are suggested to be further examined in future research. Moreover, STYLE, a frequently discussed exploratory factor in variationist studies, was re-interpreted with a new sociolinguistic meaning.

The variationist treatment of the factor STYLE commonly correlates it with a pattern that people tend to use higher prestige variants more often in more formal styles while using lower prestige variants more often in informal styles (Tagliamonte 2011: 34). The underlying assumption is that the continuum of speech styles (from formal to casual) corresponds to a single linguistic variable. However, in this study, STYLE is both explained as an internal and external factor in tonal variation. The continuum of speech styles (paragraph reading - word list reading - minimal pairs reading) is split into two parts, where each part corresponds to a single linguistic variable. Internally, STYLE (paragraph reading & word list reading) is used as an indicator of articulatory effort, which was found to be significant in constraining contour loss. Contour loss not only reflects speaker's ease of articulation but also gains social meaning when being constrained by a non-linguistic variable – REGION. The social force behind contour loss comes from speakers' attitude to hide their dialect accent. Externally, STYLE (minimal pairs) can predict the borrowing of PTH tones when speakers cannot distinguish tonal differences of morphemes in the same pair. This study proposes that the continuum of speech styles might break at some point and exert different impacts on the linguistic variations.

7.5.2. Future direction

Before closing this study, I would like to offer some suggestions for further tonal variationist research.

The relationship between the citation tone and tone sandhi has been extensively studied but due to conflicting results it is still a matter of debate. Citation tones have been generally taken as the basic or underlying tones while sandhi tones are considered derived or surface tones. For example, the interpretation of "right-dominant" or "left-dominant" takes this citation-determined view as an assumption. However, some scholars suggest treating the sandhi tone as "original" or "underlying" in the synchronic analysis or historical reconstruction of some dialects (Hashimoto, 1982; Ting, 1982), because the tonal shapes of citation tones can be influenced by its sandhi tone due to reanalysis, misperception and historical development (Z. Zhang, 1998). Both independence and interdependence were found in synchronic studies, while diachronic evidence pointed out that their interaction played a role in promoting changes of both citation tone and tone sandhi. Since tonal variation has already been identified in both the citation and sandhi systems, it is worthwhile further investigating the interplay of citation tone and sandhi tone from the dynamic perspective.

The findings of this study are based on production data and the author's transcription. This study lacks a perceptual component, however. If further research can include perceptual evidence from native Wú-speakers, some issues such as speaker's awareness, whether a merger is already completed or only near completed, and whether a token is an implicit borrowing of a PTH tone, may receive more precise and firm conclusions.

This study mainly focused on internal and external factors but involved relatively few social-psychological factors, e.g., language attitude and language identity. The interpretations relevant to language attitude and language identity are mostly based on the author's living experiences in those speech communities and the observations in interviewing participants. If further research can incorporate more social-psychological factors, then arguments of "linguistic insecurity" and "urban pride" based on quantitative analyses will be available in addition to qualitative analysis.

This study focused on the F_0 variation as it is the most important acoustic correlate of tone. It would be interesting to incorporate the variations in the other acoustic correlate of tone, e.g., duration, intensity and phonation types. This would give more support to the findings in this study.

Linguistically naïve native speakers are less aware of tones than of vowels and consonants (Yip, 2002: 1). However, Chinese speakers who master the Pīnyīn system are able to name the PTH tones of basic Chinese characters, and hence, they are a group of speakers with a higher awareness of PTH tone than those who do not know Pīnyīn. This means that proficiency in the Pīnyīn system could be an explanatory factor for the tonal borrowing in the reading tasks. This could be investigated further.

This study is the first attempt to capture the process of rapid urbanization and language change with reference to lexical tones. This study's relevance for the field of Chinese dialectology is that it has reported new findings in dialect variation and new interpretations of previous findings. It further suggested that Chinese dialectologists and sociophoneticians should cooperate in studying the linguistic consequences of urbanization in China. Relevant for the field of ULS, this study has provided new insights into the interplay of urban and suburban dialects and it has proposed ULS to adopt new methods of acoustic analysis. Relevant for the field of variationist sociolinguistics, this study has modeled acoustic analysis of tonal variations and provided new insights into the analysis of complex tonal systems and sampling in large Chinese cities. Importantly, this study has proposed several exploratory factors for studying tonal variations. After all, as stated above, new questions have emerged from those insights. Further research is really needed on the topic of tonal variation and change.

References

Adank, Patricia Martine. 2003. Vowel Normalization: A Perceptual-Acoustic Study of Dutch Vowels. Radboud University.

Adank, Patricia Martine, Smits, Roel & Van Hout, Roeland. 2004. A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America.* Volume 116, pp. 3099.

Amara, Muhammad. 2005. Language, migration, and urbanization: the case of Bethlehem. *Linguistics.* Volume 43, Issue 5, pp. 883-901.

Bailey, Guy, Wikle, Tom, Tillery, Jan & Sand, Lori. 1993. Some patterns of linguistic diffusion. *Language Variation and Change.* Volume 5, Issue 3, pp. 359-390.

Bailey, Guy, Wikle, Tom, Tillery, Jan & Sand, Lori. 1996. The linguistic consequences of catastrophic events. *Sociolinguistic Variation: Data, Theory, and Analysis: Selected Papers from NWAV.* Volume 23, pp. 435-451.

Bailey, Guy. 2004. Real and apparent time. *The Handbook of Language Variation and Change*, pp. 312-332. Blackwell Publishing Ltd.

Bào, Míngwěi (鲍明伟). 1980. Liùshínián lái Nánjīng fāngyīn xiàng Pŭtōnghuà kàolŏng qíngkuàng de kǎochá [A study of phonological developments of Nanjing dialect towards Putonghua in the last 60 years, 六十年来南京方音向普通话靠拢情况的考察]. *Zhōngguó Yŭwén [Chinese Language and Writing, 《中国语文》].* Volume 157, pp. 241-245.

Bates, Douglas, Maechler, Martin, Bolker, Ben & Walker, Steven. 2013.*Lme4: Linear mixed-effects models using S4 classes* (1.0-5 ed.) http://cran.r-project.org/web/packages/lme4/index.html.

Blankenship, Barbara. 2002. The Timing of Nonmodal Phonation in Vowels. *Journal of Phonetics.* Volume 30, Issue 2, pp. 163-191.

Boersma, Paul. 2002. Praat, a system for doing phonetics by computer. *Glot International.* Volume 5, Issue 9/10, pp. 341-345.

Bortoni-Ricardo, Stella Maris. 1985. *The Urbanization of Rural Dialect Speakers: A Sociolinguistic Study in Brazil*. London: Cambridge Studies in Linguistics.

Britain, David. 2004. Space and Spatial Diffusion. In Chambers, J., Trudgill, P. & Schilling-Estes, N. (Eds.), *The Handbook of Language Variation and Change*, pp. 603-637. Blackwell Publishing Ltd.

Bucholtz, Mary. 1999. “Why be normal?”: Language and identity practices in a community of nerd girls. *Language in Society.* Volume 28, Issue 2, pp. 203-223.

Cadora, Frederic J. 1970. Some Linguistic Concomitants of Contactual Factors of Urbanization. *Anthropological Linguistics.* Volume 12, Issue 1, pp. 10-19.

Cao, Jianfen & Maddieson, Ian. 1992. An Exploration of Phonation Types in Wu Dialects of Chinese. *Journal of Phonetics.* Volume 20, Issue 1, pp. 77-92.

Cáo, Xiǎoyàn (曹晓燕). 2003. Wúxī Fāngyán Liǎngzìzǔ Liándú Biàndiào [Wúxī tone sandhi in bisyllable words, 无锡方言两字组连读变调]. In Shànghǎishì Yǔwén Xuéhuì &Xiānggǎng Zhōngguó Yǔwén Xuéhuì Hébiān [The Chinese language society of Shànghǎi and The Chinese language society of Hong Kong,上海市语文学会，香港中国语文学会合编] (Eds.), *Wúyǔ Yánjiū [Studies on the Wú dialects, 《吴语研究》]* Shànghǎi Jiàoyù Chūbǎnshè [Shànghǎi Education Press，上海教育出版社].

Cedergren, Henrietta J. & Sankoff, David. 1974. Variable rules: Performance as a statistical reflection of competence. *Language.* pp. 333-355.

Chambers, J. K. 2003. *Sociolinguistic Theory: Linguistic Variation and its Social Significance* (second edition ed.). Oxford, Malden: Blackwell publishers Ltd.

Chan, Marjorie K. M. & Ren, Hongmo. 1989. Wuxi tone sandhi from last to first syllable dominance. *Acta Linguistica Hafniensia.* Volume 21, Issue 2, pp. 35-64.

Chang, Kun. 1975. Tonal developments among Chinese dialects. *Zhōngyāng Yánjiūyuàn Lìshǐ Yǔyán Yánjiūsuǒ Jíkān [Bulletin of the Institute of History and Philology, Academia Sinica, 《中央研究院历史语言研究所集刊》].* Volume 46, Issue 4, pp. 636-709.

Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Chao, Yuen Ren. 1967. Contrastive aspects of the Wu dialects. *Language.* Volume 43, Issue 1, pp. 92-101.

Chatterjee, Indranil, Halder, Hindol, Bari, Sayani, Kumar, Suman & Roychoushury, Amitabha. 2011. An Analytical Study of Age and Gender Effects on Voice Range Profile in Bengali Adult Speakers using Phonetogram. *International Journal of Phonosurgery and Laryngology.* Volume 2, Issue 1, pp. 65-70.

Chen, Matthew Y. 2000. *Tone Sandhi: Patterns Across Chinese Dialects*. Cambridge: Cambridge University Press.

Chén, Qíshēng (陈祺生). 1989. Wúxī huà de Liándú Biàndiào Jīqí zài Yǔfǎ shàng de Zuòyòng [Wúxī tone sandhi and its role in grammar, 无锡话的连读变调及其在语法上的作用]. In Jiāngsū shěng Yǔyán Xuéhuì [The linguistic society of Jiangsu, 江苏省语言学会] (Ed.), *Yǔyán Yánjiū Jíkān (Dìèrjí) [Collection of Linguistic Research Papers no. 2, 语言研究集刊(第二辑)]* Jiāngsū Jiàoyù Chūbǎnshè [Jiāngsū Education Press, 江苏教育出版社].

Chén, Yuán (陈原). 1980. *Yǔyán Yǔ Shēnghuó [Language and Social Life, 《语言生活》]*. Shanghai: Shànghǎi Xuélín Chūbǎnshè [Shànghǎi Xuélín Press, 上海学林出版社].

Chén, Yuán (陈原). 1983. *Shèhuì Yǔyánxué [Sociolinguistics, 《社会语言学》]*. Beijing: Sānlián Chūbǎnshè [Sānlián Press, 三联出版社].

Chen, Zhongmin. 2000. Studies on Dialects in the Shanghai Area: Their Phonological System and Historical Development. PhD dissertation. University of California, Berkeley.

Chén, Zhōngmǐn (陈忠敏). 2007. Shànghǎi Shìqūhuà Shūshēng Yángdiàolèi Hébìng de Yuányīn[On the cause of the merger of the Yangping, Yangshang and Yangqu tones in the Shanghai dialect, 上海市区话舒声阳调类合并的原因]. *Fāngyán [Dialectology, 《方言》]*. Volume 4, pp. 305-310.

Chén, Zhōngmǐn (陈忠敏). 2010. Wúyǔ Qīngyīn Zhuóliú de Shēnxué Tèzhēng jí Jiàndìng Biāozhì: Yǐ Shànghǎi Huà Wéi Lì [An Acoustic study of voiceless onset followed by breathiness of Wu dialects: Based on the Shànghǎi dialect, 吴语清音浊流的声学特征及鉴定标志——以上海话为例]. *Yǔyán Yánjiū [Studies in Language and Linguistics, 语言研究]*. Volume 30, Issue 3, pp. 20-34.

Chéng, Xù (程序). 2006. Bóyáng fāngyán de yǔyīn biànyì yánjiū [A Study on the Age Variation of Pronunciation in Poyang Dialect, 《鄱阳方言的语音变异研究》]. *Zhōngguó Shèhuì Yǔyánxué [Journal of Chinese Sociolinguistics, 《中国社会语言学》]*. Volume 2, pp. 101-110.

Dialect survey group of Jiāngsū Province and Shànghǎi (江苏省上海市方言调查指导组). 1960. *Jiāngsū Shěng Hé Shànghǎi Shì Fāngyán Gàikuàng [the General Situation of the Dialects in Jiāngsū Province and Shànghǎi 《江苏省和上海市方言概况》]*. Nánjīng: Jiāngsū Rénmíng Chūbǎnshè [Jiāngsū People's Publishing House, 江苏人民出版社].

Dīng, Shēngshù (丁声树) & Lǐ, Róng (李荣). 1966. *Gǔ-Jīn Zìyīn Duìzhào Shǒucè [Handbook of Comparison of Modern and Ancient Sounds of Characters,*

《古今字音对照手册》]. Beijing: Kēxué Chūbǎnshè [Science Press, 科学出版社].

Disner, Sandra Ferrari. 1980. Evaluation of Vowel Normalization Procedures. *The Journal of the Acoustical Society of America.* Volume 67, pp. 253-261.

Durian, David. 2007. Getting [ʃ] stronger Every Day?: More on Urbanization and the Socio-geographic Diffusion of (str) in Columbus, OH. *University of Pennsylvania Working Papers in Linguistics.* Volume 13, Issue 2, pp. 65-78.

Eckert, Penelope. 2000. *Language Variation as Social Practice: The Linguistic Construction of Identity in Belten High*. Oxford: Blackwell.

Eckert, Penelope. 2008. Variation and the Indexical Field. *Journal of Sociolinguistics.* Volume 12, Issue 4, pp. 453-476.

Eckert, Penelope. 2012. Three waves of variation study: the emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology.* Volume 41, pp. 87-100.

Esposito, Christina Marie. 2006. The Effects of Linguistic Experience on the Perception of Phonation. PhD dissertation. University of California.

Esposito, Christina Marie. 2010. The effects of linguistic experience on the perception of phonation. *Journal of Phonetics.* Volume 38, Issue 2, pp. 306-316.

Fought, Carmen. 1999. A Majority Sound Change in a Minority Community: /u/ - fronting in Chicano English. *Journal of Sociolinguistics.* Volume 3, Issue 1, pp. 5-23.

Fù, Yìróng (付义荣). 2011. *Yányǔ Shèqū Hé Yǔyán biànhuàYánjiū: Jīyú ĀnhuīFùcūn De Shèhuì Yǔyánxué Diàochá[Studies of the Speech Community and Language Change: A Sociolinguistic Investigation of the Fu Village in Anhui, 《言语社区和语言变化研究：基于安徽付村的社会语言学调查》]*. Beijing: Běijīng Dàxué Chūbǎnshè [Peking University Press,北京大学出版社].

Garellek, Marc & Keating, Patricia. 2011. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association.* Volume 41, Issue 02, pp. 185-205.

Gě, Yànhóng (葛燕红). 2005. Nánjīng shì 'Xiáojié' Chēnghūyǔ de Diàochá Fēnxī [A Survey on the Appellation 'Xiaojie' in Nanjing, 南京市“小姐”称呼语的调查分析]. *Zhōngguó Shèhuì Yǔyánxué [Journal of Chinese Sociolinguistics, 《中国社会语言学》].* Volume 2, pp. 196-206.

