

LOCAL IDENTITY AND LANGUAGE ATTITUDE IN STANDARDIZATION:
EVIDENCE FROM TIANJIN CHINESE TONE SANDHI

By

Xiaomei Wang

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Linguistics—Doctor of Philosophy

2020

ABSTRACT

LOCAL IDENTITY AND LANGUAGE ATTITUDE IN STANDARDIZATION: EVIDENCE FROM TIANJIN CHINESE TONE SANDHI

By

Xiaomei Wang

This dissertation investigates the roles of local identity, language attitude, social awareness, as well as social meanings in dialect change by examining Tianjin Chinese tone sandhi in apparent time. It studies the process by which local variants become stigmatized and also the process by which local features increase. Tianjin Chinese is in the process of standardization (Gao & Lu, 2003; Gu & Liu, 2003), but the current study finds that only stigmatized local features are disappearing, while an unmarked local feature seems to be immune to standardization. I interpret this in line with Labov's (1972) study of Martha's Vineyard, whereby traditional local features may come to index resistance to standardization and to the incursion of new people into the speech community.

Ninety sociolinguistic interviews, including a word list, were conducted in Tianjin in the local dialect in 2014-16 (48f, age 18-82). Participants were categorized as 'middle class' or 'working class' using a combined measure of occupation, education and income. Qualitative assessments of 'positive' or 'negative' were assigned to speakers' attitudes to Tianjin and to migrants.

The variables of the dissertation are three of the four traditional Tianjin tone sandhi, referred to as (53-53), (53-21) and (21-21) after their input tonal values respectively. Application of the (53-53) and (53-21) variables produce local outputs; non-application makes a speaker sound more like a Standard Chinese speaker. The old output variant of (21-21) is traditional; the new variant is closer to Standard Chinese.

7462 tokens of (53-53), 5683 tokens of (53-21), and 4117 tokens of (21-21) were extracted from the interviews and word lists (N = 17262). Tokens were impressionistically coded for the application or non-application of (53-53) and (53-21), and for the new or old variant of (21-21), and a subsample were checked in Praat.

(53-53) application and the old variant of (21-21) have decreased substantially in frequency over time, probably because their outputs include or are similar to the most salient Tianjin low tone: Tone 1 (Han 1993). Mixed effects regression shows that they are avoided in word list style, by the middle class, by women, and by people with a negative attitude to Tianjin.

In contrast, (53-21) has increased from 73.5% among speakers aged 65+, to 93% among speakers under 65. I speculate that because (53-21) is below public awareness, with little style-shifting, it is available for ‘recycling’ (Dubois & Horvath 2000) as a positive marker of ‘new’ Tianjin identity. Tianjin natives may be linguistically contrasting themselves with the many migrants who have moved to the city in the last three decades, and indeed a negative attitude to migrants significantly increases the likelihood of (53-21) application in the regress.

The reduction or elimination of prominent stereotyped dialect features has been observed in languages with large numbers of speakers, often because of migration and language contacts (Hinskens, 1998). Unmarked dialect features have been observed to persist or increase to keep local identity (Labov 1972, Dubois & Horvath 2000). These two conflicting forces might lead to a stable compromise dialect (Hinskens, 1998). Here the case study of Tianjin Chinese tone sandhi also exhibits signs of changing to a compromise dialect, with stereotyped local features disappearing and unmarked local features increasing, adding to the expanding number of non-Western case studies of language change (Stanford & Preston, 2009) that support earlier generalizations made from Western communities.

Copyright by
XIAOMEI WANG
2020

To my mother

ACKNOWLEDGMENTS

My gratitude goes to all the people who guided and encouraged me throughout my graduate study. First, I would like to express my upmost gratitude to my co-chairs, Professor Yen-Hwei Lin and Professor Suzanne Evans Wagner. I have been richly blessed to have worked with them. I cannot thank them enough for their knowledge, inspiration, patience and support. They have provided tremendous support at every stage of my graduate study along with helping me immensely with my dissertation. This study would not have been complete without their creative ideas, thought-provoking comments, corrections and revisions.

I am also indebted to my committee members, Professor Karthik Durvasula, Professor Deogratias Ngonyani and Professor Anne Violin-Wigent, who have offered their invaluable time and expertise in helping me.

I would like to thank Michigan State University for supporting me to complete my graduate study with generous grants and fellowships.

I would like to thank Professor Jilun Lu (Tianjin Normal University), Wenjie Guo, Professor Ruwei Tan (Tianjin Normal University) and Professor Hong Chen (Tianjin University) for helping me recruit participants.

Last but not least, I would like to express my everlasting gratitude to my dearest mom and dad, Yanxiu Liu (刘炎秀) and Jinquan Wang (王锦泉), who believed in me and gave me their unconditional love. I also want to thank my husband, Wei Wang (王伟), who has been there supporting me throughout this whole adventure with love, humor and wits.

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xiii
Chapter 1 Introduction	1
1.1 The variables, research questions, and hypotheses of the current study.....	2
1.1.1 The sandhi variables	2
1.1.2 Research questions.....	5
1.1.3 Hypotheses.....	5
1.2 Standardization, vernacular maintenance and migration	7
1.3 Dissertation outline	13
Chapter 2 The socio-economic environment of Tianjin dialect.....	15
2.1 Brief history of Tianjin city	16
2.2 Tianjin industry, economy and migration.....	18
2.3 The renovation of the urban area	22
Chapter 3 Tianjin dialect.....	27
3.1 Tianjin lexical tones.....	27
3.2 Tianjin disyllabic tone sandhi	28
3.3 Language environment of Tianjin dialect.....	29
3.4 Promotion of Putonghua	31
3.5 Attitudes to SC and Tianjin.....	32
3.6 Existing studies on Tianjin dialect.....	35
Chapter 4 Data collection and methods	41
4.1 Speaker sample	43
4.1.1 Sampling method	43
4.1.2 Eligibility criteria.....	45
4.1.3 Demographic characteristics.....	45
4.1.3.1 Age.....	46
4.1.3.2 Social class.....	47
4.1.4 Attitude to migrants and attitude to Tianjin dialect	49
4.1.4.1 Attitude to migrants	49
4.1.4.2 Attitude to Tianjin dialect	52
4.2 Interview and word list	54
4.2.1 The interview	56
4.2.2 The word list	58
4.3 Methodology.....	60
4.3.1 Impressionistic coding and inter-rater reliability.....	60
4.3.2 Extraction of tokens	65
4.3.3 Statistical analysis.....	66

Chapter 5 Word list Results	67
5.1 A general introduction to word list results.....	67
5.2 Regression model construction	70
5.3 The word list results of (53-53)	73
5.3.1 The effect of age group	75
5.3.2 The effect of social class.....	76
5.3.3 The effect of gender	78
5.3.4 The effect of attitudes to Tianjin dialect	80
5.4 The word list results of (53-21)	84
5.4.1 The effect of age group	85
5.4.2 The effect of attitudes to migrants	86
5.5 The word list results of (21-21)	88
5.5.1 The effect of age group	90
5.5.2 The effect of attitudes to Tianjin dialect	93
5.6 Summary	94
Chapter 6 Interview Results.....	96
6.1 A general introduction to interview results.....	96
6.2 Regression model construction	99
6.3 Factors predicting the application of the (53-53) sandhi rule	101
6.3.1 The effect of age group	103
6.3.2 The effect of social class.....	104
6.3.3 The effect of gender	106
6.4 Factors predicting the application of the (53-21) sandhi rule	108
6.4.1 The effect of age group	110
6.4.2 The effect of attitudes to migrants	111
6.4.3 The effect of social class.....	113
6.5 Factors predicting the application of the (21-21) sandhi rule	116
6.5.1 The main effects of age group	117
6.5.2 The main effects of social class	119
6.6 Summary	122
Chapter 7 Analysis and discussion	124
7.1 Style and sociolinguistic variation	124
7.2 ‘Change from above’ and ‘Change from below’	126
7.3 ‘Indicator’, ‘marker’ and ‘stereotype’	128
7.4 The change in (53-53).....	130
7.4.1 (53-53): a decrease from above.....	130
7.4.2 LL: a stereotype avoided in both tone sandhi and lexical tone.....	133
7.5 The change in (53-21).....	138
7.5.1 (53-21): an increase from below	138
7.5.2 (53-21): an increase in local variant led by people with negative attitude to migrants.....	140
7.5.3 (53-21): recycling of an unmarked local variant to mark Tianjin identity	142
7.6 The change in (21-21).....	146
7.6.1 (21-21): a change from above led by middle class speakers	146
7.6.2 Linguistic motivations for the emergence of the new variant.....	153

7.6.2.1 Metrical structure, tonal complexity and tonal stability	153
7.6.2.2 The rising in T1 and the emergence of the new variant.....	155
7.6.3 Attitude to Tianjin dialect in the emergence of the new variant.....	157
7.7 Summary	159
Chapter 8 Conclusions and future directions	161
APPENDICES	167
APPENDIX A. Interview questions	168
APPENDIX B. Word list	175
APPENDIX C. Language attitude survey results	178
BIBLIOGRAPHY	181

LIST OF TABLES

Table 2.1 Population of different districts in Tianjin, selected years from 1980 to 2017(10,000persons) (Source: Tianjin statistical bureau, 2018)	24
Table 3.1 Lexical tones in Tianjin dialect and SC	27
Table 3.2 The tone sandhi rules in Tianjin and SC.....	28
Table 3.3 'Special consonants' (Type I)	36
Table 3.4 'Special consonants' (Type II).....	36
Table 3.5 'Special consonants' (Type III).....	36
Table 3.6 'Special consonants' (Type IV)	37
Table 4.1 Entire participant sample by age, social class and gender (N=90)	46
Table 4.2 The distribution of participants by attitude to migrants and age group	50
Table 4.3 The distribution of participants by attitude to migrants and social class.....	50
Table 4.4 The distribution of participants by attitude to migrants and gender	50
Table 4.5 The distribution of participants by attitude to migrants and attitude to Tianjin	51
Table 4.6 Scoring guide for the questions about Tianjin dialect	53
Table 4.7 The distribution of participants by attitude to Tianjin and age group	54
Table 4.8 The distribution of participants by attitude to Tianjin and social class	54
Table 4.9 The distribution of participants by attitude to Tianjin and gender	54
Table 5.1 Rates of sandhi rule application for (53-53), (53-21) and (21-21) in word list style, by age group, all speakers (N=90)	68
Table 5.2 Fixed effects of generalized mixed model of (53-53) in word list	74
Table 5.3 The application frequency of LL variant in word list by age group	75
Table 5.4 The application frequency of LL variant by age and social class.....	77
Table 5.5 The application frequency of LL variant in word list by gender and age group	78
Table 5.6 The distribution of LL variant in word list by social class and gender	79