Glasberg, Brian R. & Moore, Brian CJ. 1990. Derivation of auditory filter shapes from notched-noise data. *Hearing Research.* Volume 47, Issue 1, pp. 103-138.

Goldsmith, J. A. 1976. *Autosegmental Phonology* Citeseer.

Hashimoto, Mantaro 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.* Volume 53, Issue 4, pp. 645-659.

Hè, Fúlíng (贺福凌). 2012. Cóng Hànyǔ Fāngyán Kàn “Niàn”"Sà""Xì" [On the Pronunciations of “Twenty”, “Thirty” and “Forty” in Chinese Dialects, 从汉语方言看“廿”"卅"“卌”的读音问题]. *Zhōngguó Yǔwén Tōngxùn [Newsletter of Chinese Language, 《中国语文通讯》].* Volume 91, pp. 95-106.

Hé, Huìmíng (何慧明) (Ed.). 1991. *Shànghǎishì Sōngjiāng Xiànzhì [Shànghǎishì Sōngjiāng County Annals, 《上海市松江县志》].* Shànghǎi Rénmíng Chūbǎnshè [Shànghǎi People's Publishing House, 上海人民出版社].

Ho, Dah-an (何大安). 1988. "Zhuó Shǎng Guī Qù" yǔ Xiàndài Fāngyán ["Zhuó Shǎng Guī Qù" and the modern dialects, “浊上归去”与现代方言]. *Zhōngyāng Yánjiūyuàn Lìshǐ Yǔyán Yánjiūsuǒ Jíkān [Bulletin of the Institute of History and Philology, Academia Sinica, 《中央研究院历史语言研究所集刊》].* Volume 59, Issue 1, pp. 115-140.

Holmberg, Eva B., Hillman, Robert E., Perkell, Joseph S., Guiod, Peter C. & Goldman, Susan L. 1995. Comparisons among Aerodynamic, Electroglottographic, and Acoustic Spectral Measures of Female Voice. *Journal of Speech, Language and Hearing Research.* Volume 38, Issue 6, pp. 1212-1223.

Hombert, Jean-Marie, Ohala, John J. & Ewan, William G. 1979. Phonetic Explanations for the Development of Tones. *Language.* Volume 55, Issue 1, pp. pp. 37-58.

Horvath, Barbara M. & Horvath, Ronald J. 2001. A multilocality study of a sound change in progress: The case of /l/ vocalization in New Zealand and Australian English. *Language Variation and Change.* Volume 13, Issue 01, pp. 37.

Hóu, Jīngyī (侯精一) (Ed.). 1999. *Xiàndài Hànyǔ Fāngyán Yīnkù [Phonetic Database of Modern Chinese Dialects, 《现代汉语方言音库》].* Shànghǎi Jiàoyù Chūbǎnshè [Shànghǎi Education Press，上海教育出版社].

Howard, David M. 1997. Practical Voice Measurement. In Harris T., Harris S., Rubin J.S. & Howard D.M. (Eds.), *The Voice Clinic Handbook*, pp. 323-382. London: Whurr Publishing Company.

Hugo, Q. 2012. *Statistics with R: handout of tutorial presented at the workshop on Experimental Methods in Language Acquisition Research (EMLAR VIII)*http://www.hugoquene.nl/emlar/Rtutorialemlar8.pdf.

Institute of Linguistics, CASS (Ed.). 2002. *Xiàndài Hànyǔ Cídiǎn [the Contemporary Chinese Dictionary, 《现代汉语词典》]*. 5th ed. Beijing: Shāngwù Yìnshūguǎn [Commercial Press, 商务印书馆].

International Phonetic Association. 1999. *Handbook of the International Phonetic Association: A Guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Jannedy, Stefanie & Hay, Jennifer. 2006. Modelling Sociophonetic Variation. *Journal of Phonetics.* Volume 34, Issue 4, pp. 405-408.

Johnson, Daniel Ezra. 2009. Getting off the GoldVarb Standard: Introducing Rbrul for Mixed‐Effects Variable Rule Analysis. *Language and Linguistics Compass.* Volume 3, Issue 1, pp. 359-383.

Jones, Mari C. & Esch, Edith. 2002. *Language Change: The Interplay of Internal, External and Extra-Linguistic Factors* Walter de Gruyter.

Kreiman, Jody, Gerratt, Bruce R. & Antonanzas-Barroso, Norma. 2007. Measures of the Glottal Source Spectrum. *Journal of Speech, Language and Hearing Research.* Volume 50, Issue 3, pp. 595-610.

Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language.* pp. 715-762.

Labov, William. 1975. On the use of the present to explain the past. *Proceedings of the 11th International Congress of Linguists,* Volume 2. pp. 825-851.

Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change.* Volume 2, Issue 2, pp. 205-254.

Labov, William. 1994. *Principles of Linguistic Change (I): Internal Factors*. Massachusetts: Blackwell.

Labov, William. 2001. *Principles of Linguistic Change (II): Social Factors*. Massachusetts: Blackwell.

Ladefoged, Peter. 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Malden, MA: Blackwell.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.

Lewis, M. Paul (Ed.). 2009. *Ethnologue: Languages of the World.* 16th ed. Dallas, Tex: SIL International.

Lǐ, Chūnlín (李春林). 2005. *Duànliè Yǔ Suìpiàn: Dāngdài Zhōngguó Shèhuì Jiēcéng Fēnhuà Shízhèng Fēnxī (Cleavage Or Fragment: A Empirical Analysis on the Social Stratification of the Contemporary China, 断裂与碎片：当代中国社会阶层分化实证分析)*. Beijing: Shèhuì Kēxué Wénxiàn Chūbǎnshè (Social Science Academic Press, 社会科学文献出版社).

Li, Chunling. 2009. *Profile of Middle Class in Mainland China.* Retrieved 12/13, 2012, from http://e-sociology.cass.cn/pub/pws/lichunling/grwj_lichunling/P020090525597135469507.pdf.

Lǐ, Róng (李荣). 1966. Wēnlǐng Fāngyán Yǔyīn Fēnxī [An acoustic study on the Wēnlǐng dialect, 温岭方言语音分析]. *Zhōngguó Yǔwén [Chinese Language and Writing, 《中国语文》].* Issue 1, pp. 1-9.

Lǐ, Rónggāng (李荣刚). 2012. *Chéngshìhuà yǔ Xiāngcūn Yányǔ Shèqū de Biànhuà [Urbanization and Rural Speech Community Changes, 《城市化与乡村言语社区的变化：连岛社区调查》].* Unpublished PhD dissertation, Nanjing Univerisity.

Liú, Língzhū (刘灵珠). 2005. Zuò / zuò biànyì yánjiū [The Variation between "Zuo (作)" and "Zuo (做)", 作/做变异研究]. *Zhōngguó Shèhuì Yǔyánxué [Journal of Chinese Sociolinguistics, 《中国社会语言学》].* Volume 2, pp. 214-224.

Lu, Hanchao. 2004. *Beyond the Neon Lights: Everyday Shanghai in the Early Twentieth Century*. Berkeley and Los Angeles; London: University of California Press.

Makoni, Sinfree, Brutt-Griffler, Janina & Mashiri, Pedzisai. 2007. The use of "indigenous" and urban vernaculars in Zimbabwe. *Language in Society.* Volume 36, Issue 1, pp. 25-49.

Mei, Tsu-lin. 1970. Tones and Prosody in Middle Chinese and The Origin of The Rising Tone. *Harvard Journal of Asiatic Studies.* Volume 30, pp. pp. 86-110.

Mille, Catherine, Al-Wer, Enam, Caubet, Dominique & Watson, Janet C. E. 2007. *Arabic in the City: Isses in Dialect Contact and Language Variation*. New York: Routledge.

Milroy, James. 1981. *Regional Accents of English: Belfast*. Belfast: Blackstaff Press.

Moore, Brian CJ & Glasberg, Brian R. 1983. Suggested formulae for calculating auditory‐filter bandwidths and excitation patterns. *The Journal of the Acoustical Society of America.* Volume 74, pp. 750.

Nolan, Francis. 2003. Intonational equivalence: an experimental evaluation of pitch scales. *Proceedings of the 15th International Congress of Phonetic Sciences,* Volume 771. pp. 774.

Nordberg, Bengt. 1994. *The Sociolinguistics of Urbanization: The Case of the Nordic Countries* de Gruyter.

Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge University Press.

Norman, Jerry. 2003. The Chinese dialects: phonology. In Thurgood, G. &LaPolla, R. J. (Eds.), *The Sino-Tibetan Languages*, pp. 72-83. London: Routledge.

Pān, Wùyún (潘悟云). 1982. Guānyú Hànyǔ Shēngdiào Fāzhǎn de Jǐgè Wèntí (Several problems in the development of Chinese tones, 关于汉语声调发展的几个问题). *Journal of Chinese Linguistics.* Volume 10, Issue 2, pp. 359-385.

Paolillo, John C. 2002. *Analyzing Linguistic Variation: Statistical Models and Methods*. Standford: Center for the Study of Language and Information.

Podesva, Robert J. 2007. Phonation type as a stylistic variable: The use of falsetto in constructing a persona1. *Journal of Sociolinguistics.* Volume 11, Issue 4, pp. 478-504.

Pulleyblank, Edwin G. 1978. The nature of the Middle Chinese tones and their development to Early Mandarin. *Journal of Chinese Linguistics.* Volume 6, Issue 2, pp. 173-203.

Qian, Nairong. 2010. The Spread of Shanghainese to Nanqiao Fengxian District. In Van den Berg, M. &Xu, D. (Eds.), *Urbanization and the Restructuring of Speech Communities in China and Europe*, pp. 184-190. Newcastle upon Tyne: Cambridge Scholars Publishing.

Qián, Nǎiróng (钱乃荣). 1992. *Dāngdài Wúyǔ Yánjiū [Studies in the Contemporary Wú Dialects, 《当代吴语研究》]*. Shanghai: Shànghǎi Jiàoyù Chūbǎnshè [Shànghǎi Education Press，上海教育出版社].

Qián, Nǎiróng (钱乃荣). 2006. SOV wánchéngtǐ jù hé SVO wánchéngtǐ jù zài Wǔyǔ zhōng de jiēchù jiéguǒ [The Result of the Contact of SOV and SVO with Perfective Aspect in Wu Dialect, SOV 完成体句和 SVO 完成体句在吴语中的接触结果]. *Zhōngguó Shèhuì Yǔyánxué [Journal of Chinese Sociolinguistics, 《中国社会语言学》]*. Volume 2, pp. 94-100.

R Core Team. 2013.*A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rand, David & Sankoff, David. 1990. *GoldVarb Version 2: A Variable Rule Application for the Macintosh* Université de Montréal.

Reaser, Jeffrey. 2004. A quantitative sociolinguistic analysis of Bahamian copula absence: Morphosyntactic evidence from Abaco Island, The Bahamas. *Journal of Pidgin and Creole Languages.* Volume 19, Issue 1, pp. 1-40.

Ren, Nianqi. 1992. *Phonation Types and Consonant Descriptions: Shanghai Chinese.* Unpublished PhD dissertation, The University of Connecticut.

Rietveld, Toni & Van Hout, Roeland. 1993. *Statistical Techniques for the Study of Language and Language Behaviour* Walter de Gruyter.

Robinson, John, Lawrence, Helen & Tagliamonte, Sali. 2001. Goldvarb 2001. *Department of Language and Linguistic Science. University of York.*

Rose, Phil. 1987. Considerations in the normalisation of the fundamental frequency of linguistic tone. *Speech Communication.* Volume 6, Issue 4, pp. 343-352.

Sagart, Laurent. 1998. The origin of Chinese tones. *International Symposium on "Tone Languages in the World",* Tokyo.

Sankoff, David, Tagliamonte, Sali & Smith, Eric. 2005. Goldvarb X: A variable rule application for Macintosh and Windows. *Department of Linguistics, University of Toronto.*

Sankoff, David, Tagliamonte, Sali & Smith, Eric. 2012. Goldvarb LION: A variable rule application for Macintosh. *University of Toronto. URL Http://Individual.Utoronto.Ca/Tagliamonte/Goldvarb.Htm.*

Shi, Feng, Shi, Lin & Liao, Rongrong. 1987. An experimental analysis of the five level tones of Gaoba Dong language. *Journal of Chinese Linguistics.* Volume 15, Issue 2, pp. 335-360.

Stanford, James. 2008. A sociotonetic analysis of Sui dialect contact. *Language Variation and Change.* Volume 20, Issue 3, pp. 409-450.

State Language Commission, China (Ed.). 1988. *Xiàndài Hànyǔ Chángyòng Zìbiǎo (Modern Chinese Characters used Frequently Book, 《现代汉语常用字表》).* Beijing: Yǔwén Chūbǎnshè [Language Press, 语文出版社].

Tacoli, Cecilia. 1998. Rural-urban interactions; a guide to the literature. *Environment and Urbanization.* Volume 10, pp. 147-166.

Tagliamonte, Sali. 2006. *Analysing Sociolinguistic Variation*. New York: Cambridge University Press.

Tagliamonte, Sali. 2011. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Oxford: Wiley-Blackwell.

Thomas, Erik R. 2002. Instrumental Phonetics. In Chambers, J., Trudgill, P. & Schilling-Estes, N. (Eds.), *The Handbook of Language Variation and Change*, pp. 168-200. Oxford, UK/Malden, MA: Blackwell.

Thomas, Erik R. 1997. A rural/metropolitan split in the speech of Texas Anglos. *Language Variation and Change.* Volume 9, Issue 03, pp. 309-332.

Thomas, Erik R. 2011. *Sociophonetics: An Introduction*. Basingstoke, UK: Palgrave Macmillan.

Thomason, Sarah. 2010. Contact Explanations in Linguistics. *The Handbook of Language Contact*, pp. 29-47. Wiley-Blackwell.

Tillery, Jan & Bailey, Guy. 2003. Urbanization and the evolution of Southern American English. In Nagle, S. J. &Sanders, S. L. (Eds.), *English in the Southern United States*, pp. 159-172. Cambridge: Cambridge University Press.

Ting, Pang-Hsin. 1982. Some aspects of tonal development in Chinese dialects. *Bulletin of the Institute of History and Philology.* Volume 53, Issue 4, pp. 629-644.

Ting, Pang-hsin. 1996. Tonal evolution and tonal reconstruction in Chinese. In Huang, C. J. &Li, Y. A. (Eds.), *New Horizons in Chinese Linguistics*, pp. 141-159. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Trudgill, Peter. 1999. Dialect Contact, Dialectology and Sociolinguistics= Dialectos en contacto. dialectología y sociolingüística. *Cuadernos De Filología Inglesa.* Volume 8, Issue 1, pp. 1-8.

Trudgill, Peter. 2000. *Sociolinguistics: An Introduction to Language and Society* (Fourth edition). Harmondsworth: Penguin Group.

Tsou, Benjamin K. 2012. Aspects of Residual Bilingualism in Chinese Urban Settings. *10th Urban Language Seminar,* Utrecht.

van Bergem, Dick R. 1989. Phonetic and linguistic aspects of vowel reduction. *Proceedings of the Institute of Phonetic Sciences, Amsterdam.* Volume 13, pp. 97-105.

van den Berg, Marinus &Xu, Daming (Eds.). 2010. *Industrialization and Re-Structuring of Speech Commuities in China and Europe.* Newcastle: Cambridge Scholars.

van der Harst, Sander. 2011. The Vowel Space Paradox: A Sociophonetic Study on Dutch. LOT.