Table 5.7 The application frequency of LL variant in word list by attitude to Tianjin and age group	80
Table 5.8 The application frequency of LL variant in word list by social class and attitude to Tianjin	82
Table 5.9 The distribution of LL variant in word list by gender and attitude to Tianjin	83
Table 5.10 Fixed effects of generalized mixed model of (53-21) word list	84
Table 5.11 The application frequency of the HH variant by age group	85
Table 5.12 The application means of the HH variant by attitude to migrants	87
Table 5.13 Distribution of HH variant in word list by attitude to migrants and age group	88
Table 5.14 The fixed effects of generalized mixed model of (21-21) in word list	89
Table 5.15 The application frequency of new HH variant in word list by age group	90
Table 5.16 The application frequency of old LH variant in word list by age group	91
Table 5.17 The application frequency of HH variant in word list by attitude to Tianjin and age group	93
Table 5.18 The application frequency of LH variant in word list by attitude to Tianjin and age group	93
Table 6.1 Rates of sandhi rule applications of (53-53), (53-21) and (21-21) in interview style, by age group, all speakers (N=90)	98
Table 6.2 The parameter estimates for the generalized mixed model of (53-53) interview	102
Table 6.3 The application frequency of LL variant in interview by age group	103
Table 6.4 The application frequency of LL variant by social class	104
Table 6.5 The application frequency of LL variant by social class and age group	106
Table 6.6 The application frequency of LL variant by gender	107
Table 6.7 Application frequency of LL variant in interview by social class and gender	108
Table 6.8 The parameter estimates for the generalized mixed model of (53-21) in interview style	109
Table 6.9 The application frequency of HH variant by age group	110
Table 6.10 The application frequency of HH variant by attitude to migrants	111

Table 6.11 The application frequency of HH variant by attitude to migrants and age group	112
Table 6.12 The application frequency of HH variant by social class	113
Table 6.13 The application frequency of HH variant by social class and age group	114
Table 6.14 The application frequency of HH variant by social class and attitude to migrants ..	115
Table 6.15 The fixed effects for the generalized mixed model of (21-21) interview	116
Table 6.16 The application frequency of HH variant by age group	118
Table 6.17 The application frequency of LH variant by age group.....	118
Table 6.18 The application frequency of HH variant by social class	120
Table 6.19 The application frequency of LH variant by social class.....	120
Table 6.20 Distribution of HH variant by age group and social class	122
Table 7.1 Tonal values of Tianjin lexical tones in previous and present study	135
Table 7.2 The distribution of HH variant by attitude to migrants and age group in both styles	141
Table 7.3 The distribution of (53-21) by attitude to migrants and social class in both styles	142
Table 7.4 The distribution of HH variant by age group, social class and style	151
Table 7.5 The distribution of HH variant by age group, attitude to Tianjin and style.....	158

LIST OF FIGURES

Figure 2.1 Administrative division map of Tianjin (Retrieved from https://baike.sogou.com/historylemma?lId=6808&cId=173664100 on July 15, 2019)	16
Figure 2.2 Annual GDP growth rates of Tianjin and the whole country from 1978 to 2017 (Tianjin statistical bureau, 2018)	19
Figure 2.3 Tianjin population from 1978 to 2017 (10,000 persons) (Tianjin statistical bureau 2018)	20
Figure 2.4 Layout of city planning of Tianjin (2005-2020) (Source: City construction committee of Tianjin 2005; Retrieved from http://www.cityup.org/case/general/20090114/44146.shtml) ..	23
Figure 3.1 Tianjin and neighboring dialects (Li & Han 1991)	30
Figure 4.1 Tianjin lexical tones in raw Hertz values, 16 young speakers	61
Figure 4.2 Tianjin lexical tones in normalized results, 16 young speakers	62
Figure 4.3 The output of T4+T4 → T1+T4 in raw Hertz	63
Figure 4.4 The output of T4+T1 → T2+T1 in raw Hertz	63
Figure 4.5 The output of T1+T1 → T2+T1 in raw Hertz	64
Figure 4.6 The output of T3+T3 → T2+T3 in raw Hertz	64
Figure 5.1 Rate of use by variant for (53-53), (53-21) and (21-21) in word list style, by age group, all speakers (N=90).....	69
Figure 5.2 Application frequency of LL variant in word list by age group.....	75
Figure 5.3 Application frequency of LL variant in word list by social class.....	77
Figure 5.4 Application frequency of HH variant in word list by age group.....	86
Figure 5.5 Application frequency of HH variant in word list by attitude to migrants.....	87
Figure 5.6 Application frequency of HH variant in word list by age group	91
Figure 5.7 Application frequency of LH variant in word list by age group	92
Figure 5.8 Application frequency of HH variant in word list by attitude to Tianjin	94
Figure 5.9 Application frequency of LH variant in word list by attitude to Tianjin.....	94

Figure 6.1 Rate of use by variant for (53-53), (53-21) and (21-21) in interview style by age (N=90).....	98
Figure 6.2 Application frequency of LL variant in interview by age group.....	103
Figure 6.3 Application frequency of LL variant in interview by social class.....	105
Figure 6.4 The application frequency of LL variant in interview by gender.....	107
Figure 6.5 Application frequency of HH variant in interview by age group.....	110
Figure 6.6 Application frequency of HH variant in interview by attitude to migrants.....	112
Figure 6.7 Application frequency of HH variant in interview by social class.....	114
Figure 6.8 Application frequency of HH variant in interview by age group.....	118
Figure 6.9 Application frequency of LH variant in interview by age group	119
Figure 6.10 Application frequency of HH variant in interview by social class.....	121
Figure 6.11 Application frequency of LH variant in interview by social class	121
Figure 7.1 Application rates of LL variant in word list and interview styles by age group	131
Figure 7.2 Application rates of LL variant by social class and style.....	132
Figure 7.3 Application rates of LL variant by gender and style	132
Figure 7.4 The normalized F0 results of T1 in young and old groups.....	136
Figure 7.5 The normalized F0 results in young and old groups	136
Figure 7.6 The normalized F0 results of T3 in young and old groups.....	137
Figure 7.7 The normalized F0 results of T4 in young and old groups.....	137
Figure 7.8 Application rates of HH variant in word list and interview styles by age group	138
Figure 7.9 Application rates of HH variant by social class and style.....	139
Figure 7.10 The application rates of HH variant by attitude to migrants and style.....	140
Figure 7.11 Application rates of the old LH variant in word list and interview styles.....	147
Figure 7.12 Application rates of the new HH variant in word list and interview styles.....	148
Figure 7.13 Application rates of LH variant by social class and style	149
Figure 7.14 Application rates of HH variant by social class and style	149

Figure 7.15 Application rates of LH variant by attitude to Tianjin and style	150
Figure 7.16 Application rates of HH variant by attitude to Tianjin and style	150

Chapter 1 Introduction

This dissertation is on diachronic change in three tone sandhi patterns in Tianjin Chinese, a Mandarin dialect, from a variationist sociolinguistic perspective (Tagliamonte, 2012). This approach to the study of language change and variation focuses on social variation in languages and examines how this variation is structured. In the present study, spontaneous speech in sociolinguistic interviews and reading speech from word lists were collected from 90 native speakers of Tianjin dialect. The speech data were analyzed to determine how, for each tone sandhi pattern, the sociolinguistic variants are distributed across speaker age, gender, social class and speaker attitudes. The distributions point to the potential social meanings of the variants and provide some initial insights into the likely causes of the changes over time that the sandhi patterns are undergoing.

Chinese is a tonal language and studies of tones occupy an important position in the research on Chinese languages and dialects, but most of these studies are dialectal, acoustic or phonological. Indeed, there are virtually no systematic variationist sociolinguistic studies of tone sandhi in Mandarin in mainland China (although see Gao & Lu, 2003; Feng, 2007/2015; Zhao, 2018; and Zhang, 2019 for some rare examples).

The situation is the same in the study of Tianjin Chinese. Since Tianjin Chinese is known for its intricate tone sandhi patterns, there are ample dialectal, acoustic or phonological studies on Tianjin tone sandhi (Li & Liu, 1985; Chen, 1986 & 2000; Tan, 1986; X. Wang, 2003; Wee, 2004; Lin, 2008; Zhang & Liu, 2011; Li & Chen, 2016; X. Wang & Lin, 2017; among others), but the only sociolinguistic studies on or including Tianjin tone sandhi are Gao & Lu (2003) and Feng (2007/2015), which were made more than 10 years ago. Since then, Tianjin tone sandhi phonology has changed a lot, but no systematic variationist sociolinguistic study has been made to account

for the change. Besides, both studies study the same tone sandhi and attribute its decrease to the influence of Standard Chinese (hereafter SC), but the changes of other Tianjin tone sandhi patterns have not been investigated in any sociolinguistic study. This dissertation investigates linguistic and social motivations of the change and variation in Tianjin tone sandhi, which can be used as a case study to explore how dialects change under the conflicting pressures of standardization on the one hand, and consciousness of local identity triggered by migration on the other hand.

This chapter consists of three sections. Section 1.1 introduces the purpose, hypothesis and the three tone sandhi patterns that are the primary dependent variables of the study. Section 1.2 reviews previous studies of standardization and migration, highlighting the difference between existing studies and the proposals in the current study. Section 1.3 provides an outline of the dissertation.

1.1 The variables, research questions, and hypotheses of the current study

1.1.1 The sandhi variables

Tone sandhi refers to the phonological alternations of tones due to tonal, prosodic, or morpho-syntactic environments in which the tones appear (Chen, 2000; Zhang, 2014). Traditionally four disyllabic tone sandhi patterns have been identified in Tianjin (Li, 1956; Li & Liu, 1985; Shi, 1986; Shi, 1988; Zhang & Liu, 2011; Wang & Lin, 2017), as shown in (1). T represents tone, and therefore T1, T2, T3 and T4 represent tone 1, tone 2, tone 3 and tone 4 respectively. H stands for a high tonal feature and L stands for a low tonal feature¹. The numbers are Chao numbers (Chao, 1968), with 5 indicating the highest tone degree and 1 indicating the lowest tone degree.

¹ Please consult section 3.1 and 3.2 for more detailed introduction to Tianjin lexical tone and disyllabic tone sandhi.

(1) The four traditional disyllabic tone sandhi patterns

	Tone sandhi rules	Examples	Glosses	
a. (53-53)	$T4^{(53)}+T4^{(53)}\rightarrow T1^{(21)}+T4^{(53)}$ (HL+HL $\rightarrow$ LL+HL)	<i>da hui</i>	‘meeting’	
b. (53-21)	$T4^{(53)}+T1^{(21)}\rightarrow T2^{(45)}+T1^{(21)}$ (HL+LL $\rightarrow$ HH+LL)	<i>jiao shi</i>	‘teacher’	
c. (21-21)	$T1^{(21)}+T1^{(21)}\rightarrow T3^{(13)}+T1^{(21)}$ (LL+LL $\rightarrow$ LH+LL)	<i>fei ji</i>	‘airplane’	(old)
	$T1^{(21)}+T1^{(21)}\rightarrow T2^{(45)}+T1^{(21)}$ (LL+LL $\rightarrow$ HH+LL)	<i>fei ji</i>	‘airplane’	(new)
d.	$T3^{(213)}+T3^{(213)}\rightarrow T2^{(45)}+T3^{(213)}$ (LH+LH $\rightarrow$ HH+LH)	<i>mai ma</i>	‘buy horse’	

Three of these tone sandhi patterns (1a-c) do not exist in SC, so their applications set Tianjin dialect apart from SC, and they are prone to be changed under the influence of SC. For this reason, these three tone sandhi patterns are selected as the dependent variables of the current study. Henceforward I refer to the pattern in (1a) as (53-53)², the pattern in (1b) as (53-21) and the pattern in (1c) as (21-21) after their input tonal values in Chao number respectively. The parentheses around the names (53-53), (53-21) and (21-21) follow the usual convention in variationist sociolinguistics for indicating the dependent linguistic variable. The fourth sandhi pattern, shown in (1d) is the only tone sandhi that exists in both Tianjin and SC.