VanNess Simmons, Richard. 1999. *Chinese Dialect Classification: A Comparative Approach to Harngjou, Old Jintarn, and Common Northern Wu*. Amsterdam and Philadelphia: John Benjamins Pubulishing Company.

Wang, Guangtao (editor in chief). 2012. *The state of China's cities 2012/2013*. United States: UN-HABITAT.

Wang, William S-Y. 1982. A note on tone development. In The Chinese Language Society of Hong Kong (Ed.), *Wang Li Memorial Volumns (English Volume)*, pp. 435-443. Hong Kong: Joint Publishing Company.

Wáng, Yìzhī (王逸之). 2009. *Wúxī Fāngyán Yǔyīn Yánjiū [A phonetic study of the Wuxi dialect, 《无锡方言语音研究》].* Unpublished MA dissertation, Shanghai University. Shanghai.

Wang, William S-Y. 1967. Phonological Features of Tone. *International Journal of American Linguistics.* Volume 33, Issue 2, pp. pp. 93-105.

Weinreich, Uriel, Labov, William & Herzog, Marvin. 1968. Empirical foundations for a theory of language change. In Lehmann, W. &Malkiel, Y. (Eds.), *Directions for Historical Linguistics*, pp. 95-188. Austin: University of Texas Press.

Wolfram, Walt & Schilling-Estes, Natalie. 1998. *American English: Dialects and Variation*. Malden, MA: Blackwell.

Wurm, Stephen Adolphe, Li, Rong, Baumann, Theo & Lee, W. Mei (Eds.). 1988. *Language Atlas of China.* Hong Kong: Longman Group.

Xǔ, Bǎohuá (许宝华) &Tāng, Zhēnzhū (汤珍珠) (Eds.). 1988. *Shànghǎi Shìqū Fāngyánzhì [Dialects in Shànghǎi Metropolitan Area, 《上海市区方言志》].* Shanghai: Shànghǎi Jiàoyù Chūbǎnshè [Shanghai Education Press].

Xǔ, Bǎohuá (许宝华), Tāng, Zhēnzhū (汤珍珠) & Qián, Nǎiróng (钱乃荣). 1981. Xīnpài Shànghǎi Fāngyán de Liándú Biàndiào [Tone sandhi of new style Shànghǎi dialect, 新派上海方言的连读变调]. *Fāngyán [Dialectology, 《方言》].* Volume 2, pp. 145-155.

Xǔ, Bǎohuá (许宝华) & Táo, Huán (陶寰). 1997. Shànghǎi Fāngyán Cídiǎn [Shànghǎi Dialect Dictionary, 《上海方言词典》]. In Lǐ, R. (. (Ed.), *Xiàndài Hànyǔ Fāngyán Dacídiǎn [the Dictionary of Modern Chinese Dialects, 《现代汉语方言大词典》].* Nanjing: Jiāngsū Jiàoyù Chūbǎnshè [Jiāngsū Education Press, 江苏教育出版社].

Xu, Daming. 1992. A Sociolinguistic Study of Mandarin Nasal Variation. PhD dissertation. University of Ottawa.

Xu, Daming. 2006. Urban Language Survey. *Journal of Chinese Sociolinguistics.* Volume 2, pp. 1-15.

Xu, Daming. 2010. The Formation of A Speech Community: Mandarin Nasal Finals in Baotou. In van den Berg, M. &Xu, D. (Eds.), *Industrialization and the Restructuring of Speech Community in China and Europe*, pp. 120-140. Newcastle: Cambridge Scholars.

Xu, Daming. to appear. Gender Differences in Chinese Speech Communities. In Wang, W. S. &Sun, C. (Eds.), *The Oxford Handbook of Chinese Linguistics* Oxford University Press.

Xú, Jīnyì (徐金益). 2007. *Wúxī Fāngyán Shēngdiào Shíyàn Yánjiū [An acoustic experimental study of tones in Wuxi Dialect, 《无锡方言声调实验研究》].* Unpublished Master dissertation, Nanjing Normal University. Nanjing.

Xu, Yi. 1994. Production and perception of coarticulated tones. *The Journal of the Acoustical Society of America.* Volume 95, Issue 4, pp. 2240-2253.

Xu, Yi & Liu, Fang. 2012. Intrinsic coherence of prosodic and segmental aspects of speech. *Prosodies–Context, Function, Communication.* pp. 1-26.

Xue, Caide. 2010. A study of the langauge behaviour of Shanghai residents. In Van den Berg, M. &Xu, D. (Eds.), *Urbanization and the Restructuring of Speech Communities in China and Europe*, pp. 164-183. Newcastle upon Tyne: Cambridge Scholars Publishing.

Xuē, Zhèndōng (薛振东) (Ed.). 1992. *Shànghǎishì Nánhuì Xiànzhì [Shànghǎishì Nánhuì County Annals, 《上海市南汇县志》].* Shànghǎi Rénmíng Chūbǎnshè [Shànghǎi People's Publishing House, 上海人民出版社].

Yan, Margaret Mian. 2006. *Introduction to Chinese Dialectology*. Munich: Lincom Europa.

Yán, Yìmíng (颜逸明). 1995. Wúyǔ Fēnqū Lùnhéng [Arguments and Judgement on the classification of the Wú dialects, 吴语分区论衡]. In Xú, Y. (. (Ed.), *Wúyǔ Yánjiū [Studies of the Wú Dialects, 《吴语研究》]*, pp. 21-34. Xiānggǎng Zhōngwén Dàxué Xīnyà Shūyuàn [The Chinese University of Hong Kong New Asia College, 香港中文大学新亚书院].

Yáng, Yùguó (杨玉国). 2005. Guǎngxī Wǔmíng Guānhuà Rùshēng Zì Diàochá [An Investigation into the Entering Tones of Wuming Guanhua,广西武鸣官话入声字调查]. *Zhōngguó Shèhuì Yǔyánxué [Journal of Chinese Sociolinguistics, 《中国社会语言学》].* Volume 2, pp. 167-177.

Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.

You, Rujie. 2010. Language Competition in Contemporary Shanghai. In Van den Berg, M. &Xu, D. (Eds.), *Urbanization and the Restructuring of Speech Communities in China and Europe*, pp. 141-163. Newcastle upon Tyne: Cambridge Scholars Publishing.

Yóu, Rújié (游汝杰). 2010. Shànghǎi Jiāoqū Jìn 30 Niánlái de Biànhuà [Evolution of dialects in Shanghai suburban districts since 1980s, 上海郊区语音近 30 年来的变化] . *Fāngyán [Dialectology, 《方言》].* Volume 3, pp. 194-200.

Young, Richard & Bayley, Robert. 1996. VARBRUL analysis for second language acquisition research. In Baylay, R. &Preston, D. R. (Eds.), *Second Language Acquisition and Linguistic Variation*, pp. 253-306. Amsterdam: John Benjamins.

Zhān, Bóhuì (詹伯慧) (Ed.). 2001. *Hànyǔ Fāngyán ji Fāngyán Diàochá [汉语方言及方言调查, Chinese Dialects and Dialect Surveys].* Húběi Jiàoyù Chūbǎnshè [Hubei Education Press，湖北教育出版社].

Zhang, Jie. 2007. A directional asymmetry in Chinese tone sandhi systems. *Journal of East Asian Linguistics.* Volume 16, Issue 4, pp. 259-302.

Zhang, Jisheng. 2006. *The Phonology of Shaoxing Chinese* Netherlands Graduate School of Linguistics LOT, Utrecht.

Zhang, Qing. 2010. The Discursive Construction of the Social Stratification Order in Reforming China. *Journal of Language and Politics.* Volume 9, Issue 4, pp. 508-527.

Zhāng, Shuǐqīng (张水清) & Dù, Débīn (杜德斌). 2001. Shànghǎi Jiāoqū Chéngshìhuà Móshì Tàntǎo [Analysis on Suburbanization Model of Shanghai, 上海郊区城市化模式探讨]. *Dìyù Yánjiū Yǔ Kāifā [Areal Research and Development, 《地域研究与开发》].* Volume 20, Issue 4, pp. 22-26.

Zhang, ZS. 1998. The relationship between'citation'and'sandhi'tones in a modular analysis. *Bulletin of the Institute of History and Philology Academia Sinica.* Volume 69, pp. 181-203.

Zhèngzhāng, Shàngfāng (郑张尚方). 2003. *Shanggu Yinxi [Archaic Chinese Phonology,《上古音系》]* Shànghǎi Jiàoyù Chūbǎnshè [Shànghǎi Education Press, 上海教育出版社].

Zhou, Minglang. 2010. Sociolinguistic research in China. In Ball, M. J. (Ed.), *The Routledge Handbook of Sociolinguistics Around the World*, pp. 67-80. Oxon: Routledge.

Zhū, Bǎohé (朱保和) (Ed.). 1992. *Shànghǎishì Bǎoshān Xiànzhì [Shànghǎishì Bǎoshān County Annals, 《上海市宝山县志》].* Shànghǎi Rénmíng Chūbǎnshè [Shànghǎi People's Publishing House, 上海人民出版社].

Zhuāng, Yì (庄逸). 2011. Wúxīshì 10 nián zēngjiā rénkǒu 119 wàn [1.19 million population increase in Wúxī city in the last decade, 无锡市 10 年增加人口 119 万]. *Wúxī Rìbào [Wúxī Daily, 无锡日报],* Retrieved from http://wx.xinhuanet.com/2011-05/05/content_22688302.htm database.

Appendix I Word lists of monosyllabic morphemes

σ₂ / σ₁	[ɑ](Wúxī) [ʌ](Shànghǎi)	[ɛ]	[ʌ] (Wúxī) [ɔ] (Shànghǎi)	[u]	[y]	[i]	[ɿ]
T1A	街 kia [tɕiɛ^{55}] street	单 tén/[tan^{55}] single	遭 tsau [tsau^{55}] to suffer (misfortune)	孤 ku [ku^{55}] alone	居 kiu [tɕy^{55}] to reside	低 ti [ti^{55}] low	资 Tsy [tsɿ^{55}] resource 私 sy [sɿ^{55}] private
T1B	揩 k'é [k^{h}ai^{55}] to wipe	摊 t'én/[t^{h}an^{55}] to spread out	操 ts'au [ts^{h}au^{55}] gymnastic	枯 k'u [k^{h}u^{55}] dried up	区 k'iu [tɕ^{h} y^{55}] district	梯 t'i [t^{h}i^{55}] ladder	雌 ts'y [ts^{h}ɿ^{55}] (said of plants and animals) female
T2A	排 pɦa [p^{h}ai^{35}] to be in a row	抬 tɦé/[t^{h}ai^{35}] to lift up	曹 sɦau [ts^{h}au^{35}] a Chinese surname	湖 ɦu [xu^{35}] lake	渠 kɦiu [tɕ^{h}y^{35}] channel	题 tɦi [t^{h}i^{35}] title	词 sɦy [ts^{h}ɿ^{35}] word
T2B	埋 mɦa [mai^{35}] to bury	篮 lɦén [lan^{35}] basket	牢 lɦau [lau^{35}] secure	吴 ng [u^{35}] wu	余 ɦiu [y^{35}] extra	迷 mɦi [mi^{35}] be crazy about	
T3A	摆 pa [pai^{214}] to pose	胆 tén [tan^{214}] courage	早 tsau [tsau^{214}] early	古 ku [ku^{214}] ancient	举 kiu [tɕy^{214}] to hold up	底 ti [ti^{214}] bottom	子 tsy [tsɿ^{214}] son 死 sy/[sɿ^{214}]

							dead
T3B	扯 ts'a [tsʰɤ^{214}] to chat	坦 t'én [tʰan^{214}] frank	草 ts'au [tsʰau^{214}] grass	苦 k'u [kʰu^{214}] bitter	取 ts'iu [tɕʰy^{214}] to bring	体 t'i [tʰi^{214}] body	此 ts'y [tsʰɿ^{214}] this/these
T4A	罢 pɦa [pa^{51}] to quit	诞 tɦén [tan^{51}] birth	造 sɦau [tsau^{51}] to make	户 ɦu [xu^{214}] household	聚 sɦiu [tɕy^{51}] to get together	弟 tɦi [ti^{51}] young brother	似 zy [tsʰɿ^{51}] similar
T4B	买 mɦa [mai^{214}] to buy	懒 lɦén [lan^{214}] lazy	老 lɦau [lau^{214}] old	午 ng [u^{214}] noon	雨 ɦiu [y^{214}] rain	米 mɦ [mi^{214}] rice	
T5A	拜 pa [pai^{51}] to worship	旦 tén/[tan^{51}] daybreak	照 tɕau [tʂau^{51}] to shine	故 ku [ku^{51}] old	句 kiu [tɕy^{51}] sentence	帝 ti [ti^{51}] king	四 sy [sɿ^{51}] four
T5B	派 p'a [pʰai^{51}] to send	炭 t'én [tʰan^{51}] carbon	糙 ts'au [tsʰau^{55}] coarse	库 k'u [kʰu^{51}] storehouse	趣 ts'iu [tɕʰy^{51}] interesting	替 t'i [tʰi^{51}] to substitute	刺 ts'y [tsʰɿ^{51}] to penetrate
T6A	败 pɦa [pai^{51}] failure	蛋 tɦén [tan^{51}] egg	召 sɦau [tʂau^{51}] to call together	护 ɦu [xu^{51}] to protect	具 kɦiu [tɕy^{51}] tool	第 tɦi [ti^{51}] a prefix before a number	寺 zy [tsʰɿ^{51}] temple
T6B	卖 mɦa [mai^{51}] to sell	烂 lɦén [lan^{51}] rotten	闹 nau [nau^{51}] noisy	误 ng [u^{51}] mistake	预 ɦiu [y^{51}] in advance	谜 mɦi [mi^{35}] puzzle	

σ_2 / σ_1	[ʌʔ](Wúxī) [ɐʔ](Shànghǎi)	[ɔʔ]	[ɪʔ](Wúxī) [iɪʔ](Shànghǎi)	[əʔ]
T7A	百 paq [pai^{214}] a hundred	剥 poq [po^{55}] to peel 博 poq [po^{35}] plentiful	笔 pieq [pi^{214}] pen	拨 peq [po^{55}] to poke; to stir 失 ɕeq [ʂi^{55}] to lose
T7B	拍 p‘aq [p^{h}ai^{55}] to pat 插 tɕ‘aq [tʂ^{h}a^{55}] to insert	扑 pɦoq [p^{h}o^{55}] to flap	匹 p‘ieq [p^{h}i^{214}] classifier of horses 吃 k‘ieq [tʂ^{h}i^{55}] to eat	泼 p‘eq [p^{h}o^{55}] to spill out 赤 tɕ‘eq [tʂ^{h}i^{51}] red
T8A	白 pɦāq [pai^{35}] white 石 sɦāq [ʂi^{35}] stone	薄 pɦoq [po^{35}] thin	别 pɦiēq [piɛ^{35}] another	孛 (孛相=玩) pɦeq [po^{35}] to play
T8B	麦 mɦāq [mai^{51}] wheat 蜡 lɦāq [la^{51}] candle	木 mɦoq [mu^{51}] wood 绿 lɦoq [lv^{51}] green	蜜 mɦieq [mi^{51}] honey 力 lɦieq [li^{51}] force	沫 mɦeq [mo^{51}] froth

tone	char.	Wú Pron.	PTH Pron.	gloss
T1A	召	tɕau	[tʂɑu^{55}]	to summon, convene
T1A	颠	tien	[tian^{55}]	to bump
T1B	亲	ts‘ing	[tɕ^{h}in^{55}]	intimate
T1B	开	k‘é	[k^{h}ai^{55}]	to open
T2A	咸	ɦén	[xian^{35}]	salty
T2B	元	djy	[uan^{35}]	currency unit Yuan
T2B	毛	mɦau	[mau^{35}]	hair, feather
T3B	考	k‘au	[k^{h}ɑu^{214}]	test
T4A	腐	fɦu	[fu^{214}]	rotten
T4A	上	djang	[ʂɑŋ^{51}]	to go up
T4A	道	tɦau	[tau^{51}]	road
T4B	满	mɦu	[man^{214}]	full
T4B	冷	lɦang	[ləŋ^{214}]	cold
T4B	礼	lɦi	[li^{214}]	gift
T5A	蒜	sun	[suɑn^{51}]	garlic
T5A	记	ki	[tɕi^{51}]	remember
T5B	汽	k‘i	[tɕ^{h}i^{51}]	steam
T5B	靠	k‘au	[k^{h}ɑu^{51}]	to rely on
T6A	字	zy	[tsɿ^{51}]	word, character
T6A	大	tɦū	[ta^{51}]	big, great
T6A	附	fɦu	[fu^{51}]	to attach
T6B	面	mɦien	[mian^{51}]	flour
T6B	闹	nau	[nau^{51}]	noisy
T7A	说	ɕueq	[ʂuo^{55}]	to talk
T7B	出	tɕ‘ueq	[tʂ^{h}u^{55}]	out, beyond
T8A	合	ɦeq	[xɤ^{35}]	to combine
T8A	毒	tɦoq	[tu^{35}]	poison
T8A	绝	sɦiueq	[tɕuɤ^{35}]	to cut off

Appendix II Word lists of bisyllabic words

Table 1 Bisyllable words both used in PTH and the Wu dialects

σ₁ \ σ₂	T1	T2	T3	T4	T5	T6	T7	T8
T1A	遭殃 tsau øiang [tsɑu^{55} iɑŋ^{55}] to suffer a disaster	居民 kiu mɦing [tɕy^{55} min^{35}] resident 私人 sy nging [sɿ^{55} ʐən^{35}] personal	低保 ti pau [ti^{55} pɑu^{214}] subsistence allowance	街道 kia tɦau [tɕiɛ^{55} tɑu^{51}] street	颠倒 tien tau [tian^{55} tɑu^{214}] upside down	单调 tén tɦiau [tan^{55} tiɑu^{51}] monotonous and dull	资格 tsy kāq [tsɿ^{55} kɤ^{35}] qualifications	孤立 ku lɦieq [ku^{55} li^{51}] isolated
T1B	开销 k'é siau [k^{h}ai^{55} ɕiɑu^{55}] expenditure	梯形 t'i ɦing [t^{h}i^{55} ɕiŋ^{35}] ladder-shaped 雌雄 ts'y ɦiong [ts^{h}ɿ^{55} ɕyŋ^{35}] male and female 揩油 k'a ɦiou [k^{h}ai^{55} iou^{35}] to take advantages	开水 k'é sỹ [k^{h}ai^{55} ʂuei^{214}] boiling water	清淡 ts'ing tɦén [tɕ^{h}iŋ^{55} tan^{51}] (taste) light	枯燥 k'u sau [k^{h}u^{55} tsɑu^{51}] dull and dry	操练 ts'au lɦien [ts^{h}ɑu^{55} lian^{51}] to drill (military training) 摊派 t'én p'a [t^{h}an^{55} p^{h}ai^{51}] to apportion	区别 k'iu pɦiēq [tɕ^{h}y^{55} piɛ^{35}] difference	亲热 ts'ing ngieq [tɕ^{h}in^{55} ʐɿ^{51}] intimate
T2A	曹操 sɦau ts'au [ts^{h}au^{35} ts^{h}au^{55}]	排球 pɦa kɦiou [p^{h}ai^{35} tɕ^{h}iou^{35}] volleyball	词组 sɦy tsu [ts^{h}ɿ^{35} tsou^{214}] phrase	渠道 kɦiu tɦau [tɕ^{h}y^{35} tau^{51}] channel	咸菜 ɦén ts'é [ɕiɛn^{35} ts^{h}ai^{51}] brined	原旧 ngiuon kɦiou [yɛn^{35} tɕiou^{51}] as usual	湖北 ɦu poq [xu^{35} pei^{214}] a province	题目 tɦi mɦoq [ti^{35} mu^{51}] topic

	Cao Cao, a famous historical figure in china 台灯 t'é teng [t^{h}ai^{35} tən^{55}] desk lamp				vegetables		name in china	
T2B	迷宫 mi kong [mi^{35} kuŋ^{35}] maze	梧桐 ɦu tɦong [xu^{35} tuŋ^{35}] chinese parasol tree	篮板 lɦén pén [lan^{35} pan^{214}] Backboard 元宝 djy pau [yɛn^{35} pɑu^{214}] (gold or silver) ingot	厘米 lɦi mɦi [li^{35} mi^{214}] centimeter	牢靠 lɦau k'au [lɑu^{35} kɑu^{51}] secure	余地 ɦiu tɦi [y^{35} ti^{51}] leeway	毛笔 mɦau pieq [mau^{35} pi^{214}] writing brush	埋没 mɦa mo [mai^{35} mo^{51}] to bury
T3A	早操 tsau ts'au [tsɑu^{214} ts^{h}ɑu^{55}] morning exercises	子孙 tsy sueng [tsɿ^{214} suən^{55}] offspring	古董 ku tong [ku^{214} duŋ^{214}] antique	摆动 pa tɦong [pai^{214} duŋ^{51}] to swing	底细 ti si [ti^{214} ɕi^{51}] exact details	举例 kiu lɦi [tɕy^{214} li^{51}] to give an example	胆汁 tén tɕeq [tan^{214} tʂʅ^{55}] bile	死活 sy ɦueq [sɿ^{214} xuo^{35}] anyway
T3B	取笑 ts'iu siau [tɕ^{h}y^{214} ɕiɑu^{51}] to tease	扯皮 ts'a pɦi [tʂ^{h}ɛ^{214} p^{h}i^{214}] dispute over trifles	草纸 ts'au tɕy [ts^{h}ɑu^{214} tʂi^{214}] toilet paper	苦心 k'u sing [k^{h}u^{214} ɕin^{55}] painstaking efforts	考试 k'au si/ɕi [k^{h}ɑu^{214} ʂʅ^{51}] examination	草地 ts'au tɦi [ts^{h}ɑu^{214} ti^{51}] grassland	毯子 t'én tsy [t^{h}an^{214} tsɿ] blanket	体育 t'i ɦioq [t^{h}i^{214} y^{51}] physical training
T4A	罢工 pɦa kong [pa^{51} kuŋ^{55}] strike	领头 lɦing tɦou [liŋ^{214} dou^{35}] collar	户口 ɦu k'ou [xu^{51} k^{h}ou^{214}] registered residence	腐乳 ɦu djy [fu^{214} ʐu^{214}] preserved beancurd	上进 djáng tsing [ʂɑŋ^{51} tɕin^{51}] enterprising 造句	道地 tɦau tɦi [tɑu^{51} ti^{51}] authentic	道德 tɦau teq [tɑu^{51} tɤ^{35}] morality	聚集 sɦiu sɦieq [tɕy^{51} tɕi^{35}] to gather

					ɦau kiu [tsɑu^{51} tɕy^{51}] to make sentences			
T4B	买方 mɦa fang [mai^{214} faŋ^{55}] buyer 午休 ng hiou [u^{214} ɕiou^{55}] noon break 雨披 ɦiu p'i [y^{214} p^{h}i^{55}] poncho	冷盆 lɦang pɦeng [ləŋ^{214} p^{h}ən^{35}] cold dishes	老虎 lɦau hu [lau^{214} xu^{214}] tiger	蚂蚁 mo ngi [ma^{214} i^{214}] ant	礼拜 lɦi pa [li^{214} pai^{51}] week	懒惰 lɦén tū [lan^{214} tuo^{51}] lazy	米色 mɦi seq [mi^{214} sɤ^{51}] cream color	满月 mɦu ngiueq [man^{214} yɛ^{51}] a baby's completion of its first month of life
T5A	故居 ku kiu [ku^{51} tɕy^{55}] former residence	蒜苗 sun mɦiau [suɑn^{51} miɑu^{35}] garlic stem	担子 tén tsy [tan^{51} tsi] load	干部 kun pɦu [kan^{51} pu^{51}] cadre	照顾 tɕau ku [tʂau^{51} ku^{51}] to take care of	句号 kiu ɦau [tɕy^{51} xɑu^{51}] full stop	拜托 pa t'ōq [pai^{51} t^{h}uo^{55}] to request	记录 ki lɦoq [tɕi^{51} lu^{51}] record
T5B	替身 t'i ɕeng [t^{h}i^{51} ʂən^{55}] stand-in	库存 k'u sɦueng [k^{h}u^{51} ts^{h}ən^{35}] stock	汽水 k'i ɕÿ [tɕ^{h}i^{51} ʂuei^{214}] soda water	刺猬 ts'y ɦué [ts^{h}ɿ^{51} uei^{214}] hedgehog	靠背 k'au pé [k^{h}ɑu^{51} pei^{51}] backrest	趣味 ts'iu ɦué [tɕ^{h}y^{51} uei^{51}] interest	派别 p'a pɦiēq [p^{h}ai^{51} piɛ^{35}] faction	快活 k'ua ɦueq [k^{h}uai^{51} xuo^{35}] cheerful
T6A	召开 sɦau k'é [tʂɑu^{51} k^{h}ai^{55}] to convoke	蛋黄 tɦén ɦuang [tan^{51} xuɑŋ^{35}] yolk	字典 zy tien [zɿ^{51} tian^{214}] dictionary	护士 ɦu sɦy [xu^{51} ʂʅ^{51}] nurse	大蒜 tɦū sun [ta^{51} suan^{51}] garlic	寺庙 zy mɦiau [ts^{h}ɿ^{51} miɑu^{51}] temple	败笔 pɦa pieq [pai^{51} pi^{214}] a faulty expression in writing	附属 fɦu djoq [fu^{51} ʂu^{214}] affiliated
T6B	闹钟 nau tɕong	烂糊 lɦén ɦu	面孔 mɦien k'ong	号码 ɦau mo	卖相 mɦa siang	自愿 zy ngiuon [tsɿ^{51}	预测 ɦiu ts'e	卖力 mɦa lɦieq

	[nɑu^{51} tʂuŋ^{55}] alarm clock	[lan^{51} xu^{35}] pulpy	[miɛn^{51} kuŋ^{214}] face	[xau^{51} ma^{214}] number	[mai^{51} ɕiɑŋ^{51}] appearance	yɛn^{51}] voluntarily	[y^{51} ts^{h}ɤ^{51}] to forecast	[mai^{51} li^{51}] to work hard
T7A	说书 ɕueq ɕÿ [ʂuo^{55} ʂu^{55}] Chinese storytelling (by a professional storyteller)	拔河 pɦāq ɦū [pa^{35} xɤ^{35}] tug of war	笔挺 pieq t'ing [pi^{214} t^{h}iŋ^{214}] (standing) very straight	博士 poq zy [po^{35} ʂʅ^{51}] doctor (degree)	百姓 paq sing [pai^{214} ɕiŋ^{51}] common people	失败 ɕeq pɦa [ʂʅ^{55} pai^{51}] failure	剥削 poq siaq [po^{55} ɕyɛ^{55}] to exploit	
T7B	插销 tɕ'aq siau [tʂ^{h}a^{55} ɕiɑu^{55}] bolt	噱头 siaq tɦou [ɕuɤ^{51} t^{h}ou^{35}] tricks meant to deceive	拍板 p'aq pén [p^{h}ai^{55} pan^{214}] to give the final decision	出动 tɕ'ueq tɦong [tʂ^{h}u^{55} tuŋ^{51}] to dispatch troops	吃醋 k'ieq ts'u [tɕ^{h}ʅ^{55} ts^{h}u^{51}] jealousy	匹配 p'ieq p'é [p^{h}i^{214} p^{h}ei^{51}] matched	扑灭 pɦoq mɦiēq [p^{h}u^{55} miɛ^{51}] to put out (the fire)	泼辣 p'eq lɦāq [p^{h}o^{55} la^{51}] daring and resolute
T8A	石灰 sɦāq hué [ʂi^{35} xuei^{55}] lime	舌头 ɕɦeq tɦou [ʂɤ^{35} t^{h}ou^{35}] tongue	白酒 pɦāq tsiou [pai^{35} tɕiu^{214}] white spirit	别墅 pɦiēq ɕɦÿ [piɛ^{35} ʂu^{51}] villa	绝对 sɦiueq té [tɕuɤ^{35} tuei^{51}] absolutely	薄利 pɦoq lɦi [po^{35} li^{51}] small profits	蹩脚 pɦiēq kiaq [piɛ^{51} ɕiɑu^{214}] inferior	毒药 tɦoq ɦiaq [tu^{35} jɑu^{51}] poison
T8B	蜜蜂 mɦieq fong [mi^{51} fəŋ^{55}] bee	木头 mɦoq tɦou [mu^{51} t^{h}ou^{35}] wood	木板 mɦoq pén [mu^{51} pan^{214}] wood board	合理 ɦeq li [xɤ^{35} li^{214}] reasonable	力气 lɦieq k'i [li^{51} tɕ^{h}i^{51}] energy	麦片 mɦāq p'ien [mai^{51} p^{h}ian^{51}] oatmeal 末代 mɦeq tɦé [mo^{51} tai^{51}] last reign of a dynasty	蜡烛 lɦāq tɕoq [la^{51} tʂu^{35}] candle	绿叶 lɦoq ɦiēq [lv^{51} jɤ^{51}] green leaves