² In previous versions of this work, e.g. Wang (2017), Wang & Wagner (2019), I referred to the three dependent variables using acronyms of their English tonal configurations. (53-53), a falling-falling tone sequence, was (FF), (53-21), a falling-low tone sequence, was (FL), and (21-21), a low-low tone sequence, was (LL).

The application versus non-application of (53-53), as shown in (1a), is the first variable. When there is a pair of falling tones, the first tone variably becomes a low tone. The (53-53) tone sandhi is reportedly disappearing from Tianjin, with the young and those of higher socioeconomic status leading the disappearance (Gao & Lu, 2003).

The second variable is the application versus non-application of (53-21), as shown in (1b). The (53-21) tone sandhi was the least frequent among the four Tianjin tone sandhi changes observed in the 1980s (Shi, 1988), but in contrast to the first variable, it is reportedly applied almost without exception among young Tianjin speakers today (Shi & P. Wang, 2004).

The third variable (21-21) is reported to have two application output variants, shown in (1c). Non-application of (21-21) is highly infrequent, so the focus in this dissertation will be on its two application patterns. The LL.LL→HH.LL rule traditionally was LL.LL→LH.LL, the old variant in (1c) (Li & Liu, 1985). But Lu (1997) indicated the rule had a tendency to be LL.LL→HH.LL, the new variant in (1c), among young female speakers. This new variant was subsequently also attested in (Zhang & Liu, 2011), (Li & Chen, 2016) and (X. Wang & Lin, 2017), but these three are acoustic studies based on data mainly from young speakers. My pilot experiments on old speakers show that the output of old speakers is LH.LL, the old variant, indicating the change in (21-21).

The social motivations for the rise of traditional Tianjin (53-21) and the decline of traditional Tianjin (53-53) are not currently well-understood. It is also unclear why the innovative variant of (21-21) is apparently replacing the old variant. By studying the three variables among Tianjin speech community members from different age groups, occupations, social classes as well as their language attitudes, it will be possible to figure out the linguistic and social motivations of the observed generational changes.

1.1.2 Research questions

The dissertation has two main research questions:

1. Are the changes ongoing in Tianjin tone sandhi above consciousness or below consciousness? Answering this question will be helpful in revealing the social meanings of different variants.
2. Why are local features of the dialect changing in different directions over time and with different social patterns? The answers to the first question will be used to answer this second question.

1.1.3 Hypotheses

With regard to the first research question, the dissertation tests the hypothesis that the second variable (53-21, or HL.LL→HH.LL) is undergoing a change from below, while the other two variables (53-53, or HL.HL→LL.HL and 21-21, or LL.LL→HH.LL) undergoing changes from above (Labov, 1966). A change from below or above refers to whether a change is below or above public awareness (Labov, 1966).

Intra-speaker stylistic shifts are a useful diagnostic for determining whether a change is above consciousness or below consciousness. The study is going to compare and contrast speakers' use of the three variables in word list reading style and interview style. Comparing word list style with interview style will capture differences in the speaker's attention to his or her own speech. The speaker will be paying much less attention to his own speech in interview style than in word list reading style. If a variable is involved in a change from above, the more prestigious variant should appear more frequently in the more careful word list style. If there are no stylistic shifts or the prestigious variant does not appear more frequently in the more careful style, it suggests it is a change from below (Trudgill, 1972).

With regard to the second question, all the three sandhi variables are part of the local dialect, so the different changes in them cannot simply be attributed to standardization i.e. convergence with SC. The dissertation addresses the hypothesis that the changes result from the complicated interaction between standardization, language ideology and the desire to keep local identity which is triggered by migration.

Standardization is an on-going process of suppressing variation and a drive towards uniformity, in speech and in some other forms of social practice (Milroy & Milroy, 1997). It is countered by a process called ‘vernacular maintenance’ (Milroy & Milroy, 1997, p. 53), which is a drive to maintain vernacular speech. The language ideology in favor of the standard variety works as the force leading to the disappearance of salient feature of the dialect, while the vernacular language ideology in favor of the dialect and the emphasis of the local identity drives to keep the local feature of the dialect.

The (53-53) decrease seems to be the result of standardization, and the (53-21) increase is probably associated with community resistance to migrants. The local Tianjin variant of (53-21) has reportedly undergone a simple drop, caused by the influence from SC (Gao & Lu, 2003). Unlike (53-53), the Tianjin variant of (53-21) seems to show a curvilinear distribution with age (X. Wang, 2016). Previous studies (Shi, 1988; Shi & Wang, 2004) suggest that it declined over time like (53-53), probably also because of the influence from SC, but it has since risen after the decline. The rise of (53-21) may be due to its status as a ‘marker’ (Labov, 1972) related to “new” Tianjin identity; and the decline of (53-53) may be due to its status as a ‘stereotype’ (Labov, 1972) of ‘old-fashioned’ Tianjin identity and speech.

In line with previous studies of language change reversal (e.g. Labov, 1972; Dubois & Horvath, 2000), this dissertation shows the effects of local identity in keeping traditional local

features. It is the unmarked local features that are kept by local identity because of language ideology, which explains why (53-21) is selected as a marker to keep local identity while (53-53) is not.

With adequate data from three generations, it is possible to find out not only how each language change is happening in the present, but also how each change happened in the past, and therefore investigate the motivation of the changes. This dissertation connects language variation and change with linguistic and social factors. It is hoped that the study of the variation and change in Tianjin tone sandhi will contribute to the still very limited field of variationist Chinese sociolinguistics, as well as to a more general understanding of the mechanisms of language change, especially the change of dialect under the competing influences of standard language and ideological resistance to in-migration. In the next section, I review previous studies in standardization and vernacular maintenance, and I highlight the effects of language ideology in the conflict.

1.2 Standardization, vernacular maintenance and migration

According to J. Milroy (2001: 531) “the process of standardization works by promoting invariance or uniformity in language structure.” The Chinese government adopted the policy of promoting Putonghua (literally ‘common speech’, i.e. SC) in 1956, which promotes uniformity in Chinese languages and leads Chinese dialect change toward Putonghua, prompting the process of language standardization in China.

Chinese dialects vary greatly. According to Yuan (2000), there are seven dialect groups: Guan (Mandarin dialects, including Beijing, Tianjin and Nanjing variants), Wu (including Shanghai and Suzhou variants), Yue (including Cantonese), Min (including the Hokkien and Fuzhou variants), Kakka (kejia), Xiang (Hunanese) and Gan (Jiangxinese). All these dialects share

the same writing system, but the sounds in these dialects are so different that speakers from different dialect groups are not mutually intelligible³.

The enormous diversity among Chinese dialects is perhaps one reason why the Chinese government adopted the policy of promoting Putonghua in 1956. The government made the argument that, as the common national speech, Putonghua would be helpful in overcoming the communication barriers between people from different dialect groups.

Since the adoption of the policy, Putonghua has become the only official language in all schools, colleges and universities in China. The influence of SC on Chinese dialects is significant (Zhang, 1995; Gao & Lu, 2003; Gu & Liu, 2003; Jiang, 2008; Feng, 2007/2015; among others). However, although most standardization studies in China emphasize the influences of Putonghua on dialects, they ignore “vernacular maintenance” (J. Milroy & L. Milroy, 1997:53), the force that can maintain dialect features (except for rare example like Feng 2007/2015).

Standardization is studied not only in China, but in other countries too (J. Milroy & L. Milroy, 1997; Barrett, 2008; J. Milroy, 2001; Auer, 2005; among others). These studies, in contrast, do not ignore “vernacular maintenance” (J. Milroy & L. Milroy, 1997:53). For example, Barrett (2008) examines the degree of Spanish influence in the speech of three generations of Sipakapense-Spanish bilinguals in Guatemala by quantitative analyses of syntactic variation and code-switching patterns, and finds younger speakers show patterns that suggest a resistance to Spanish influence in Mayan in the process of language standardization. “In particular, the use of SVO word order, which has been an iconic index of Maya assimilation and shift to Spanish, is entirely absent in the youngest generation even though it occurs quite frequently in the speech of

³ Because they are not mutually intelligible, scholars outside Mainland China tend to call them different languages instead of different dialects. Whether they should be called dialect or language is not what this dissertation is interested in, so I leave this question open.

older generations of Sipakapense speakers” (Barrett 2008: 276). Barrett (2008) attributes the hyper-differentiating of the two languages among younger generation to ethnic identity, highlighting the importance of indigenous identity in cases of language change due to contact and language shift.

The influences of indigenous identity on language change can be traced back to Labov’s Martha’s Vineyard study (Labov, 1972), which illustrates the consciousness of local identity, triggered by tourists, could make traditional local features maintain. Similarly, in a study of three generations of ethnic Cajuns in a rural area of Louisiana, Dubois and Horvath (2000) reported on a ‘recycling’ of French-derived features in the Cajun variety of English. The older generation used many French-derived Cajun features in their English, and the middle-aged generation dramatically decreased their use of all of these features. But some of the Cajun features revived among the young generation. The young speakers are influenced by the so-called “Cajun Renaissance”, and they are proud to be Cajuns. Besides, they are able to profit from the important economic benefits of display of Cajun culture associated with tourism. Dubois and Horvath (2000) argue that the Cajun features increased among younger Cajun speakers because the use of Cajun variants has become a sociolinguistic marker of Cajun identity, just as the use of traditional Martha’s Vineyard variants marked island identity.

More recent studies have continued to examine the connection between place identity and language use. Becker (2009) examines the use of non-rhoticity in syllable coda among speakers of English from New York City’s Lower East Side. She states that non-rhoticity, a feature that is still vigorous in that area, is viewed as an indication of authenticity and localness by Lower East Siders. She argues that Lower East Siders use non-rhoticity to claim membership to their local membership.

In these studies, local or dialect features are used to gain recognition of local identity or to signal to other speakers their identification with the place. Labov (1972) relates the retaining of local features to group identity. Labov (1972) noticed a seemingly conflicting phenomenon: people who use local dialects often believe that their dialects are ‘bad’, but they persist in using the ‘bad’ dialects. Labov identifies recognition of group identity as the underlying reason making them use the ‘bad’ dialects. The desire to keep group identity makes covertly prestigious local variants as ideological resistance to outsiders.