Table 2 Bisyllabic words only used in the Wu dialects

Tone	word	Wú Pron.	gloss
T1A.T1	家生	ka sáng	furniture
T1A.T1	身胚	ɕeng p‘é	body
T1A.T2	扳头	pén tɦou	spanner
T1B.T2	坍台	t‘én t‘é	to collapse
T1B.T2	开洋	k‘é ɦiang	dried shrimp
T1B.T2	摊头	t‘én tɦou	booth
T1B.T3	推板	t‘é pén	inferior
T1B.T3	欢喜	huon hi	happy and joyous
T1A.T4	搬场	pūn tɕ‘áng	to move
T1A.T5	关照	kuén tɕau	to keep an eye on
T1A.T5	心相	sing siang	interest
T1B.T5	亲眷	ts‘ing kiuon	relatives
T1A.T6	新妇	sing vu	bride
T1B.T6	粢饭	ts‘y vén	glutinous rice
T1A.T7	收作	ɕou tsōq	to tidy up
T1A.T8	生活	sáng ɦueq	drubbing
T1A.T8	精肉	tsing ngioq	lean meat
T2A.T1	辰光	ɕɦeng kuang	time
T2A.T1	馋胚	sɦén p‘é	glutton
T2A.T1	调羹	tɦiau káng	[tʰiau^{35} kən^{55}]
T2A.T1	停当	tɦing tɦang	[tʰiŋ^{35} tɑŋ^{55}]
T2B.T1	洋机	ɦiang ki	sewing machine
T2B.T1	莲心	lɦi sing	lotus seed
T2B.T2	名堂	ming tɦang	variety
T2B.T2	胡咙	ɦu long	throat
T2B.T2	洋盘	ɦiang pɦūn	an amateur who is prone to be fooled
T2B.T2	浑堂	ɦueng tɦang	N/A
T2B.T3	难板	nén pén	seldom
T2B.T3	门坎	meng k‘én	door sill
T2B.T5	文旦	fɦeng tén	pomelo
T2B.T6	闲话	ɦén ɦuo	gossip
T2B.T8	寒热	ɦun ngieq	fever
T3A.T1	总归	tsong kué	after all
T3A.T1	死腔	sy k‘iang	smugness
T3B.T1	坦眼	t‘én ngén	be shy of
T3B.T4	炒米	tɕ‘au mɦi	parched rice
T3A.T5	爽气	suang k‘i	straightforward
T3B.T5	讨打	t‘au táng	so naughty as to invite parental discipline
T3B.T6	讨骂	t‘au mɦo	to have done something unfavorable
T3A.T7	准足	tɕueng tsoq	on time
T3B.T7	挺刮	kuaq kuaq	(cloth) stiff and smooth
T3B.T8	火着	hū ɕɦaq	on fire
T4A.T1	像腔	ziang k‘iang	to behave properly
T4A.T1	上腔	djáng k‘iang	to retort sarcastically
T4B.T1	懒胚	lɦén p‘é	lazy bone
T4B.T1	眼泡	ngén p‘au	eyelid
T4B.T3	老酒	lɦau tsiou	liquor
T4A.T5	上劲	djáng king	enthusiastic about doing something
T4B.T5	惹气	dja k‘i	angry
T4B.T5	有劲	ɦiou king	interesting

T4A.T6	上路	djáng lɦu	reliable
T4B.T8	暖热	nun ngieq	warm
T5A.T1	粪箕	feng ki	dustpan
T5A.T1	灶间	tsau kien	kitchen
T5A.T1	傧相	ping siang	groomsman & bridesmaid
T5A.T2	寄娘	ki ngiang	adopted mother
T5A.T2	绢头	kiuon tɦou	handkerchief
T5A.T5	对过	té kū	opposite
T5B.T5	睏裤	k'ueng k'u	pyjama trousers
T5A.T7	信壳	sing k'ōq	envelope
T5A.T8	壮肉	tsuang ngioq	fat meat
T6A.T1	自家	zy ka	oneself
T6A.T1	便当	pɦien tɦang	convienient
T6A.T2	寿头	djou tɦou	a fool
T6A.T2	旧年	kɦiou ngien	last year
T6B.T2	弄堂	lɦong tɦang	alley
T6B.T2	用场	ɦiong tɕ'áng	function
T6A.T3	事体	sɦy t'i	affair; matter
T6A.T4	运道	ɦiuing tɦau	fortune
T6B.T4	料重	lɦiau ɕɦiong	(burden) heavy
T6B.T4	闹猛	nau mɦeng	bustling with noise and excitement
T6A.T5	硬劲	ngáng king	manage to do sth with difficulty
T6B.T6	夜饭	ɦia fɦén	supper
T6B.T6	烂饭	lɦén fɦén	rice cooked with too much water
T6A.T7	硬扎	ngáng tsāq	firm, strong
T6B.T8	赖学	lɦa ɦōq	to play truant
T7A.T1	瘪三	piēq sén	a wretch or tramp who lives by begging or stealing
T7A.T1	识相	ɕeq siang	be sensible
T7A.T1	搭腔	tāq k'iang	respond
T7B.T1	豁边	huaq p'ien	to overdo
T7B.T2	豁拳	huaq kɦiuon	finger-guessing game
T7A.T3	作死	tsōq sy	seek death
T7A.T4	脚布	kiaq pu	cloth wore on foot
T7A.T4	结棍	kieq kueng	strong
T7A.T4	搭界	tāq ka	be related
T7B.T4	赤佬	tɕ'eq lɦau	rogues
T7A.T5	笃定	toq tɦing	be sure of
T7A.T5	适意	ɕeq Ǿi	comfortable
T7A.T6	百脚	paq kiaq	centipede
T7B.T7	促掐	ts'oq k'āq	mean
T7B.T7	镯头	ɕɦoq tɦou	bracelet
T7B.T8	吃力	k'ieq lɦieq	strenuous
T8A.T1	贼腔	sɦeq k'iang	dubious-looking
T8B.T1	落班	lɦōq pén	to go off work
T8A.T2	着棋	ɕɦaq kɦi	play chess
T8B.T2	学堂	ɦōq tɦang	school
T8A.T3	蚀本	ɕɦeq peng	lose one's capital
T8B.T3	辣手	lāq ɕou	ruthless method
T8B.T4	日里	ngieq lɦi	daytime
T8B.T4	屋里	Ǿueq lɦi	in house
T8B.T4	落雨	lōq ɦiu	to rain

T8B.T5	入调	ɕɦueq tɦiau	accord with rules
T8A.T6	特为	tɦeq ɦué	specially
T8B.T6	力道	lɦieq tɦau	power
T8B.T6	物事	veq sɦy	affair
T8A.T7	折脚	ɕɦeq kiaq	faild twice in examination
T8B.T7	日脚	ngieq kiaq	sunlight
T8B.T8	热络	ngieq lɦōq	intimate
T8B.T8	活络	ɦueq lɦōq	activating collaterals

Table 3 Reduplicated Verbs

σ₂ / σ₁	T1	T2	T3	T4	T5	T6	T7	T8
A	遮遮 tɕo tɕo [tʂɤ^{55} tʂɤ] to hide	陪陪 pɦé pɦé [pʰei^{35} pʰei] to company	改改 ké ké [gai^{214} gai] to revise	动动 tɦong tɦong [tuŋ^{51} tuŋ] to move	照照 tɕau tɕau [tʂɑu^{51} tʂɑu] to look into the mirror	垫垫 tɦie tɦie [tian^{51} tian] to pad with	刮刮 kuaq kuaq [guɑ^{55} guɑ] to scrape	读读 tɦoq tɦoq [tu^{35} tu] to read
B	猜猜 ts'é ts'é [tsʰai^{55} tsʰai] to guess	量量 liang liang [liaŋ^{35} liaŋ] to measure	炒炒 tɕ'au tɕ'au [tʂʰɑu^{214} tʂʰɑu] to fry	咬咬 ngau ngau [iau^{214} iau] to bite	跳跳 tɦiau tɦiau [tʰiɑu^{51} tʰiɑu] to jump	练练 lie lie [lian^{51} lian] to practice	吃吃 k'ieq k'ieq [tʂʰi^{55} tʂʰi] to eat/taste	热热 ngieq ngieq [zɤ^{51} zɤ] to heat (food)

Table 4 Morphemes for eliciting expressions with "verb + resultative or directional complements"

tone	char.	Wú Pron.	PTH Pron.	gloss
Part 1: σ_1				
T1	敲	k'au	[tɕhiɑu^{55}]	to knock
T2	磨	mū	[mo^{35}]	to grind
T3	剪	tsien	[tɕian^{214}]	to cut by scissors
T4	斩	tsén	[tʂan^{214}]	to chop
T5	炸	tso	[tʂa^{51}]	to explode
T6	弄	long	[noŋ^{51}]	to make, to do
T7	切	ts'iēq	[tɕ^{h}iɛ^{55}]	to cut by knife
T8	划	ɦuaq	[xuɑ^{51}]	to slice
Part 2: σ_2				
T1	开	khé	[kai^{55}]	open
T2	平	pɦing	[piŋ^{35}]	flat
T3	好	hau	[hau^{214}]	the action has been done (satisfyingly)
T4	断	tɦun	[tuan^{51}]	broken into two parts
T5	破	phū	[po^{51}]	broken into pieces
T6	坏	ɦua	[huai^{51}]	broken
T7	出	tɦhueq	[tɕ^{h}u^{55}]	out
T8	落	lōq	[luo^{51}]	the action has been done

Table 5 Morphemes for eliciting expressions with "number + classifier"

tone	char.	Wú Pron.	PTH Pron.	gloss
Part 1: σ_1				
T1	三	sén	[san^{55}]	three
T2	/	/	/	/
T3	九	kiou	[tɕiu^{214}]	nine
T4	两	liang	[liaŋ^{214}]	two
T5	四	sy	[sɿ^{51}]	four
T6	廿	ngien	[niɛn^{51}]	twenty
T7	八	pāq	[pa^{55}]	eight
T8	六	lɦoq	[liu^{51}]	six
Part 2: σ_2				
T1	斤	king	[tɕin^{55}]	a weight equal to 0.5 kg
T2	年	ngien	[niɛn^{35}]	year
T3	把	po	[pa^{214}]	a handful of
T4	米	mi	[mi^{214}]	meter
T5	块	kɦua	[khuai^{51}]	a measure for pieces and monetary units
T6	段	tɦun	[tuan^{51}]	paragraph
T7	只	tseq	[tʂʅ^{55}]	a measure for animals like hens, birds,etc.
T8	粒	lieq	[li^{51}]	a measure for small round pieces

Table 6 Bisyllable words using exceptional tone sandhi patterns in the Wúxī dialects

tone combination	word	Wú Pron.	PTH Pron.	gloss
T1A.T2	当然	tɦang nien	[tɑŋ^{55} ʐan^{35}]	absolutely
T1A.T3	光火	kuang hū	N/A	irritated
T1A.T3	刚巧	kang k‘iau	[gɑŋ^{55} tɕʰiɑu^{214}]	by coincidence
T1B.T5	喷嚏	p‘eng ti	[pʰən^{55} tʰi^{51}]	sneeze
T1B.T5	牵记	k‘ien ki	[tɕʰian^{55} tɕi^{51}]	to care and miss
T1B.T5	钞票	tɕ‘au p‘iau	[tʂʰɑu^{55} pʰiɑu^{51}]	currency note
T1B.T7	听说	t‘ing ɕueq	[tʰiŋ^{55} ʂuo^{55}]	to hear of
T2A.T1	层单	sɦeng tén	N/A	sheet
T2A.T1	提高	tɦi kau	[tʰi^{35} kau^{55}]	to raise or improve
T2B.T1	阳台	ɦiang t‘é	[iɑŋ^{35} tʰai^{35}]	balcony
T2A.T3	团长	tɦun tɕáng	[tʰuɛn^{35} tʂɑŋ^{214}]	team leader
T2A.T5	残废	sɦén fi	[tsʰan^{35} fei^{51}]	disabled
T2B.T6	和调	ɦū tɦiau	N/A	to play a joke
T2B.T6	迷露	mɦi lɦu	N/A	fog
T2B.T7	难得	nén teq	[nan^{35} tɤ ^{35}]	rare
T3A.T1	本生	peng sáng	N/A	itself
T3A.T1	始终	sy tɕong	[ʂi^{214} tʂuŋ^{55}]	from beginning to end
T3A.T2	感情	kūn sɦing	[gan^{214} tɕhiŋ^{35}]	sentiment
T3A.T2	本来	peng lɦé	[pən^{214} lai^{35}]	original
T3B.T2	可怜	k‘ū lɦien	[khɤ^{214} lian^{35}]	pitiful
T3B.T2	可能	k‘ū neng	[khɤ^{214} nəŋ^{35}]	possible
T3B.T3	可耻	k‘ū tɕ‘y	[khɤ^{214} tʂhi^{214}]	shameful
T3A.T4	所以	sū Ǿi	[suo^{214} ji^{214}]	therefore
T3B.T4	好像	hau ziang	[xɑu^{214} ɕiɑŋ^{51}]	seem to be, as if
T3B.T5	可靠	k‘ū k‘au	[khɤ^{214} khɑu^{51}]	reliable
T3A.T6	享受	hiang ɕɦou	[ɕiɑŋ^{214} ʂou^{51}]	to enjoy
T3B.T6	讨论	t‘au lɦeng	[thɑu^{214} luən^{51}]	to discuss
T3A.T7	享福	hiang foq	[ɕiɑŋ^{214} fu^{35}]	to live in ease and comfort
T3A.T7	蝙蝠	pien foq	[pian^{214} fu^{35}]	bat
T3B.T7	可惜	k‘ū sieq	[khɤ^{214} ɕi^{55}]	unfortunately
T4B.T1	老官	lɦau kuon	N/A	husband
T4B.T2	理由	lɦi ɦiou	[li^{214} iou^{35}]	reason
T4B.T2	羽毛	iu mɦau	[y^{214} mɑu^{35}]	feather
T4A.T3	户口	ɦu k‘ou	[xu^{51} kʰou^{214}]	registered residence
T4B.T3	老酒	lɦau tsiou	N/A	liquor
T4B.T3	耳朵	ngi tū	[ɚ^{214} tuo]	ear
T4A.T4	丈姆	ɕɦáng m	N/A	mother-in-law
T4A.T5	负数	fu su	[fu^{51} ʂu^{51}]	negative number
T4A.T5	鳝片	ɕɦūn p‘ien	[ʂan^{51} pʰian^{51}]	eel slice
T4A.T5	上照	djáng tɕau	N/A	photogenic
T4A.T6	近视	kɦing sɦy	[tɕin^{51} ʂi^{51}]	myopic
T4A.T6	部队	pɦu tɦé	[pu^{51} tuei^{51}]	army
T4A.T8	序幕	sɦiu mɦu	[ɕy^{51} mu^{51}]	prologue
T5A.T2	帝王	ti ɦuang	[ti^{51} wɑŋ^{35}]	emperor
T5A.T4	最近	tsué kɦing	[tsuei^{51} tɕin^{51}]	recently
T6A.T2	调查	tɦiau sɦo	[tiau^{51} tʂʰa^{35}]	to investigate

T6A.T3	具体	kɦy t‘i	[tɕy^{51} tʰi^{214}]	specific
T6B.T3	效果	ɦiau kū	[ɕiau^{51} kuo^{214}]	effect
T6A.T4	导演	tɦau ien	[tau^{214} iɛn^{214}]	directors of film, drama, etc.
T6A.T4	万岁	mɦén sué	[øuɑŋ^{51} suei^{51}]	long live
T6B.T5	夜快	ɦia k‘ua	N/A	at night fall
T6B.T5	浪费	lɦang fi	[lɑŋ^{51} fei^{51}]	waste
T6B.T6	预料	ɦiu lɦiau	[y^{51} liau^{51}]	to predict
T6B.T6	议论	ngi lɦeng	[i^{51} lun^{51}]	to discuss
T7A.T1	作兴	tsōq hing	N/A	be willing to
T7B.T5	一向	øieq hiang	[ji^{55} ɕiɑŋ^{51}]	always
T8A.T3	白酒	pɦāq tsiou	[pai^{35} tɕiu^{214}]	Kaoliang spirit
T8A.T3	月饼	ngiueq ping	[juɤ^{51} piŋ^{214}]	moon cake
T8A.T3	局长	kɦioq tɕáng	[tɕu^{35} tʂɑŋ^{214}]	director general
T8B.T3	墨水	mɦeq ɕÿ	[mo^{51} ʂuei^{214}]	ink
T8B.T3	热水	ngieq ɕÿ	[ʐɤ^{51} ʂuei^{214}]	hot-water
T8B.T3	辣手	lāq ɕou	[la^{51} ʂou^{214}]	ruthless method

Appendix III Paragraphs

* Words underlined are target words for testing the citation tone and tone sandhi. Words in bold are the filler words.

中国人的姓名

中国人的姓名：中国人的姓有一个字的，也有两个字和两个字以上的。一个字的是单姓，两个字或两个字以上的是复姓。十三亿中国人到底有多少个姓，到现在也没有 准确的统计数字。国家公安部最近统计出来，中国现在使用的姓氏有四千七百多个，其中 "王"是第一大姓，有近九千六百万人，占全国 人口 总数的百分之七点二五。人口数在两千万人以上的姓一共有十个，分别是：王、张、李、刘、陈、赵、周、吴、杨、黄。另外，像孙、林、石、廖、戚、董、易、郑、罗也都是排名前一百位的姓。

中国人的姓名都是姓在前，名在后。名字有一个字的，也有两个字的。名字往往有一定的含义，表示某种 愿望。有的名字 包含出生时的地点、时间或天气。如"**京**"、"晨"、"雪"、"**冬**"等。有的名字表示希望 具有某种美德，如"**忠**"、"义"、"礼"、"仁"、"**信**"等。有的名字表示希望长寿、健康、幸福、快乐的意思，如"**松**"、"寿"、"健"、"福"、"**乐**"等。男人的名字和女人的名字也不一样，男人的名字多用表示威武、英勇的字，如"**刚**"、"虎"、"龙"、"勇"、"**强**"等。女人的名字常用表示漂亮 好看的字，如"凤"、"花"、"彩"、"玉"、"秀"、"**娟**"等。

Translation:

Names of Chinese people

Most surnames of Chinese people have only one character, however, it exists that the double-character surnames or surnames with three or more characters. Up to now, there is no exact statistic on how many surnames there are in China for 1.3 billion people. The recent statistical figures from the Ministry of Public Security show that the surnames using nowadays exceed 4700, among which *Wang* is the most common surname shared by almost 93 million people and made up 7.25% of the country's total population. There are ten surnames used by more than 20 million people respectively, they are ***Wang, Zhang,*** *Li, Liu, Chen, Zhao, Zhou,* ***Wu, Yang, Huang***. In addition, surnames such as ***Sun, Lin,*** *Shi, Liao, Qi, Dong, Yi,* ***Zheng, Luo*** are all within the top 100 Chinese surnames.

In China, the surname comes first followed by the given name, and the latter has its own traditions and features. It can have one or two characters. Chinese name usually have a certain meaning, expressing some kind of wish. Some names embody the location, time or a natural phenomenon when the person was born, such as "***Jing*** (Beijing)", "*Chen* (morning)", "*Xue* (snow)", "***Dong*** (winter)". Some names indicate the expectation of possessing some virtues, such as "***Zhong*** (loyalty)", "*Yi* (righteousness)", "*Li* (etiquette)", "*Ren* (benevolence)"and "***Xin*** (faith)". Some names convey the meaning of longevity, healthy and happiness, such as "***Song*** (pine, representing longevity)", "*Shou* (longevity) ", "*Jian* (health)", "*Fu* (good fortune)" and "***Le*** (happiness)". Male names are different from female ones: men's names usually have the character meaning power and bravery, such as "***Gang*** (hardness)", "*Hu* (tiger)", "*Long* (dragon)", "*Yong* (courage)", "***Qiang*** (strength)". And names of females usually use characters representing beauty, such as "***Feng*** (phoenix)", "*Hua* (flower)", "*Cai* (colors)", "*Yu* (jade)", "*Xiu* (elegent)", "***Juan*** (graceful)".

Monosyllabic morphemes in the paragraphs

tone	char.	Wú Pron.	PTH Pron.	gloss
T1A	周	tɕou	[tʂou^{55}]	A Chinese surname
T1B	花	huo	[xuɑ^{55}]	flower
T1B	龙	long	[luŋ^{35}]	dragon
T2A	陈	ɕɦeng	[tʂʰən35]	a surname
T2A	晨	ɕɦeng	[tʂʰən35]	morning
T2B	刘	lɦiou	[liou35]	a surname
T2B	仁	djeng	[ʐən35]	benevolence
T3A	董	tong	[tuŋ^{214}]	A Chinese surname
T3B	彩	ts'é	[tsʰai^{214}]	colorful
T3B	虎	hu	[xu^{214}]	tiger
T4A	赵	ɕɦau	[tsau51]	a surname
T4B	李	lɦi	[li214]	a surname
T4B	礼	lɦi	[li214]	etiquette
T4B	勇	ɕɦiong	[ioŋ214]	courage
T5A	秀	siou	[ɕiu^{51}]	excellent
T6A	寿	djou	[ʂou51]	longevity
T6A	健	kɦien	[tɕian51]	health
T6B	廖	lɦiau	[liau51]	a surname
T6B	义	ngi	[ji51]	righteousness
T7A	福	foq	[fu^{35}]	blessing and good fortune
T7A	雪	siueq	[ɕuɤ^{214}]	snow
T7B	易	ɦieq	[ji^{51}]	be prone to
T7B	戚	ts'ieq	[tɕʰi^{51}]	A Chinese surname
T8A	石	sɦāq	[ʂi35]	stone
T8B	玉	ngioq	[ju51]	jade

Bisyllabic words in the paragraphs

Tone	word	Wú Pron.	PTH Pron.	gloss
T1A.T2	包含	pau ɦūn	[pɑu^{55} xan^{35}]	to include
T1A.T5	单姓	tén sing	[tan^{55} ɕiŋ^{51}]	one-character surname
T1A.T7	中国	tɕong kū	[tʂuŋ^{55} guo^{35}]	China
T1B.T5	天气	t'ien k'i	[tʰian^{55} tɕʰi^{51}]	weather
T1B.T6	希望	hi øuang	[ɕi^{55} wɑŋ^{51}]	hope
T2A.T1	其中	kɦi tɕong	[tɕʰi^{35} tʂuŋ^{55}]	among
T2A.T1	时间	ɕɦy kien	[ʂi^{35} tɕiɛn^{55}]	time
T2A.T2	排名	pɦa mɦing	[pʰai^{35} miŋ^{35}]	rank
T2A.T6	含义	ɦūn ngi	[xan^{35} i^{51}]	meaning
T2A.T6	长寿	ɕɦáng djou	[tʂʰɑŋ^{35} ʂou^{51}]	longevity
T2A.T6	常用	ɕɦáng ɦiong	[tʂʰɑŋ^{35} yŋ^{51}]	often used
T2A.T7	全国	sɦiuon kū	[tɕʰuɛn^{35} kuo^{35}]	the entire country
T2B.T2	男人	nūn nging	[nan^{35} ʐən^{35}]	man
T2B.T2	威武	ué fɦu	[uei^{55} u^{214}]	powerful
T2B.T3	人口	nging k'ou	[ʐən^{35} kʰou^{214}]	population
T2B.T6	名字	mɦing zy	[miŋ^{35} tsɿ^{51}]	name
T3A.T6	表示	piau sɦy	[piɑu^{214} ʂi^{51}]	to mean
T3A.T6	使用	sy ɦiong	[ʂi^{214} juŋ^{51}]	to use
T3A.T6	总数	tsong su	[tsuŋ^{214} ʂu^{51}]	sum
T3A.T7	准确	tɕueng k'iaq	[tʂuən^{214} tɕʰuɤ^{51}]	accurate
T3B.T5	好看	hau k'un	[xɑu^{214} kʰan^{51}]	beautiful

T3B.T5	统计	t‘ong ki	[t^{h}uŋ^{214} tɕi^{51}]	to calculate
T4B.T2	女人	ngiu nging	[ny^{214} ʐən^{35}]	woman
T4B.T3	某种	mɦou tɕong	[mou^{214} tʂuŋ^{214}]	some kind
T4B.T5	两个	lɦiang ke	[liɑŋ^{214} ke^{51}]	two
T5A.T2	姓名	sing mɦing	[ɕiŋ^{51} miŋ^{35}]	name
T5A.T3	到底	tau ti	[tɑu^{51} ti^{214}]	on earth
T5A.T4	最近	tsué kɦing	[tsuei^{51} tɕin^{51}]	recently
T5A.T6	数字	su zy	[ʂu^{51} tsi^{51}]	number
T5B.T6	漂亮	p‘iau lɦiang	[p^{h}iɑu^{51} liɑŋ^{51}]	pretty
T5B.T8	快乐	k‘ua lɦo̅q	[k^{h}uɑi^{51} lɤ^{51}]	joyful
T6A.T1	健康	kɦien k‘ang	[tɕiɛn51 kʰan55]	healthy
T6A.T4	地点	tɦi tien	[ti^{51} tiɛn ^{214}]	location
T6A.T4	具有	kɦiu ɦiou	[tɕy^{51} iou^{214}]	have, process
T6B.T6	另外	lɦing nga	[liŋ^{51} uai^{51}]	moreover
T6B.T6	愿望	ngiuon fɦuang	[yɛn^{51} uɑŋ^{51}]	desire, wish
T7A.T5	一个	øieq ke	[ji^{55} gɤ^{51}]	one
T7A.T6	一定	øieq tɦing	[ji^{55} tiŋ^{51}]	definitely
T7A.T7	幸福	?? foq	[ɕiŋ^{51} fu^{35}]	happy
T7B.T2	出来	tɕ‘ueq lɦé	[tʂ^{h}u^{55} lai^{35}]	to come out
T7B.T7	六百	lɦoq paq	[liu^{51} pai^{214}]	six hundred
T8A.T1	十三	ɕheq sén	[ʂi^{35} san^{55}]	thirteen
T8A.T5	复姓	foq sing	[fu^{51} ɕiŋ^{51}]	compound surname
T8A.T5	十个	ɕheq ke	[ʂi^{35} gɤ^{51}]	ten
T8B.T4	没有	mɦeq ɦiou	[mei^{35} jou^{214}]	no, without

Appendix IV Interview questionnaire

版本一（供会说普通话和方言的讲话人用）

方言和普通话学习经历

1. 你是哪里人？出生、生长在哪个区？你觉得自己说的是什么话？
2. 是不是从小说本地方言？如果不是，还能记得是从几岁开始说的吗？
   你的本地方言是谁教的？
3. 是不是从小说普通话？如果不是，还能记得是从什么时候开始说的吗？
4. 你的普通话是谁教的？
5. 你的父母会说普通话吗？
6. 你在外地住过半年以上吗？

语言能力

1. 你还会说什么其他方言吗？
2. 你觉得自己的本地方言说得怎么样？
3. 你在说本地方言时会混入普通话或者其他方言吗？
   如果会，这种情况经常发生吗？
4. 当你说方言的时候要表达一些社会上新近产生的词，你是会用本地方言的音调说出还是直接转普通话说出？

语言使用

1. 现在每天用本地方言吗？平时说普通话多还是本地方言多？这种状况持续多久了？
2. 当你和同事谈工作方面的事的时候，你用普通话还是本地方言多？
3. 当你和朋友谈一些比较私人的话题的时候，你用普通话还是本地方言多？
4. 问老年人：你和老伴讲话用普通话多还是方言多？和兄弟姐妹呢？和你的孩子呢？
   问年轻人：你和爷爷奶奶讲话用普通话多还是方言多？和父母呢？和兄弟姐妹呢？
5. 以下场景你更倾向于用普通话还是本地方言？

（1）（在当地）问路；（2）（在当地）商店买东西 （3）（在当地）政府机构办事

语言态度

1. 你觉得本地方言和普通话哪个更好听？
2. 你觉得说本地方言的人显得诚恳，值得信任呢还是说普通话的人更诚恳，值得信任呢？
3. 对你来说，本地方言和普通话听起来哪个更亲切？
4. 你觉得说本地方言的人比说普通话的人更受尊重吗？
5. 你觉得在无锡，本地方言和普通话哪个更有用？
6. 你觉得说本地方言的人比说普通话的人更有社会地位，有权威吗？

个人信息

性别（观察，记录，不问）；出生年月；职业；最高学历

本地特色食品访谈题目包括：

1. 你喜欢吃什么?你不喜欢吃什么?
2. 你最喜欢的本地小吃是什么?
3. 你最喜欢的本地特色菜是什么? 哪里的比较正宗?
4. 你会做饭吗? 你有没有什么拿手菜? 你能告诉我做好这道菜的诀窍是什么吗?
5. 你觉得怎么吃才有益健康? 你觉得什么是健康的饮食习惯?

Version I (for people who can speak both local dialect and PTH)

Language Study and Environment

1. Where are you from? Which district of this city were you born in? Which dialect do you speak?
2. Have you spoken dialect since you were very young?
   If not, could you please recall from when? Who taught you the dialect?
3. Have you spoken PTH since you were very young?
   If not, could you please recall from when? Who taught you PTH?
4. Are your parents able to speak dialect?
5. Are there any other places have you stayed for a long time?

Language Aptitude

1. What other dialects could you speak?
2. How do you think of your dialect aptitude?
3. Do you mix in words of PTH or other dialects when you speak dialect?
   If so, do you do that often?
4. When you come across words you do not use often in dialect, such as neologism, will you speak them out in the dialect accent or switch over into PTH?

Language Use

1. Do you use dialect every day? Which one do you use more often, PTH or dialect? How long has this been going on?
2. Which one of PTH and dialect do you use more often when you talk to your colleague/classmates about work/study-related topics?
3. Which one of PTH and dialect do you use more often use when you talk to your friends about personal matters?
4. ***For old people:*** which one of PTH and dialect do you use more often when you talk to your (1) wife/husband, (2) brothers/sisters, (3) children, (4) grandchildren?
   For young people: Which one of PTH and dialect do you use more often when you talk to your (1) grandparents, (2) parents, (3) cousins?
5. Please recall, which of PTH and dialect do you use more often when you (1) asked for road directions, (2) shopped in a shopping center (3) spoke to a government officer at a public counter?

Language Attitude

1. Which one do you prefer to hear between PTH and local dialect?
2. Which one more useful for you between PTH and local dialect?
3. Which one is friendlier for you between PTH and local dialect?
4. Which one do you think is more authoritative between PTH and local dialect?
5. Which one do you think is more cordial between PTH and local dialect?
6. Whose speaker do you think is more prestigious between PTH and local dialect?

Personal information

sex; age; occupation; level of education

Food topic

1. What are your favorite local snacks?
2. What's the typical local dish you like most? Where can we find them?
3. Are you able to cook? Do you have a specialty? What's your secret recipe?
4. What is your favorite food and what are your dislikes?
5. What do you think of health diet? What do you think is healthy eating habit? How to eat healthy?

版本一（供不会普通话，只会方言的讲话人用）

方言和普通话学习经历

1. 你是哪里人？出生、生长在哪个区？你觉得自己说的是什么话？
2. 是不是从小说本地方言？如果不是，还能记得是从几岁开始说的吗？
   你的本地方言是谁教的？
3. 你的父母会说本地方言吗？
4. 你在外地住过半年以上吗？
5. 你能听懂普通话吗？

语言能力

1. 你还会说什么其他方言吗？
2. 你觉得自己的本地方言说得怎么样？
3. 你在说本地方言时会混入其他方言吗？如果会，这种情况经常发生吗？
4. 当你说方言的时候要表达一些社会上新近产生的词，你怎么说？

语言使用

1. 你在各种场合都说本地方言吗？有例外吗？

语言态度

1. 你觉得本地方言和普通话哪个更好听？
2. 你觉得说本地方言的人显得诚恳，值得信任呢还是说普通话的人更诚恳，值得信任呢？
3. 对你来说，本地方言和普通话听起来哪个更亲切？
4. 你觉得说本地方言的人比说普通话的人更受尊重吗？
5. 你觉得在无锡，本地方言和普通话哪个更有用？
6. 你觉得说本地方言的人比说普通话的人更有社会地位，有权威吗？

个人信息

性别（观察，记录，不问）；出生年月；职业；最高学历

本地特色食品访谈题目包括：

1. 你喜欢吃什么?你不喜欢吃什么?
2. 你最喜欢的本地<u>小吃</u>是什么?
3. 你最喜欢的本地<u>特色菜</u>是什么? 哪里的比较正宗?
4. 你会做饭吗? 你有没有什么拿手菜? 你能告诉我做好这道菜的诀窍是什么吗?
5. 你觉得怎么吃才有益健康? 你觉得什么是健康的饮食习惯?