Trudgill (1972) illustrates the influences of group identity and covert prestige from the perspective of gender in his studies of English speakers in Norwich. He found that male speakers use more non-standard working class speech variants than female speakers do. He argues that male speakers seem to be more concerned with achieving group solidarity. Kiesling (1988) echoes Trudgill (1972) in the study of word-final *-ing* versus *-in* among college fraternity men in the United States. Based on the observation that fraternity men preferred *-in* to *-ing*, Kiesling concludes that fraternity men used *-in* to indicate working class values such as ‘hardworking’ or ‘casual’, signifying group identity of their own.

The current dissertation studies the interaction between regional group identity and language use in a situation of standardization and in-migration. In Tianjin speech community, although most speakers agreed that Tianjin dialect is not so socioeconomically advantageous as SC, they persisted in using the less advantageous dialect. The study argues that it is covert prestige-related local or group identity that drives locals to retain some dialect features. The desire to keep local identity, or group recognition related to a place, makes the unmarked local feature maintain or even strengthen.

The awakening of the local identity in Tianjin speech community is not related to tourism (Labov, 1972; Dubois & Horvath, 2000) nor ethnic difference (Barret, 2008), nor social class difference (Trudgill, 1972; Kiesling, 1998), but large-scale in-migration. In the 1980s, the Chinese government enacted the “Reform and Open-up Policy”, which introduced market principles into socialism, creating a so-called socialist market economy. After the adoption of this policy, foreign investment and private industry were allowed in China, and Chinese citizens were allowed to leave their hometowns and villages to try their luck in different cities. Tianjin has attracted more and more migrants from other Mainland provinces and cities since the 1980s. Different dialects and cultures came together with these migrants and, as this dissertation will show, the large-scale migration to Tianjin in the past thirty years has made Tianjin locals more conscious of their local identity. This appears to have motivated Tianjin locals to maintain the unmarked Tianjin feature to keep their local identity, forming a force in favor of keeping dialect features.

Speakers’ relationship with places can become especially highlighted in circumstances of displacement and migration (Cotter & Horesh, 2015). Some studies on displacement and migration are focused on situations in which speakers cross political or national borders, facing integration into a new community with a different culture and/or language (Schleef, Meyerhoff, & Clark, 2011). Some studies focus on movement between dialectal boundaries within the same country (Carmichael, 2017). Carmichael (2017) shows the effects of place identity by examining rates of r-lessness in the wake of Hurricane Katrina, which caused large-scale displacement in the region. Variable r-lessness is a salient feature tied to New Orleans English. Participants came from Chalmette, a suburb of New Orleans proper, where r-lessness is more commonly used than in New Orleans proper. Carmichael (2017) argues that the marked quality of r-lessness makes it available

for identity-driven use to express a connection to Chalmette. The current study examines movement between dialectal boundaries within the same country.

The influences of migration on language change not only shown in the strengthening of local identity or place identity, but also in the decreasing of local stereotypical features (Thomas, 1997). Thomas (1997) finds the effects of migration on the accents of Texas by conducting a random-sample survey of Texas and a systematic set of surveys of high schools in Texas. Thomas (1997) reports that young rural Anglos preserve two stereotypical features of the Texas accent, monophthongal /ai/, as in *night*, and the lowered onsets of /e/, as in *day*, while young Anglos from metropolitan centers lack these features. But this difference between young rural and metropolitan Anglos is not shown in the middle-aged and old speakers in these areas, suggesting the difference observed in young speakers is caused by in-migration from other parts of the country, which is concentrated in metropolitan centers. In the current study, the decrease of stereotyped Tianjin feature in the process of standardization might be accelerated by the large-scale in-migration to Tianjin, which caused more situations to speak Putonghua among speakers from different dialects.

In addition to migration, the neighborhood network of Tianjin was destroyed because of the renovation of the urban area and the following relocation of its original residents during the period from the 1980s to the 2000s. Tianjin dialect has historically been used in the six central districts, the urban area of Tianjin. But most original residents were relocated to different places, and therefore their neighborhood network was destroyed. According to J. Milroy and Lesley Milroy (1997), vernacular maintenance is an ideology that is brewed within dense social networks. Here this dissertation suggests that local features can still be kept even if a neighborhood network is destroyed, giving prominence to the effects of local identity in dialect maintenance.

Whether a dialect is lost or maintained depends on the conflict between the force in favor of standard language and the force in favor of dialect. The current dissertation illustrates that in spite of the strong influences from SC, and the destruction of neighborhood vernacular networks caused by renovation of the urban area and the relocation of its original residents, language ideology in favor of local dialect still can maintain or even strengthen the unmarked local feature, but it cannot help to preserve local features that have become overtly stigmatized. The change and variation in Tianjin tone sandhi relates in a complex way to migration, local identity, language ideology and standardization.

1.3 Dissertation outline

In Chapter 2, I introduce the city of Tianjin, including its history, current development and migration, sketching out the social-economic setting in which Tianjin dialect is used. In Chapter 3, I describe the speech community of Tianjin and its distinctive regional dialect. I will introduce the origin and formation of Tianjin dialect, the salient features of Tianjin dialect, Tianjin tone sandhi patterns, the language environments of Tianjin and previous sociolinguistic studies on Tianjin dialect. In Chapter 4, I outline the methods used in obtaining the data, including speaker sample, interview method and coding. In Chapter 5, I present the word list results, looking firstly at the overall main effects and then providing a more detailed presentation of the linguistic and social factors affecting the application of the three variables. In Chapter 6, I report the results of interview style, looking firstly at the overall results and then I provide a more detailed presentation of the linguistic and social factors influencing the application of the three variables in interview. In Chapter 7, I discuss intra-speaker and inter-speaker variation and their social meanings, with the goal of answering the two main research questions raised in this chapter regarding the linguistic

and social motivations for the change in Tianjin tone sandhi patterns. In Chapter 8, I conclude the dissertation and suggest future research.

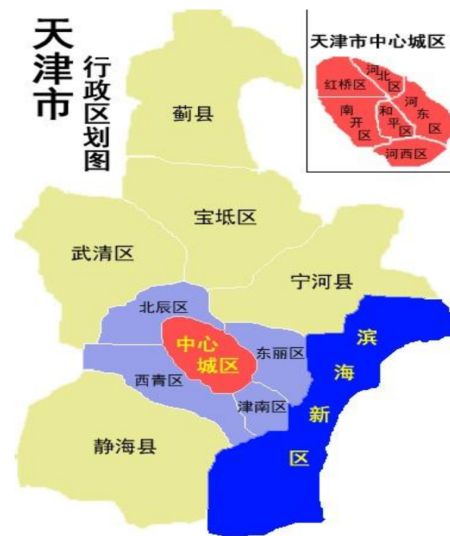
Chapter 2 The socio-economic environment of Tianjin dialect

Tianjin is one of the four province-level municipalities under the direct control of Chinese Central Government, parallel with Beijing, Shanghai, and Chongqing. It is the most important manufacturing center and largest coastal city in Northern China. Situated in the eastern part of the North China Plain, Tianjin borders Hebei province to the north, south and west, Beijing to the northwest, and the Bohai Gulf to the east. The city of Tianjin covers 11,919 square km and consists of 16 administrative districts with a total population of 15.56 million (Tianjin Municipal Bureau of Statistics, 2018).

The Tianjin urban area consists of six districts: Heping District, Nankai District, Hexi District, Hedong District, Hebei District and Hongqiao District. The other twelve districts are called ‘new districts’. They were formerly in the rural area around Tianjin. Tianjin dialect is mainly used in urban area of Tianjin (Li & Han, 1991). Figure 2.1 is an administrative division map of Tianjin.

As shown in Figure 2.1 below, the red area in the big map is the urban area, or the six central districts of Tianjin, which today takes up a very small proportion of Tianjin city’s total area. The purple-blue area, the yellow green area and the sapphire blue area are the twelve formerly rural districts.

Figure 2.1 Administrative division map of Tianjin (Retrieved from <https://baike.sogou.com/historylemma?IIId=6808&cId=173664100> on July 15, 2019)



Since the founding of the People's Republic of China in 1949, Tianjin has undergone big socio-economic changes, which provides the social environment for a series of changes in various aspects of Tianjin dialect. This chapter will give an overview of Tianjin city, with a purpose of providing socio-economic settings for variation and change in Tianjin dialect. After providing a brief introduction to Tianjin history in §2.1, I will introduce the in-migration to Tianjin from other mainland provinces and cities in §2.2, and the renovation of the urban area in §2.3. The in-migration and renovation cause the change in language environments in Tianjin, which in turn has great impact on the change and variation of Tianjin tone sandhi.

2.1 Brief history of Tianjin city

The brief history of Tianjin city provided in this section is drawn largely from Kwan (2001) and Tan (2015).

The development of Tianjin is closely related to water transportation. The land where Tianjin lies today was historically the silting of various rivers entering the sea at the Bohai Bay. It was the construction of the Grand Canal during the Sui Dynasty (518-618) that made Tianjin

develop into a trading center. The Grand Canal makes it possible to transport grain from the south to the north by water since Chinese main rivers are from west to east. Tianjin was originally only a wharf where grain barges going up one segment of the canal were transferred to other segments to continue their northward journey. As the trading center became prosperous, a fortified town was built during the Song Dynasty (960-1279) to protect the region and trade from raiders. It was in that period that Tianjin obtained its first name *Zhigu*, which means ‘straight port’.

In 1404, Tianjin acquired its current name, Tianjin, which means ‘heavenly port’. The name is linked with Yongle Emperor of the Ming Dynasty, who crossed the Haihe river, the most important river in Tianjin, before he overthrew the reign of his nephew. After Yongle Emperor moved the capital to Beijing, a fort was established at Tianjin, known as *Tianjinwei*, ‘Fort Tianjin’. Large number of soldiers and their families came with Emperor Yongle from Anhui province and settled down at Tianjin. The soldiers and their families took the center and prosperous parts of Tianjin, and the original residents were driven to neighboring poor areas. This is the first mass immigration in Tianjin history, which had great influences on Tianjin dialect. Currently the Tianjin dialect which is used in the urban area of Tianjin city is still very similar to Anhui dialect, but it is different from all geographically neighboring dialects (Li & Han 1991).

During the Second Opium War, the Treaties of Beijing were signed in 1860, which opened Tianjin to foreign trade. From 1860 to 1900, Britain and France were joined by the empires of Japan, Germany, Russia, Austria-Hungary, Italy and Belgium in establishing self-contained concessions in Tianjin, each with its own prisons, schools, barracks and hospitals. These concessions in Tianjin took up 3867 acres, 7 times bigger than the Old Tianjin Town, and contributing to the overall expansion of the city. Tianjin developed rapidly and was established as

a municipality of China in 1927. During the period of Japanese occupation (1937-1945), Tianjin was ruled by a puppet state based in Beijing.

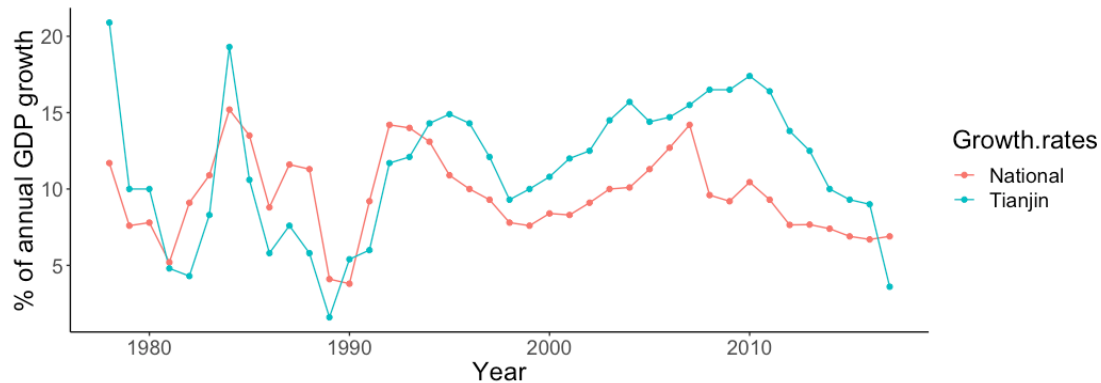
Since the foundation of the People's Republic of China in 1949, Tianjin has been a municipality of China, except between 1958 and 1967, when it was the capital of Hebei province. Tianjin experienced rapid economic development after the reform and open-up policies were adopted by Chinese government, which has important influences on Tianjin dialect. Tianjin Section 2.2 will be a brief account of the economic development of Tianjin and the arrival of migrants.

2.2 Tianjin industry, economy and migration

As explained in the previous section, Tianjin has long been an industrial center in China. As one of the places where Chinese modern industry originated, Tianjin was famous for its salt industry, soda industry, carpet industry and textile industry in the 19th century. Later, the chemical industry, textile industry, steel industry and machinery manufacturing became the main industries in Tianjin. Tianjin is still a stronghold of the chemical industry, but machinery and textile manufacturing have been replaced with a much wider range of endeavors, including medical biotechnology, new energy, new materials, electronic industry, petrochemical technology, astronautics and aeronautics (D. Wang, 2011).

Tianjin has witnessed a high-speed development since the adoption of the Open-up and Reform policy in 1978. Its GDP increased at a double-digit growth rate annually from 1992 to 2015. Figure 2.2 compares the annual GDP growth rates of Tianjin (blue line) to the annual national growth rates of China (red line) from 1978 to 2017. The annual GDP growth rates of Tianjin are shown to be higher than national growth rates from 1995 to 2016, although its economy growth slows down in 2017 (Tianjin statistical Bureau 2018).

Figure 2.2 Annual GDP growth rates of Tianjin and the whole country from 1978 to 2017 (Tianjin statistical bureau, 2018)

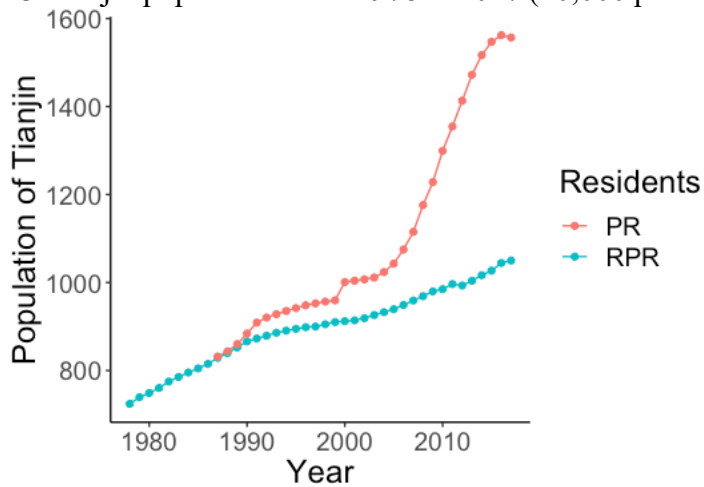


The Tianjin population, especially migrants, has increased rapidly as its economy has developed. By the end of 2017, the permanent residents of Tianjin were about 15.62 million people, of whom approximately one third (5 million) were migrants (Tianjin Statistics Bureau, 2018) from other mainland cities or provinces who work in Tianjin.

For official purposes in China, people are categorized as migrants not by how long they stay in a city, but by whether they have a hukou for a city. A hukou is a record in the system of household registration required by law in Mainland China. A household registration record officially identifies a person as a resident of an area and includes identifying information such as name, parents, spouse, and date of birth. People who do not have a permanent hukou for Tianjin are officially considered to be migrants, although they may have stayed in Tianjin for decades.

Since one hypothesis of the dissertation is that migration is an influential factor of the changes in Tianjin dialect, the size of the migrant and non-migrant Tianjin population from 1978-2017 is provided in Figure 2.3. The data are from the Tianjin statistical yearbook 2018 (Tianjin Statistics Bureau, 2018). ‘PR’ stands for permanent residents, and RPR stands for registered permanent residents. PR includes all people who live in Tianjin, while RPR refers only to those having Tianjin hukou. The number of migrants is obtained by subtracting RPR from PR.

Figure 2.3 Tianjin population from 1978 to 2017 (10,000 persons) (Tianjin statistical bureau 2018)



As shown in Figure 2.3, only registered permanent residents are recorded for the years from 1978-1986. But both PR and RPR information is provided from 1987 onwards. The statistical yearbook does not explain why PRs do not appear in the records until 1987, so I can only speculate. The reason perhaps is that permanent residents are equal in number to registered permanent residents from 1978-1986, or at least the difference between permanent residents and registered permanent residents in these years is negligible. This population information echoes Chinese residence policies. There were almost no migrants in Mainland Chinese cities between the founding of the People's Republic of China and the adoption of the "Reform and Opening-up" economic policy in 1978 because people were supposed to stay where their hukou was in those years. Migrants therefore had not started to arrive at all in Tianjin until the early 1980s.

Figure 2.3 also indicates there is an overlap between PRs and RPRs in the years from 1987 to 1990, indicating that the difference between PR and RPR is very small. In other words, the number of migrants is very limited in those years. Shortly after the adoption of the "Reform and Open-up" policy, although Chinese were allowed to migrate freely within mainland China, people were too cautious to try their luck in other provinces or cities in case the policy changed again.

The situation did not change until the important speech made by the former President of China Deng Xiaoping in 1992, which confirmed the “Reform and Open-up” policy would not be reversed. Chinese people became bold, and more and more people left their hometowns for better opportunities. This is also reflected in Figure 2.3. PR and RPR can be separated clearly since the early 1990s, suggesting large numbers of migrants started to arrive after that initial wait-and-see hesitation.

Figure 2.3 also indicates that there was a sharp increase in the number of migrants from 2005-2016, which may be related to government policy too. In 2005, the Sixth Plenum of the 16th Central Committee of the Communist Party of China passed a decision, according to which the development of Binhai New Area of Tianjin would be the third pole to promote national economic growth. The other two poles to promote national economic growth were the development of Shenzhen New Area in the 1980s and Pudong New Area of Shanghai in the 1990s. The successful examples of Shenzhen and Pudong New Areas led Chinese people in general to have very high expectations of the development of Binhai New Area of Tianjin. After the passing of the decision, large numbers of migrants arrived in Tianjin to try their luck, which is reflected in the abrupt jump of PR since 2005 in Figure 2.3.

In addition to government policies, the Tianjin population, especially the number of migrants, fluctuates with its economic development. Figure 2.2 suggests that Tianjin GDP growth rates are higher than national growth rates from 1992 to 2016, but its annual GDP growth rate is slower than the national growth rate in 2017, for the first time in the past 16 years. Figure 2.3 indicates that migrants increase steadily and continuously in Tianjin from 1992 to 2016, but decrease in 2017 as Tianjin economy declines. As shown in Figure 2.3, there is a drop in PR in

2017, while RPR keeps a steady rising trend, suggesting the minor decrease in the year is caused by the decrease of migrants.

2.3 The renovation of the urban area

The renovation of the urban area and the following relocation of its original residents destroy the neighborhood networks of using Tianjin dialect, exerting major influences on the change of Tianjin dialect. So I will introduce the renovation of the urban area here.

Tianjin has sprawled like all other cities in China with the socialist construction and economic development since the founding of People's Republic of China in 1949. Since the city was not developed with a clear layout, the urban functional partition is not clear. As a result, industrial areas, business areas and financial areas are mixed together with residential areas. Besides, with half-a-century-old living quarters and slums, Tianjin central districts were very old and shabby at the end of the 1970s and beginning of the 1980s. To make things worse, Tianjin suffered a heavy blow in the 1976 Tangshan earthquake. More than half of the buildings in Tianjin were damaged during the earthquake and Tianjin people were forced to live in anti-quake thatched houses.

In order to change the disorderly layout of the city and improve people's living conditions, the Tianjin government started a reconstruction of the urban area in the beginning of the 1980s. The renovation of the urban area consisted of two stages. First, in the beginning of 1980s, the renovation began with the removal of anti-quake thatched houses. Then all run-down areas started to be rebuilt since 1994. By the end of 2000, the large densely-populated and most difficult areas, including Nanshi in Nankai and Heping Districts, Xiguangkai in Nankai Districts and Zhongshanmen in Hedong District, had been rebuilt and the original residents relocated.

Figure 2.4 Layout of city planning of Tianjin (2005-2020) (Source: City construction committee of Tianjin 2005; Retrieved from <http://www.cityup.org/case/general/20090114/44146.shtml>)



As shown in Figure 2.4, the bigger red circle on the left is the urban area, incorporating the six central districts of Tianjin. The smaller red circle at the right is Binhai New Area. The six central districts and Binhai New Area have formed the two cores of Tianjin since 2005. Yellow circles represent extended new districts that were originally suburbs and rural areas. The residents who originally lived in the urban area (big red circle) were encouraged to be relocated in new districts or Binhai New Area.

The relocation policy successfully reduced the density of the central districts. Most original residents of the central districts did not move back after the reconstruction of the urban area. Instead they settled down in the new districts in their more spacious apartments. Only less than 10% of the original residents in Nanmenwai (part of the reconstruction areas in Nanshi, Nankai district) have moved back, while the majority have chosen to live in other residence, either in central districts or new districts (Fu, 1991:119).

The population changes in the central districts and newly developed districts are given in Table 2.1, indicating the movement from central districts to the extended new districts. PR represents permanent residents, including permanent residents with or without Hukou. Newly developed areas include Binhai New Area and other new districts. Since the tendency is from the urban area to newly developed areas, Binhai New Area and other new districts, it is not necessary to distinguish between Binhai New Area and other new districts.

Table 2.1 Population of different districts in Tianjin, selected years from 1980 to 2017(10,000persons) (Source: Tianjin statistical bureau, 2018)

Year	Six Central Districts		Newly Developed Areas	
	PR	% in total population	PR	% in total population
1980	302.61	40.41%	446.29	59.59%
1985	333.98	41.5%	470.82	58.5%
1990	360.2	41%	518.34	59%
2000	390.28	39%	610.6	61%
2003	380.52	37.63%	630.78	62.37%
2010	434.58	33.45%	864.71	66.55%
2014	504.69	33.27%	1012.12	66.73%
2015	488.12	31.55%	1058.83	68.45%
2016	492.53	31.53%	1069.59	68.47%
2017	490.87	31.53%	1066	68.47%

As shown in above Table 2.1, the proportion of the residents who live in the six central districts has been decreasing since 1985, when the reconstruction and relocation of the urban area

started. At the same time, the proportion of residents living in newly developed areas has been increasing since 1985. The resident proportions of the old urban districts and newly developed district suggest that at least some of the original residents of the old urban area have not moved back.