Version II (for people who cannot speak PTH)

Language Study and Environment

1. Where are you from? Which district of this city were you born in? Which dialect do you speak?
2. Have you spoken dialect since you were very young?
   If not, could you please recall from when?
3. Who taught you the dialect?
4. Are your parents able to speak dialect?
5. Are there any other places have you stayed for a long time?
6. Could you understand PTH?

Language Aptitude

1. What other dialects could you speak?
2. How do you think of your dialect aptitude?
3. Do you mix in words of other dialects when you speak dialect? If so, do you do that often?
4. When you come across words you do not use often in dialect, such as neologism, how do you deal with?

Language Use

Do you use dialect in all kinds of occasion? Are there any other exceptions?

Language Attitude

1. Which one do you prefer to hear between PTH and local dialect?
2. Which one more useful for you between PTH and local dialect?
3. Which one is friendlier for you between PTH and local dialect?
4. Which one do you think is more authoritative between PTH and local dialect?
5. Which one do you think is more cordial between PTH and local dialect?
6. Whose speaker do you think is more prestigious between PTH and local dialect?

Personal information

sex; age; occupation; level of education

Food topic

1. What are your favorite local snacks?
2. What's the typical local dish you like most? Where can we find them?
3. Are you able to cook? Do you have a specialty? What's your secret recipe?
4. What is your favorite food and what are your dislikes?
5. What do you think of health diet? What do you think is healthy eating habit? How to eat healthy?

Appendix V Frequency of sandhi patterns in NHD, BSD and SJD

sandhi patterns in NHD			AGE	
			O	Y
T2.T1	24.31	调羹		1
	22.42	曹操	10	10
		馋胚	10	10
		辰光	10	9
		调羹	10	9
		莲心	10	10
		迷宫	10	10
		其中	10	9
		时间	10	9
		台灯	10	10
		提高	10	9
		停当	10	10
		阳台	10	10
		洋机	10	10
T2.T2	24.31	男人	1	
	22.42	胡咙	10	9
		浑堂	10	10
		名堂	10	10
		男人	9	9
		排名	10	9
		排球	10	10
		梧桐	10	10
		洋盘	10	9
T2.T3	24.31	词组	3	
		篮板	1	
	22.42	词组	7	9
		篮板	9	10
		门坎	10	10
		难板	10	10
		人口	9	8
		团长	10	10
		元宝	10	10
T2.T4	24.31	厘米	3	3
	22.24	厘米	3	
		渠道	7	
	22.42	厘米	4	7
		渠道	3	9
T2.T5	22.24	文旦		1
	22.42	残废	10	9
		牢靠	10	9
		文旦	10	9
		咸菜	10	10
T2.T6	24.31	常用	2	2
	22.24	含义	1	0
		和调	1	2
		名字	1	0
		余地	2	0
	22.42	长寿	10	9
		常用	8	6
		含义	9	8
		和调	9	6
		名字	9	9
		闲话	10	9
		余地	8	10
		原旧	10	10
T2.T7	24.31	全国		1
	22.42	湖北	10	9
		毛笔	10	9
		难得	10	10
		全国	10	8
T2.T8	22.42	寒热	10	10
		埋没	10	10
		题目	10	9
T4.T1	24.31	上腔	1	
	22.24	像腔	1	0
		眼泡	0	2
	22.42	罢工	10	10
		懒胚	10	10
		买方	10	9
		上腔	9	10
		午休	10	10
		像腔	9	10
		眼泡	10	8
		雨披	10	10
T4.T2	22.42	冷盆	10	10
		理由	10	9
		领头	10	10
		女人	10	9
		羽毛	10	9
T4.T3	24.31	耳朵	10	7
		户口	9	4
		老虎	9	6
		老酒	9	7
		某种	0	2
	22.24	户口		2
	22.42	耳朵	0	1
		户口	0	4
		老虎	0	1
		老酒	0	3
		某种	5	6
	55.31	某种	5	1
T4.T4	24.31	腐乳	1	4
		蚂蚁	10	7
		上下	8	5
		丈姆	7	6
	22.24	腐乳	8	1
		上下	1	1
		丈姆	1	1
	22.42	腐乳	1	2
		蚂蚁	0	2
		上下	1	4
		丈姆	2	3
T4.T5	24.31	负数	1	0
		鳝片	0	2
		上劲	0	1
		上进	0	1
		上照	0	2

		有劲	0	4
		造句	2	1
	22.24	负数	7	4
		礼拜	9	8
		惹气	7	8
		鳝片	3	3
		上劲	9	8
		上进	9	7
		上照	9	6
		有劲	9	5
		造句	6	9
	22.42	负数	2	5
		礼拜	1	1
		惹气	1	2
		鳝片	7	3
		上劲	1	1
		上进	1	2
		上照	1	1
		有劲	1	0
		造句	2	0
	55.31	鳝片		1
		有劲		1
T4.T6	24.31	道地		1
		上路		3
	22.24	部队	8	8
		道地	9	8
		近视	9	6
		懒惰	8	5
		上路	9	6
	22.42	部队	2	2
		道地	1	1
		近视	1	3
		懒惰	1	5
		上路	1	1
T4.T7	22.42	道德	10	10
		米色	10	10
T4.T8	24.31	序幕		1
	22.24	序幕	5	
	22.42	聚集	10	10
		满月	10	9
		暖热	10	9
		序幕	5	7
	55.31	序幕		1
T6.T1	22.24	便当	1	
	22.42	便当	9	8
		健康	10	7
		闹钟	10	10
		召开	10	6
		自家	9	8
	55.31	召开		4
T6.T2	22.24	调查		1
	22.42	蛋黄	10	9
		调查	10	8
		旧年	10	9
		烂糊	10	9
		弄堂	10	9
		寿头	10	10
		用场	10	9
T6.T3	24.31	具体	9	7
		事体	1	0
		效果	9	8
		字典	1	2
	22.24	面孔	1	0
		事体	1	0
		效果	0	1
		字典	5	1
	22.42	具体	0	3
		面孔	9	10
		事体	8	9
		效果	0	1
		字典	4	6
T6.T4	24.31	导演	0	7
		地点	9	6
		号码	9	10
		谜语	0	1
		具有	9	0
		运道	2	2
	22.24	导演	9	3
		护士	9	5
		具有	0	1
		闹猛	1	0
		运道	7	6
	22.42	导演	1	0
		地点	0	1
		护士	1	3
		谜语	10	8
		具有	0	7
		闹猛	8	9
		运道	1	2
	55.31	号码	1	0
		护士	0	1
T6.T5	24.31	浪费	1	0
		万岁	10	2
	22.24	大蒜	1	3
		浪费	3	6
		卖相	2	4
		万岁	0	4
		夜快	1	3
		硬劲	5	3
	22.42	大蒜	9	5
		浪费	6	4
		卖相	8	6
		万岁	0	4
		夜快	9	6
		硬劲	5	7
T6.T6	24.31	烂饭	0	2
		另外	2	0
		寺庙	1	1
	22.24	烂饭	0	2
		另外	4	5
		寺庙	3	1
		议论	9	5
		预料	9	6
		愿望	4	8
		自愿	7	7
	22.42	烂饭	10	5
		另外	4	4
		寺庙	4	2
		夜饭	10	10
		议论	1	0
		预料	1	3
		愿望	6	0

		自愿	3	2
	55.31	寺庙	2	5
		议论	0	3
		愿望	0	1
T6.T7	24.31	败笔		2
	22.42	败笔	10	7
		硬扎	10	10
		预测	10	10
T6.T8	22.42	附属	10	6
		赖学	10	8
		卖力	10	10
	55.31	附属		1
		赖学		1

sandhi patterns in BSD			AGE	
			O	Y
T2.T1	24.31	曹操	10	1
		馋胚	0	4
		辰光	0	1
		调羹	0	1
		莲心	0	1
		迷宫	0	1
		台灯	0	1
		提高	0	1
		停当	10	7
		阳台	0	1
		洋机	0	1
	22.42	曹操	0	9
		馋胚	10	6
		辰光	10	8
		调羹	10	9
		莲心	10	9
		迷宫	10	8
		其中	10	9
		时间	9	9
		台灯	10	9
		提高	10	9
		停当	0	2
		阳台	10	9
		洋机	10	9
T2.T2	24.31	胡咙	0	1
		浑堂	0	1
		名堂	10	2
		排名	0	4
		排球	0	1
		梧桐	0	2
		洋盘	0	1
	22.42	胡咙	10	9
		浑堂	10	9
		名堂	0	8
		男人	10	10
		排名	10	6
		排球	10	9
		梧桐	10	8
		洋盘	10	9
T2.T3	24.31	词组	9	9
		篮板	8	9
		门坎	9	8
		难板	3	2
		人口	10	10
		团长	9	8
		元宝	10	8
	22.42	词组	0	1
		篮板	2	1
		门坎	1	0
		难板	7	8
		团长	0	1
		元宝	0	1
T2.T4	24.31	厘米	2	9
		渠道	3	3
	22.24	渠道	7	5
	22.42	厘米	7	1
	55.31	渠道		1
T2.T5	24.31	残废	0	5
		牢靠	9	8
		文旦	2	4
		咸菜	10	8
	22.24	残废	10	5
	22.42	牢靠	0	1
		文旦	7	6
		咸菜	0	1
T2.T6	24.31	长寿	9	8
		常用	10	5
		含义	7	3
		和调	10	2
		名字	9	9
		闲话	8	8
		余地	3	8
		原旧	8	4
	22.24	含义	0	1
		和调	0	3
		闲话	0	2
		余地	7	2
	22.42	长寿	1	2
		常用	0	3
		含义	2	5
		和调	0	5
		原旧	2	4
	55.31	名字	1	1
T2.T7	24.31	湖北	8	3
		毛笔	8	2
		难得	2	1
		全国	9	7
	22.42	湖北	2	6
		毛笔	1	5
		难得	3	9
T2.T8	24.31	寒热	10	5
		埋没	8	7
		题目	9	2
	22.42	寒热	0	5
		埋没	1	3
		题目	1	8
T4.T1	24.31	罢工		3
		懒胚		1
		上腔		1
	22.24	懒胚		1
		眼泡		4
	22.42	罢工	10	6
		懒胚	10	8
		买方	10	10
		上腔	9	9
		午休	4	7
		像腔	10	9
		眼泡	10	6

		雨披	10	9
	55.31	午休	6	3
		雨披	0	1
T4.T2	22.24	羽毛		3
	22.42	冷盆	10	10
		理由	10	8
		领头	10	10
		女人	9	10
		羽毛	0	3
	55.31	理由	0	2
		羽毛	9	4
T4.T3	24.31	户口	10	8
		老虎	4	6
		老酒	8	4
		某种	3	4
	22.24	老酒		2
		某种		1
	22.42	耳朵	10	10
		户口	0	2
		老虎	4	3
		老酒	0	1
		某种	6	3
	55.31	某种	1	
T4.T4	24.31	腐乳	1	2
		蚂蚁	10	8
		上下	10	10
		丈姆	2	5
	22.24	腐乳	1	2
		蚂蚁	0	1
		丈姆	0	1
	22.42	腐乳	8	3
		丈姆	8	2
	55.31	腐乳		2
		蚂蚁		1
		丈姆		2
T4.T5	24.31	负数	6	4
		惹气	0	1
		鳝片	5	4
		上劲	9	5
		上进	5	4
		上照	2	4
		有劲	6	3
		造句	0	5
	22.24	负数	3	3
		礼拜	10	9
		惹气	10	9
		鳝片	4	4
		上劲	1	4
		上进	5	6
		上照	8	5
		有劲	4	7
		造句	9	5
	22.42	负数	1	0
		礼拜	0	1
		造句	1	0
	55.31	负数		3
		鳝片		1
T4.T6	24.31	部队	0	4
		道地	1	3
		近视	0	2
		懒惰	0	2
		上路	5	8
	22.24	部队	9	5
		道地	9	7
		近视	0	3
		懒惰	10	2
		上路	5	2
	22.42	近视	10	5
		懒惰	0	5
T4.T7	24.31	道德	7	3
		米色	7	1
	22.42	道德	1	4
		米色	2	6
T4.T8	24.31	聚集	9	1
		满月	9	3
		暖热	10	5
		序幕	4	6
	22.24	序幕	5	
	22.42	聚集	1	7
		满月	1	7
		暖热	0	5
		序幕	0	1
	55.31	序幕		1
T6.T1	24.31	便当	10	7
		健康	0	2
	22.42	便当	0	3
		健康	10	8
		闹钟	10	10
		召开	9	0
		自家	10	8
	55.31	召开		8
T6.T2	24.31	烂糊	5	2
		寿头	0	2
		用场	0	2
	22.42	蛋黄	10	10
		调查	10	10
		旧年	9	10
		烂糊	4	8
		弄堂	9	10
		寿头	10	8
		用场	10	8
T6.T3	24.31	具体	6	4
		面孔	10	5
		事体	9	6
		效果	9	8
		字典	10	8
	22.24	具体	2	5
	22.42	面孔		4
		事体		2
		效果		2
		字典		1
T6.T4	24.31	导演	4	3
		地点	10	7
		号码	9	8
		护士	0	2
		闹猛	10	3
		谜语	10	9
		运道	9	6
	22.24	导演	6	5
		护士	10	8
		运道	0	2
	22.42	地点	0	2
		号码	0	2
		具有	10	10
		闹猛	0	7
		运道	0	2
	55.31	导演		2
T6.T5	24.31	大蒜	9	7