Although the reconstruction and relocation of the urban area has improved the living conditions of the original residents and made the city appear more modernized, its impact on Tianjin dialect and local culture is huge. The most densely populated residential areas like Nanshi in Heping and Nankai districts, Old Town of Tianjin in Nankai district, Xiguangkai in Nankai district and Zhongshanmen in Hedong districts are also the places where the features of Tianjin dialect are best maintained. Tianjin natives even disagree about whether the original residents of Old Town of Tianjin in Nankai district speak the purest Tianjin dialect or the original residents of Zhongshanmen in Hedong district speak the purest Tianjin dialect. In my investigation, some participants suggested that I recruit the original residents of Old Town of Tianjin, Nanshi or Zhongshanmen because what they speak is the ‘purest’ Tianjin dialect. The dissertation is not interested in who speaks the purest Tianjin dialect, but the disagreement reflects the importance of these reconstructed and relocated residences in the use of Tianjin dialect.

Now since the reconstruction of the urban area, the original residents are separated and therefore the residential networks for speaking Tianjin dialect have been destroyed. In other words, the original neighborhood networks for speaking Tianjin dialect have disappeared as the original residents were relocated in different places. Some of them, with better economic situations, moved into some communities that are close to the city center, with better services and better environments. Those with less money have to live in extended districts which cannot provide equally convenient facilities. As we know, Tianjin dialect is only used in the six central districts.

When the original residents of the central districts are relocated in newly developed districts, they live in a community in which most people do not speak Tianjin dialect.

The influences of the reconstruction of the urban area and the relocation of its original residents are also mentioned by some participants. For example, some participants said that they did not speak Tianjin dialect as frequently as they used to since they relocated in their new residences. They said they were separated from their old neighbors and friends, and people in their new communities did not speak Tianjin dialect, so gradually they spoke Tianjin dialect less and less. They spoke Putonghua because Putonghua is the language used by everyone.

In summary, the socio-economic background of the Tianjin speech community has great influences on change and variation in the Tianjin dialect, including endangerment of the dialect. The in-migration to Tianjin in the past thirty years is proposed to be a main force that has triggered the consciousness of Tianjin local identity, which in turn blocks the process of standardization on unmarked local features. Yet the renovation of the urban area of Tianjin affects the social environment of Tianjin dialect, and as a result, accelerates the process of standardization. They are conflicting forces affecting the change and variation of Tianjin tone sandhi.

After presenting the socio-economic background of Tianjin dialect, I will introduce Tianjin dialect in the next chapter. The introduction mainly includes the features of Tianjin dialect, the language environment of Tianjin dialect and Tianjin natives' attitudes to SC and Tianjin dialect.

Chapter 3 Tianjin dialect

Tianjin dialect belongs to the Northern Mandarin group of dialects, on which SC is based (Li & Han 1991). Tianjin dialect is not grammatically different from SC and they are mutually comprehensible. Tianjin dialect differs from SC mainly in consonants and tones (Gu & Liu, 2003; Gao & Lu, 2003). Since the linguistic variables of the present dissertation are Tianjin tone sandhi rules, I will introduce Tianjin lexical tones (§3.1) and tone sandhi processes (§3.2) first. In what follows, I will introduce the language environment of Tianjin dialect in §3.3, promotion of Putonghua (SC) in § 3.4, attitudes to SC and Tianjin dialect in §3.5 and existing sociolinguistic studies on Tianjin in §3.6. At the end of §3.6, I will discuss the limitations of the existing sociolinguistic studies on Tianjin and show how the current study can better account for the variation and change in Tianjin tone sandhi.

3.1 Tianjin lexical tones

Although there are four lexical tones in Tianjin dialect, and the four tones are cognates to the four lexical tones in SC, the pitch values are different between Tianjin dialect and SC, as shown in Table 3.1. The numbers are Chao numbers (Chao, 1968), with 5 indicating the highest tone degree and 1 indicating the lowest tone degree. The letters in parentheses are phonological representations.

Table 3.1 Lexical tones in Tianjin dialect and SC

	Tone1	Tone2	Tone3	Tone4
Tianjin	11/21 (LL)	45 (HH)	213/13 (LH)	53 (HL)
SC	55 (HH)	35 (HH)	214 (LH)	51 (HL)

Among the four lexical tones (henceforth T1, T2 etc.), T1 has been taken as the tone that distinguishes Tianjin dialect from Mandarin and geographically neighboring dialects (Li & Han,

1991). This criterion has been applied to studies of the presence or absence of Tianjin dialect within the city's districts (Gao & Lu, 2003:346; Shi & Wang, 2004:176). Han (1983) drew out the first dialect map of Tianjin using T1 as the dividing line.

The view that T1 is a distinctive feature of Tianjin dialect is consistent with the investigation I conducted. In every interview, I asked: “What do you think is the most significant feature of Tianjin dialect?” All participants mentioned that the low T1 was the most significant feature of Tianjin dialect, although they usually didn't explicitly say tones. Most participants said that Tianjin dialect was “low” compared to SC and Beijing dialect. The examples they gave to support that Tianjin dialect was “low” were T1 words. “Low” means low in pitch, judging from their answers. Although most of them did not say tones explicitly, they are clear that low T1 is the most distinctive feature of Tianjin dialect.

3.2 *Tianjin disyllabic tone sandhi*

Tianjin disyllabic tone sandhi patterns are much more complex than those in SC. Traditionally there were 4 disyllabic tone sandhi patterns in Tianjin dialect (Li, 1956; Li & Liu, 1985; Shi, 1986; Shi, 1988; Zhang & Liu, 2011), while SC only has one disyllabic tone sandhi pattern, as shown in Table 3.2.

Table 3.2 The tone sandhi rules in Tianjin and SC

	Tone sandhi rules in Tianjin	Tone sandhi rules in SC	Examples	Glosses	Variable
a.T4+T4	T4+T4→T1+T4		<i>jiao shou</i>	‘professor’	(53-53)
b.T4+T1	T4+T1→T2+T1		<i>jiao shi</i>	‘teacher’	(53-21)
c.T1+T1	T1+T1→T3+T1 →T2+T1		<i>fei ji</i>	‘airplane’	(21-21)
d.T3+T3	T3+T3→T2+T3	T3+T3→T2+T3	<i>mai ma</i>	‘buy horse’	

Pattern (3d), the only sandhi rule that exists in both SC and Tianjin, is robust and occurs frequently (Shi & Wang, 2004; Zhang & Liu, 2011). It will not be discussed further in the present dissertation. As noted in Chapter 1, (3b), henceforward (53-21) has reportedly been increasing in Tianjin, while (3a), henceforward (53-53) is in decline (Shi & Wang, 2004; Zhang & Liu, 2011). However, (3c) has reportedly developed a new variant, $LL+LL \rightarrow HH+LL$ ($T1+T1 \rightarrow T2+T1$) (Lu, 1997; Zhang & Liu, 2011; X. Wang & Lin, 2017). As mentioned in the introduction, the rules in (3a), (3b) and (3c) are the three variables of the dissertation.

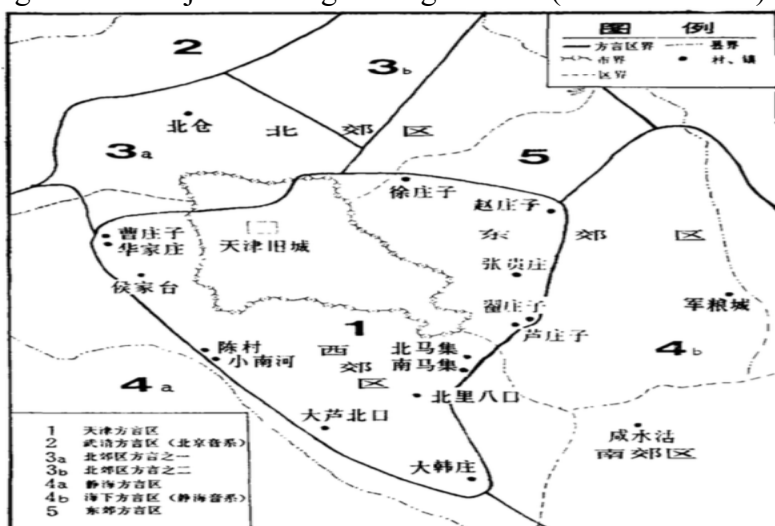
Among the aforementioned studies of Tianjin disyllabic tone sandhi, Shi (1988) was the first to note that (53-21) was the least frequently applied rule among the four sandhi patterns. Previous papers had all simply described (53-21) as one of the four basic tone sandhi rules, which implied (53-21) had no notably different application frequency than the other three tone sandhi patterns. This may be because the authors did not attend to distributional frequencies, but it may also be because there really were no differences between the application rates of (53-21) and the rates of the other three tone sandhi rules at the time of publication. I therefore assume that the application rates of (53-21) were similar to the rates of the other three rules before 1988.

3.3 Language environment of Tianjin dialect

According to Li and Han (1991), Tianjin dialect is not used in all areas of Tianjin, but mainly in the six central districts introduced in Chapter 2, known as Heping district, Nankai district, Hexi district, Hebei district, Hedong district and Hongqiao district, and also in parts of Xiqing district and Dongli district. The presence of low T1 in the dialect is the diagnostic used by Li and Han (1991) to decide whether a speaker speaks Tianjin dialect. They did not report explicitly in the paper how they obtained this conclusion. If speakers of an area use low T1, they classify that area as an area using Tianjin dialect. Although Li and Han (1991) is not the latest dialect study of

Tianjin, their standard has been used in all studies on Tianjin dialect (Gao & Lu, 2003; Shi & P. Wang, 2004; Zhang & Liu, 2011; among others). The Arabic numbers in Figure 3.1 indicate dialect areas.

Figure 3.1 Tianjin and neighboring dialects (Li & Han 1991)



The triangular shaped area with Arabic number 1 in the middle is the area where Tianjin dialect is spoken, comprised mainly of the 6 central districts. Tianjin dialect is surrounded by Jinghai dialect, Wuqing dialect and Jindong dialect, as shown in Figure 3.1. The area marked 2 is where Wuqing dialect is used, which is very similar to Beijing dialect. The areas marked 3a and 3b are where Beijiao dialect is used. The areas marked 4a and 4b are where Jinghai dialect is used. 5 indicates the area where Dongjiao dialect is spoken. The areas marked 2, 3, 4 and 5 were originally suburbs, and have now become the new districts of Tianjin. All those dialects are still spoken in those areas.

The only salient phonological feature that can distinguish Tianjin dialect from all the geographically neighboring dialects is low level T1 (Li & Han, 1991). Although lack of retroflex is a salient feature of Tianjin consonant, the feature exists in Jinghai dialect too (Li & Han, 1991:71).

3.4 Promotion of Putonghua

SC or Putonghua is formally defined as the standard form of Modern Chinese, with the Beijing phonological system as its norm of pronunciation, and northern dialects as its base dialect, and modern classics of literature as its grammatical norms (P. Chen, 1999).

The promotion of Putonghua policy was adopted by the Chinese government in 1956. Since then the government strongly encouraged that everyone must learn Putonghua. The State Council issued an official directive ordering the teaching of Putonghua in schools of all levels since the fall of 1956. The directive also contained extensive instructions on how Putonghua was to be promoted in various areas of national life: broadcasting, journalism, industry and commerce.

In order to help people speaking various dialects to learn Putonghua, fieldwork surveys highlighting the similarities and differences between dialects and Putonghua were carried out in 1950s. As one of those efforts, Li (1956) concentrates on the contrastive analysis of phonetic systems between Tianjin dialect and Putonghua, mostly for the purpose of popularizing Putonghua.

Promotion of Putonghua was resumed after the Cultural Revolution (1966-1976). The policy was even added to the revised Constitution of the People's Republic of China (PRC) in 1982. In 1986, the State Education Commission and the State Language Commission held a joint National Conference on the promotion of Putonghua, aiming to promote Putonghua at the national level (Tsung, 2014:39). The following proposals were passed in the conference:

- a. Putonghua is to become the language of instruction in all schools;
- b. Putonghua is to become the working language in government at all levels;
- c. Putonghua is to be the language used in broadcasting and journalism;
- d. Putonghua is to become the lingua franca among speakers of various local dialects.

Putonghua was given the status of a national language in the 2001 PRC National Common Language and Script Law (*Zhonghua renmin gongheguo guojia tongyong yuyan wenzifa*).

Promotion of Putonghua has huge influences on Tianjin speakers and Tianjin dialect. More than one young participant told me they had spoken Tianjin dialect before they went to primary school. After they went to primary school, they were asked to speak SC in school. They were corrected in class whenever they answered questions in Tianjin dialect. One extreme example was a girl who told me she was asked to stay in school after class because she could not speak SC. The teacher asked her to stay so that she could teach her SC. She said although the teacher was very patient and friendly, she felt very unhappy because she could not go home at the same time as her classmates. She started to speak SC from then onwards.

The pressure from the job market to speak Putonghua is even heavier than from education. Anyone who wants to secure a teaching position in China, no matter in primary schools or universities, has to take the National Oral Proficiency Test of Putonghua and at least has to pass level II. All government employees have to take the test and at least have to pass level III.

Under the influences of the dominant and effective measures of promoting Putonghua, more and more Tianjin locals are fluent in Putonghua. It is hard to find a mono dialectal Tianjin dialect speaker among young and middle-aged Tianjin locals.

3.5 Attitudes to SC and Tianjin

Most Tianjin speakers, especially young and middle-aged speakers, can speak both SC and Tianjin dialect. Tianjin is effectively a diglossic community (Ferguson, 1959). In my personal experience of living in Tianjin, Tianjin people generally think SC is more prestigious than Tianjin dialect. All the interviewees said they would use SC in school, job interviews and other formal situations, while they would speak Tianjin dialect with their families, relatives, and friends. Some

young participants told me their parents wanted them to speak SC instead of Tianjin dialect because their parents thought SC can make speakers sound better educated and from a higher social class. Some middle-aged and old speakers said they did not want their children or grandchildren to speak Tianjin dialect, so they did not speak Tianjin dialect with their children or grandchildren. Speaking SC is more helpful for their children to get a good position in society. Some participants said since their children had to speak SC in school, it was better to use SC at home too.

But it does not mean Tianjin people's attitude to SC is completely positive. Some participants think that Tianjin people speaking SC is 'pretentious', 'snobbish' and 'affected'. Likewise, Tianjin dialect is not viewed entirely negatively. Some participants said Tianjin dialect makes them feel intimate, cordial and warm. Some participants said they like Tianjin dialect because it is humorous.

In addition, although SC is more prestigious, Tianjin dialect has a function that SC does not have—Tianjin dialect can mark residents of the six central districts from those who lived in the suburbs in the past, and mark Tianjin locals from migrants when the presence of migrants is evident. As the lingua franca of China, SC does not have this function. This function is reflected in the interviews of those participants who have close contact with residents of 'new districts'. Quotations from # 67 are given below. She is an old female speaker. She had worked at a vegetable company before she retired. Her job was to buy vegetables from villagers of the rural areas, so she had close contact with them.

Hi, laoxiang, nin zeidiao xing a, women santian mei zuan guolai nin nei wei.

'Hi, villagers, your dialects made us confused for three days'.

Women jue de benwei de yuyan tebie zhan qiangshi.

'We think Tianjin dialect is an advantageous language, compared to the neighboring

rural dialects’.

Yinwei women jiena jiushi sijiao wuxian dou guolai, gen tamen bi, benwei de yuyan
hen qiangshi.

‘We worked with people from suburbs and rural areas. Compared to their dialects,
Tianjin dialect is a dominant dialect’.

(# 67, an old female participant)

Because of the big urban-rural gap in China in the Pre-reform period, urban residents had some privileges over rural residents, not only in Tianjin but all over mainland China. Since the adoption of the Open-up and Reform policy, the urban-rural gap has become smaller. But the old prejudice does not disappear completely.

With more and more migrants coming in, Tianjin natives felt their presence sharply. Some participant attitudes to Tianjin dialect are changing, which might be related to the migration. For example, below is what # 60 said about Tianjin dialect.

Wo xiaoshi ting, duo nan ting.

‘When I was a little girl, Tianjin dialect was very unpleasant to me’

Xianzai ting, you gexing, ting hao de.

‘When I hear Tianjin dialect now, I feel it is quite good because it has its features’

Yiting jiushi Tianjin ren, you daibiaoxing, tinghao.

‘As soon as you speak Tianjin dialect, others know that you are a Tianjin native.

It’s good to speak Tianjin dialect since it can indicate you’re a Tianjin native’

(#60, a middle-aged female participant)

Tianjin dialect seems to be on the second level of a three-level prestige system, with SD on top of it and neighboring dialects at the bottom. After large-scale in-migration took place, Tianjin dialect has a new function---to mark Tianjin locals as different from migrants.

People from other areas, especially people from Beijing, tend to find Tianjin dialect interesting and amusing. I have heard Beijing people say: “Tianjin Hua, zhen dou”. *Hua* means dialect, and *zhen dou* means amusing. Beijing people tend to laugh in a friendly way when they hear Tianjin dialect.

Tianjin people are generally reluctant to leave their hometown. I interviewed 42 young (college-aged) Tianjin speakers in the summer of 2014. When I asked whether they wanted to leave for Beijing or Shanghai after graduation since there were more opportunities in those cities, almost all the interviewees said no. The reason the participants gave why they preferred to stay in Tianjin was mainly that they were reluctant to be away from their families. Besides, they thought that it was too competitive to live in Beijing, and the cost of living in Beijing was too high.

These attitudes, both to the external Chinese urban landscape and to the internal city demographics, are important for an understanding of the analyses I will present in this dissertation. I provide more detail on the questions I asked in the interviews in Chapter 4.

3.6 Existing studies on Tianjin dialect

Previous studies on Tianjin dialect are rich and fruitful, but most of the studies were conducted by dialectologists, phoneticians and phonologists without discussion of social factors (Li, 1956; Li & Liu, 1985; Li & Han, 1991; Li & Han, 1992; Chen, 2000; J. Wang, 2002; X. Wang, 2003; Zhang & Liu, 2011; X. Wang & Lin, 2017). Sociolinguistic studies of Tianjin were not conducted until the late 1990s. For example, Zhou (1997) studies kinship terms in Tianjin, and proposes the principles of using kinship terms: order of seniority of a family or clan and hearer

comfortable. Later sociolinguistic studies on Tianjin (Gu & Liu, 2003; Gao & Lu, 2003; Feng, 2007/2015) focus on dialect change, which will be introduced in detail. Shi & P. Wang (2004) is an acoustic study on Tianjin, but this work incorporates speaker's age, so it will also be introduced here.

Gu & Liu (2003) studied what they called 'special consonants' in Tianjin. For example, [tʂ], [tʂʰ] and [ʂ] in SC are pronounced as [ts], [tsʰ] and [s] in Tianjin dialect. The [ts], [tsʰ] and [s] replacing [tʂ], [tʂʰ] and [ʂ] are called 'special consonants' of Tianjin dialect (Gu & Liu, 2003). Some of these 'special consonants' are shown in Tables 3.3–3.6 ((Gu & Liu, 2003). These so-called 'special consonants' could mark the difference between Tianjin dialect and SC, but similar phenomena exists in geographically neighboring Hebei dialects (Li & Han, 1991).

Table 3.3 'Special consonants' (Type I)

Pinyin	SC	Tianjin	Gloss
Zhayan	[tʂa]	[tsa]	'blink'
Zhuzhai	[tʂai]	[tsai]	'house'

Table 3.4 'Special consonants' (Type II)

Pinyin	SC	Tianjin	Gloss
rengrang	[rəŋ ran]	[iəŋ ian]	'still'
guangrong	[ruŋ]	[iuŋ]	'glory'

Table 3.5 'Special consonants' (Type III)

Pinyin	SC	Tianjin	Gloss
youyong	[iuŋ]	[ruŋ]	'useful'

In Tianjin dialect, [i] in SC is pronounced as [r] when it is in front of some rhymes. Type III ‘Special Consonants’ as shown in the above Table 3.5 reflects this type of change.

Table 3.6 ‘Special consonants’ (Type IV)

Pinyin	SC	Tianjin	Gloss
Anquan	[an]	[nan]	‘safe’
Ouran	[ou]	[nou]	‘by chance’

When [ai], [an], [aŋ], [au], [ou], [əŋ] and [ɤ] are syllable initial, a [n] is inserted in front of them in Tianjin. This insertion is the Type IV ‘Special Consonants’ as shown in Table 3.6.

Gu and Liu’s analysis of word list data from 100 participants showed that younger, white-collar workers and people with higher education produced fewer ‘special consonants’ than the older, blue-collar workers and people with lower education. Interpreting this age and social class distribution as a generational change in progress, they predicted that the ‘special consonants’ in Tianjin will eventually disappear.

Gao and Lu (2003) study Tianjin tone sandhi changes. Their analysis of 10-minute interviews and word list readings by 104 participants showed that the Tianjin sandhi rule HL.HL→LL.HL (the (53-53) variable in the present study) had become a sociolinguistic marker because it “carries the social information of the speakers” (Gao & Lu, 2003:257). A marker, as defined by Labov (1972), is a linguistic variable that speakers are less aware of than a stereotype, but which shows consistent style effects. The non-application of the Tianjin rule is considered a more prestigious form because it is less local-sounding and more like SC, while the application of the rule is taken as a less prestigious form. Gao and Lu (2003) report that the less-prestigious Tianjin variant appears much more frequently than the more standard, prestigious variant in interview style than in wordlist style. Like Gu and Liu (2003)’s claim for ‘special consonants’,

they predicted that this salient Tianjin feature would disappear completely from Tianjin dialect in the future.