		浪费	10	9
		卖相	10	1
		万岁	10	4
		夜快	0	1
		硬劲	8	9
	22.24	浪费	0	1
		卖相	0	4
		万岁	0	4
		夜快	0	3
		硬劲	1	1
	22.42	大蒜	0	1
		卖相	0	3
		夜快	1	5
	55.31	夜快	9	1
T6.T6	24.31	烂饭	8	10
		寺庙	9	5
		夜饭	9	7
		议论	0	3
		预料	8	4
		愿望	10	2
		自愿	8	9
	22.24	烂饭	1	0
		另外	10	10
		夜饭	0	1
		议论	9	3
		预料	2	4
		愿望	0	5
	22.42	寺庙	0	1
		夜饭	1	2
		愿望	0	2
		自愿	1	1
	55.31	寺庙	1	4
		议论	0	3
		预料	0	2
		愿望	0	1
T6.T7	24.31	败笔	7	4
		硬扎	10	5
		预测	9	3
	22.24	预测		1
	22.42	败笔	3	4
		硬扎	0	4
		预测	1	6
T6.T8	24.31	附属	7	3
		赖学	7	3
		卖力	8	2
	22.42	附属	3	1
		赖学	2	7
		卖力	2	8
	55.31	附属		4

sandhi patterns in SJD			AGE	
			O	Y
T2.T1	24.31	曹操	7	0
		馋胚	1	0
		停当	2	3
	22.42	曹操	2	10
		馋胚	9	9
		辰光	10	10
		调羹	10	10
		莲心	10	10
		迷宫	10	10
		其中	9	10
		时间	10	10
		台灯	10	10
		提高	10	10
		停当	8	7
		阳台	9	10
		洋机	10	10
T2.T2	24.31	浑堂	2	0
		梧桐	0	1
	22.42	胡咙	10	10
		浑堂	8	10
		名堂	10	10
		男人	10	10
		排名	9	8
		排球	10	10
		梧桐	10	9
		洋盘	10	9
T2.T3	24.31	词组	1	
	22.42	词组	9	10
		篮板	10	10
		门坎	10	10
		难板	10	10
		人口	9	10
		团长	10	10
		元宝	10	10
T2.T4	24.31	厘米	3	1
		渠道	0	5
	22.24	渠道	10	5
	22.42	厘米	7	9
T2.T5	24.31	残废	8	7
		牢靠	9	7
		文旦	8	4
		咸菜	10	7
	22.42	残废	2	1
		牢靠	0	2
		文旦	2	6
		咸菜	0	2
T2.T6	24.31	长寿	10	8
		常用	10	9
		含义	6	10
		和调	8	4
		名字	10	10
		闲话	10	8
		余地	9	8
		原旧	7	7
	22.24	含义	4	
	22.42	长寿	0	2
		和调	2	6
		闲话	0	2
		余地	1	2
		原旧	3	3
T2.T7	24.31	湖北	10	10
		毛笔	10	10
		难得	10	9
		全国	10	8
T2.T8	24.31	寒热	10	10
		埋没	1	10
		题目	10	10
	55.31	埋没	9	
T4.T1	24.31	罢工	7	7
		懒胚	7	7
		买方	9	5
		上腔	8	3

		午休	2	7
		像腔	8	4
		眼泡	1	0
		雨披	1	0
	22.24	上腔	0	1
		眼泡	0	5
		雨披	7	4
	22.42	罢工	1	0
		懒胚	2	3
		买方	1	5
		上腔	2	6
		午休	6	3
		像腔	2	5
		眼泡	9	5
		雨披	2	5
	55.31	罢工	0	2
		午休	2	0
T4.T2	24.31	冷盆	10	5
		理由	10	7
		领头	10	8
		女人	9	9
		羽毛	0	4
	22.24	羽毛		4
	22.42	冷盆	0	4
		理由	0	3
		领头	0	2
		女人	1	1
		羽毛	10	2
T4.T3	24.31	耳朵	10	7
		户口	9	7
		老虎	7	8
		老酒	10	7
		某种	8	10
	22.42	耳朵	0	2
		户口	1	3
		老酒	0	3
		某种	2	0
T4.T4	24.31	腐乳	2	3
		蚂蚁	10	9
		上下	0	3
		丈姆	9	5
	22.24	腐乳	4	6
		上下	10	7
	22.42	腐乳	4	0
		蚂蚁	0	1
		丈姆	1	5
	55.31	腐乳		1
T4.T5	24.31	鳝片		1
		上照		1
	22.24	礼拜	6	7
		惹气	3	10
		上劲	0	7
		上进	8	10
		上照	1	6
		造句	10	10
	22.42	负数	9	9
		礼拜	4	3
		惹气	7	0
		鳝片	10	9
		上劲	10	3
		上进	2	0
		上照	9	3
		有劲	10	10

T4.T6	24.31	道地	1	
	22.24	部队	7	7
		道地	9	10
		近视	10	10
		懒惰	10	10
	22.42	部队	3	3
		上路	10	10
T4.T7	24.31	道德	6	7
		米色	9	8
	22.42	道德		1
		米色		1
T4.T8	24.31	聚集	10	9
		满月	10	8
		暖热	10	7
		序幕	0	7
	22.42	聚集	0	1
		满月	0	2
		暖热	0	3
		序幕	10	3
T6.T1	24.31	召开	5	3
	22.24	便当	10	10
		闹钟	0	2
	22.42	健康	10	9
		闹钟	10	8
		召开	5	7
		自家	10	8
T6.T2	24.31	蛋黄	10	8
		调查	0	4
		旧年	3	6
		烂糊	1	4
		弄堂	8	8
		寿头	8	6
		用场	10	9
	22.24	旧年	7	4
		烂糊	9	6
		弄堂	2	2
		寿头	2	4
	22.42	蛋黄	0	2
		调查	10	6
		用场	0	1
T6.T3	24.31	具体	8	4
		面孔	10	9
		事体	7	4
		字典	0	1
	22.24	具体	1	0
		效果	10	10
		字典	6	5
	22.42	具体	1	6
		面孔	0	1
		事体	3	6
		字典	4	4
T6.T4	24.31	导演	1	6
		地点	4	1
		号码	10	9
		护士	1	0
		具有	10	0
		谜语	9	3
		闹猛	9	9
	22.24	导演	9	2
		护士	9	10
		运道	10	10
	22.42	导演	0	2
		地点	5	8

		号码	0	1
		具有	0	10
		谜语	1	7
		闹猛	1	1
		阳台	1	0
T6.T5	22.24	大蒜	5	6
		浪费	10	10
		卖相	0	5
		万岁	10	10
		硬劲	10	10
	22.42	大蒜	5	4
		卖相	10	5
		夜快	1	7
	55.31	夜快	9	3
T6.T6	24.31	寺庙	3	0
		愿望	0	1
	22.24	烂饭	10	10
		另外	10	10
		寺庙	1	9
		夜饭	4	6

		议论	9	5
		预料	10	10
		愿望	10	9
		自愿	10	10
	22.42	寺庙	6	1
	55.31	夜饭	6	4
		议论	0	5
T6.T7	24.31	败笔	10	7
		硬扎	9	7
		预测	9	6
	22.42	硬扎		3
		预测		4
	55.31	败笔		2
T6.T8	24.31	附属	9	5
		赖学	10	9
		卖力	10	9
	22.42	附属	1	3
		赖学	0	1
		卖力	0	1

Samenvatting

Dit boek onderzoekt de rol van sociale, stilistische en linguïstische factoren in tonale variatie en verandering in Wú dialecten zoals die gesproken worden in Shangai en Wuxi, twee steden gelegen aan de monding van de Yangtze rivier. Deze regio ligt in het zuidoosten van China en is het hart van de Chinese economische expansie, waar de verstedelijking zich met een ongeziene omvang en snelheid voltrekt. Het onderzoek is een zogenaamde 'urban language study' (Xu, 2006), waarbij inzichten en technieken uit de sociolinguïstiek, de fonetiek en de dialectologie gecombineerd worden. Wereldwijd is het een van de eerste sociofonetische studies naar veranderingen in toonsystemen, waardoor het onderzoek ook belangrijke methodologische implicaties heeft.

In het eerste hoofdstuk wordt de achtergrond van het onderzoek geschetst en ingegaan op de relatie tussen verstedelijking en taalverandering (1.1). Vervolgens worden de onderzoeksdoelstellingen voorgesteld (1.2). Deze studie heeft twee algemene doelstellingen: een beschrijving van de tonale veranderingen in de dialecten van Shanghai en Wuxi en het blootleggen van de verschillende factoren die deze veranderingen sturen. Daarnaast heeft het onderzoek twee methodologische doelstellingen. In de eerste plaats moet vastgesteld worden over welk deel van een woord de tonen geanalyseerd moeten worden in de bestudeerde Wú dialecten. In de tweede plaats moet een normaliseringsmethode ontwikkeld worden die ons in staat stelt de metingen van het toonhoogteverloop van verschillende sprekers met elkaar te vergelijken. Het eerste hoofdstuk wordt afgesloten met een schets van de rest van de dissertatie (1.3).

Hoofdstuk 2 introduceert het notatiesysteem voor tonen dat in deze dissertatie gebruikt wordt (2.1). We hebben gekozen voor Chao's numeriek systeem in plaats van de in het IPA gebruikte toonsymbolen. Vervolgens wordt in 2.2 het Wú dialectgebied geschetst. In 2.3 gaan we in op het toonsysteem van het Middelchinees en op de historische toonverandering Zhuó Shǎng Guī Qù (浊上归去). Het hoofdstuk wordt afgesloten (2.4) met een schets van de taalgemeenschappen in Wúxī en Shànghǎi en een overzicht van de daar eerder geattesteerde toonveranderingen.

In hoofdstuk 3 wordt de onderzoeksmethodologie toegelicht. We schetsen eerst het onderzoeksdesign. De kenmerken van de deelnemers aan dit onderzoek worden beschreven in 3.1. Daarbij wordt onder andere ingegaan op de factoren in het onderzoeksdesign (sekse, leeftijd, stad en mate van verstedelijking) en enkele andere kenmerken van de deelnemers aan het onderzoek (zoals socio-economische status, opleidingsniveau en etniciteit). Vervolgens beschrijven we hoe de deelnemers aan dit onderzoek geselecteerd zijn (3.2) en hoe de opnames gemaakt zijn (3.3). In 3.4 staan we uitgebreid stil bij het verzamelde spraakmateriaal, en de linguïstische factoren die in het onderzoeksdesign opgenomen zijn. We sluiten dit hoofdstuk af in 3.5 met een korter toelichting op de statistische analyses.

Hoofdstuk 4 besteedt uitgebreid aandacht aan de akoestische analyse van toon en aan de twee methodologische doelstellingen van dit onderzoek: het tonale domein en toonnormalisatie. In toontalen zijn de tonale kenmerken (de trilfrequentie van de stemplooien) gebonden aan het stemhebbende deel van een syllabe. Maar, niet alle stemhebbende delen van een syllabe dragen toon. Bovendien is het van belang dat de metingen van de verschillende dialecten uit dit onderzoek vergelijkbaar zijn. Daartoe dienen we eerst te achterhalen in welk deel van de syllabe de F_0 gemeten moet worden. Om dit te achterhalen zijn in een deel van de data (uit zowel Wúxī als Shànghǎi) de F_0 contouren en de duren van de syllabes gemeten en systematisch met elkaar vergeleken. Hieruit blijkt dat de syllabe-initiële stemhebbende consonant geen cruciale tooninformatie bevat, maar de nasale slotconsonant in de coda wel. De prenucleaire halfvocaal bevat enkel belangrijke tooninformatie in Wúxī, maar niet in Shànghǎi. Maar om de data uit beide steden goed te kunnen vergelijken is het nodig het tonale domein in de Wú dialecten te definiëren als het deel van de syllabe dat bestaat uit de (optionele) prenuclear onglide, de kern en de (optionele) nasaal in de coda.

Om valide uitspaken te kunnen doen over tonale verschillen tussen sprekers is het nodig de F_0 te normaliseren. Wij hebben ons onderzoek naar de evaluatie van verschillende normalisatieprocedures gebaseerd op gelijksoortig onderzoek naar normalisering van formantwaarden van klinkers (voor een uitgebreid overzicht, zie Van der Harst, 2011). Een goede normaliseringsprocedure moet (1) de fonematische variatie (dit zijn de verschillen tussen de tonen) bewaren, (2) de door anatomische verschillen veroorzaakte verschillen weghalen en (3) de sociolinguïstische variatie behouden. Uiteindelijk hebben we voor ST-AvgF_0 (een semitoon transformatie op basis van de gemiddelde F_0 in Hz) gekozen omdat deze normalisatieprocedure niet allen goed presteert maar ook voor vergelijkingen tussen talen kan worden gebruikt.

Hoofdstukken 5 en 6 focussen op de veranderingen in de tonen 2, 4 en 6 in respectievelijk Wúxī en Shànghǎi. Daarbij is telkens eerst op zoek gegaan naar de tonale variabelen die in een veranderingsproces betrokken zijn. Vervolgens zijn de factoren geanalyseerd die deze veranderingen sturen. In citatietoon (toon in losse woorden) zijn dat: (1) het samenvallen van ongespannen tonen in het Yáng register, (2) contourverlies; (3) DelayRising en (4) toonontlening uit Pǔtōnghuà. In woordgroepen (tone sandhi) zijn dat (1) contourverlies van het sandhipatroon en (2) de ontlening van het /55.31/ patroon van het Yīn register. Het samenvallen van ongespannen tonen in het Yáng register en *DelayRising* worden gestuurd door taalinterne factoren, terwijl contourverlies en toonontlening (in zowel citatietoon als toonsandhi) door externe factoren gestuurd worden. De dialecten uit de minder verstedelijkte gebieden convergeren met de stadskerndialecten bij de taalinterne veranderingen, maar divergeren ervan als het om door externe factoren gestuurde veranderingen gaan. We maken in dit onderzoek een onderscheid tussen twee taalinterne drijfveren: het economiciteitsprincipe en tonale convergentie. Het economiciteitsprincipe zorgt ervoor dat tonale vormen fonetisch gereduceerd worden, tonale convergentie is dan weer verantwoordelijk voor fonematische reductie. Beide taalinterne principes resulteren in een vereenvoudiging van het toonsysteem, wat een algemene tendens is in de Wú dialecten. Daarnaast bracht het onderzoek ook nog een reeks andere effecten aan het licht die van invloed zijn op de realisatie van tonen zoals spreekstijl, regio, sekse, de tooncategorie in het Middelchinees en de tooncategotie in het Pǔtōnghuà.

Hoofdstuk 7 biedt een samenvatting en discussie van de factoren die tonale variatie en verandering in Wúxī en Shànghǎi sturen. We gaan daarbij ook in op de methodologische en theoretische consequenties voor de Chinese dialectologie, ‘urban language studies’ en de sociolinguïstiek. De dissertatie wordt afgesloten met enkele perspectieven voor verder onderzoek. Deze studie heeft nieuwe variatiepatronen in de Chinese dialecten aan het licht gebracht en ons in staat gesteld om bestaande inzichten bij te stellen. De talige gevolgen van de verstedelijking in China kunnen slechts adequaat bestudeerd worden door een combinatie van inzichten uit de sociofonetiek en de dialectologie. Op het terrein van ‘urban language studies’ laat dit onderzoek op overtuigende wijze zien dat het nodig is om de minder verstedelijkte gebieden in het onderzoek te betrekken en die te bestuderen in relatie tot de stadskernen. Tevens doen we de aanbeveling om binnen meer sociofonetisch onderzoek te doen. Deze dissertatie heeft tenslotte paden uitgezet voor onderzoek naar taalvariatie en taalverandering in grote (Chinese) steden en de sociofonetische analyse van toon..

Curriculum vitae

Jǐngwěi Zhāng (张璟玮) was born on May 14, 1985 in Wúxī, China, where she grew up. In 2003 she started her undergraduate study in liberal arts in the Kuang Yaming Honors School of Nanjing University. She studied Chinese Language and Literature, History and Philosophy and finally got interested in Chinese linguistics. She started her linguistic research in 2005 at the Sociolinguistic Laboratory of Nanjing University. After getting her bachelor's degree in 2007, she continued to study at the same university, where she obtained her Master's degree in linguistics at the School of Liberal Arts in 2009. With a grant from the China Scholarship Council, she started in October 2009 her PhD research at the Utrecht Institute of Linguistics OTS, of which this dissertation is the result.

In addition to her sociophonetic research, Jǐngwěi Zhāng has also been involved in research on language policy at the China Centre for Linguistic and Strategic Studies of Nanjing University since 2007.