In contrast, Shi and P. Wang (2004) found that (53-21) applies in the speech of younger speakers more frequently than in other age groups. In a study of 204 speakers' word list data, Shi and Wang report that almost all young Tianjin speakers have the HL.LL→HH.LL pattern (53-21), while only some middle-aged and old Tianjin speakers have this pattern. The proportion of speakers having this pattern was higher in the middle-aged group than that in the old group. However, although Shi and P. Wang (2004) incorporated speaker ages in their study, they did not collect information on speaker social classes, genders or styles, nor on speakers' attitudes to SC and Tianjin dialect.

By comparing the linguistic behavior between employees in state-owned and foreign companies, Feng (2007/2015) studies the variation and changes in Tianjin dialect from three aspects: salient lexical items, 'peculiar consonants' realization and FF rule application (53-53 variable in this dissertation). Feng's 'peculiar consonants' are the 'special consonants' in Gu and Liu (2003). The methods Feng (2007/2015) adopted in consonants and FF analysis were sociolinguistic interviews in the variationist tradition, each interview including a casual speech and a passage reading. The interviews were not one-to-one interview, but group interview. Since Feng does not speak Tianjin dialect, the interviews were conducted by her friends. Her group interviews were conducted "either at a participant's home when there was a family reunion, at a restaurant when there was a get-together or birthday party, at the workplace during the lunch breaks or at the recreation center when informants got together for fun" (Feng, 2007: 114).

The method for lexical analysis was salient lexical word list checking. Feng (2007/2015) made a list of salient Tianjin words and asked her informants to check whether they knew the

words. Feng (2007/2015) finds that attitude to Tianjin dialect and frequency of using SC have the greatest power in explaining peculiar consonant realizations. Feng (2007/2015) also reports that people who work in foreign enterprises are more likely to use SC than those work in state-owned enterprises (interestingly, in contrast with the findings of Zhang 2005 for Beijing, in her study of these two types of employee). Regardless of the enterprise the participants worked for, the application frequency of (53-53) in casual speech is higher than it is in passage reading, that is, fewer Tianjin features are kept when more attention is paid to speech, which agrees with Gao & Lu (2003)'s reporting.

Gu & Liu (2003) and Gao & Lu (2003) attribute the loss of traditional Tianjin features to the standardizing influence of SC, but neither show any counter influence such as local identity and local prestige that maintains or recycles some features of Tianjin dialect. Both directly predicted that Tianjin features would disappear completely in the future. Different from these two studies, Feng (2007/2015) noticed the force of attitude to Tianjin dialect, and proposed that the attitude can result in maintaining more salient feature like 'peculiar consonant'. But it is unexplained why this attitude does not affect the application of FF rule, which is also a salient feature of Tianjin dialect. More important is that the attention-to-speech model (Labov, 1966) in Feng (2007/2015) seems to have no control over the influences from audience, speaker, environment and the relation between group members.

Even the latest sociolinguistic study on Tianjin was made 10 years ago. Tianjin speech community and Tianjin dialect have changed since then. As shown in §2.3, Tianjin witnessed a sharp rise in migrants during 2005-2016, but the influences of migrants on Tianjin locals and Tianjin dialect have not been studied. The current dissertation studies how local identity and

attitudes to language affect the change and variation of Tianjin tone sandhi under standardization.

In the next chapter, I describe the methods I used to conduct the study.

Chapter 4 Data collection and methods

One purpose of the present dissertation is to study the distribution of Tianjin dialect tone sandhi patterns across generations, social class and gender. In variationist sociolinguistics, by studying inter- and intra-speaker variation, we can discern different social meanings of different variants, and thereby make some inferences about the causes of the change and variation. This chapter consists of three sections. Section 4.1 introduces the participant sample, including recruitment methods and demographics of the participants. Section 4.2 reports on interviews, providing information on the methods and questions. Section 4.3 presents impressionistic coding, token extraction and statistical analysis methods. The data come from 90 native Tianjin speakers in Tianjin, China, interviewed in the summers of 2014 - 2017.

Although the sociolinguistic interview is an effective method for studying stylistic variation (Wolfram, 1969; Labov, 1972; Trudgill, 1972; Reid, 1978), no sociolinguistic interview can avoid the Observer's Paradox. That is, the interviewer's purpose is to observe how people talk while they are not observed, but the only way to find it out is to observe them. Although the Observer's Paradox cannot be solved completely, it can be minimized. Since the early 1960s, sociolinguists have developed a number of techniques for minimizing the Observer's Paradox (e.g. Labov, Cohen & Robins, 1965; Labov, 1966; Wolfram & Fasold, 1974). Many are summarized in Labov's (1984) description of the methods used for the Language Change and Variation project conducted in Philadelphia in the 1970s.

According to Labov (1984), the development in sociolinguistic interview involves two technical devices: a. the module and b. the conversational network. The module is a group of questions focusing on a particular topic such as school, game, fight, fear etc. The modules are combined into a conversational network by the interviewer. The requirements of the interview are

more detailed and specified. For example, “One governing principle is that module questions should take less than five seconds to deliver and in many cases, less than one second” (Labov 1984:34). The purpose of this requirements is to make the role of the interviewer as unobtrusive as possible. The modules are combined into a conversational network by the interviewer. There is no single pathway through the modules, but interviewers are guided by the participant’s own interests. By moving through different modules, the speaker may immerse in his or her thoughts or feelings, and the effects of the interviewer are minimized. The speakers’ speech may also shift between different styles while their thoughts or emotions shift between happiness, excitement, anger and fear. In a typical sociolinguistic interview, the interviewer is willing at the first opportunity to give up control over the conversation and follow the subject’s lead, and ideally an hour or more of recorded spontaneous speech is the result of the effort (Labov, 2001b: 88). This is also a possible way leading to “vernacular” mode of speech, in which the minimum attention is paid to the speech (Labov, 1984).

I select interview method instead of experimental techniques because “face-to-face interviews are the only means of obtaining the volume and quality of recorded speech that is needed for quantitative analysis” (Labov, 1984: 29). This attention to speech model was criticized because of its unidimensional focus on the linear relation between attention and style-shift (Bell, 1984; Coupland, 1985; Eckert & Rickford, 2001). An ‘audience design’ model was put forward in Bell (1984) as an alternative perspective, indicating that style-shifts occur to suit the audience. “In audience design, speakers accommodate primarily to their addressee” (Bell, 1984:145). The fact that I adopt the sociolinguistic interview as the main field method of the dissertation does not imply that the effects of audiences are ignored. Instead I agree ‘both adaptation to different audiences and different degrees of audio-monitoring are involved in style-shifting’ (Labov, 2001b:87). In the

interviews carried out in the present study, the audience was always the author, with no other individuals present besides the interviewee, so the effects of the audience are well controlled across participants. The questions to every participant are at least very similar, so the influences from topic (Sharma, 2018) are also controlled.

4.1 Speaker sample

90 native speakers of Tianjin dialect were recorded for this study. In this section, I describe my sampling method and eligibility criteria, the demographic design of the sample, and how I recruited participants.

4.1.1 Sampling method

84 of the 90 speakers were included in the study via the ‘friend of a friend’ or ‘snowball sampling’ method (cf Milroy L.,1980). The ‘friend of a friend’ method gave me some connection with participants. In other words, to name a friend made the speakers trust me and made them comparatively relaxed in the interview.

Random sampling can provide for broad sociological representation of the speech community, but it is extremely difficult to find the subsection of the “right” population. That is, it is hard to end up with a sample that fits into a specific demographic class that the researcher needs, as such a subsection might be ‘geographically and socially distributed amongst the population in a non-random way’ (Milroy, 1980:24). On the other hand, sociolinguistic studies have indicated that relatively small samples are adequate to account for language variation in a large speech community. “Speech communities tend to consist of many varieties spoken by groups containing very different numbers of individuals, so that uniform sampling leads to redundancy for some groups and risks missing others entirely (Sankoff, 1988:900). Since “even so called random sampling used in linguistic surveys was not ‘random’ in the strict sociological sense anyway”

(Tagliamonte S. , 2006:22), I follow many prior studies in using the friend-of-a-friend technique but with a pre-planned sample design, so that I only included participants who met the right eligibility criteria, and who could help ensure that each subgroup by age, sex and social class had enough representatives in the sample.

Most (n=42) young speakers were recruited from Tianjin Normal University. I recruited younger speakers from Tianjin Normal University because I had some connections with that university, and therefore it was comparatively convenient for me to recruit young speakers from that university. Although Tianjin Normal University attracts students from across China, I only interviewed students who were from Tianjin itself. I recruited by saying that I was looking for people who had lived in Tianjin from childhood onwards. Other speakers (n = 48) were recruited from among my personal friends, or friends of friends.

Six of the 90 participants had previously been selected as “maintainers” of Tianjin dialect by Tianjin Archives and Tianjin TV Station. In 2014, Tianjin Archives started a project aiming to preserve and record Tianjin dialect. They advertised in the main media in Tianjin (Yu & Zhang, 2015; Wang L., 2016). Some experts on Tianjin dialect were invited to select “pure” Tianjin speakers from the participants. They were invited to make recordings at Tianjin TV Station and their recordings might be kept in Tianjin Archives as samples of Tianjin dialect. Six of those “maintainers” participated in my fieldwork.

I recruited these “pure” speakers because at that time, I seemed to have drained my personal social network in finding enough old (aged 65+) Tianjin speakers. I had heard of the Tianjin dialect voice archiving project sponsored by Tianjin Archives and Tianjin TV Station. I originally wanted to obtain the recordings from Tianjin Archives, but I was told the project had not been completed and they could not lend the recordings to me. They gave me the contact information of two experts

on Tianjin dialect (Mr. Wenjie Guo and Prof. Ruwei Tan) who could contact some of the “maintainers”. Mr. Wenjie Guo helped me to contact six “maintainers” and they all agreed to take part in the present study.

In order to avoid “performance speech” (Schilling-Estes, 1998) when working with “pure” dialect speakers, I analyzed the data from the six “pure” Tianjin speakers separately before I included them with the data from other old speakers. No significant difference was observed between the results of the six “pure” Tianjin speakers and other old participants over the application of the three variables, so their data are aggregated in the final analysis.

4.1.2 Eligibility criteria

As introduced in Chapter 3, Tianjin dialect is mainly spoken in the six inner-city “central” districts, so I only recruited speakers from those districts. All the participants speak Tianjin dialect as their native language, though most speakers can speak SC too, which is unavoidable because of the promotion of SC described earlier. None of the participants spoke any other Chinese dialect, and none of them had ever lived in other language environments in their childhood. This is consistent with the Chinese fieldwork tradition, which always asks participants’ childhood language environment. I only recruited adults, and thus the minimum age of participants is 18.

4.1.3 Demographic characteristics

The sample was designed to include three generations of Tianjin speakers, male and female, across two social classes. The age, gender and social class distributions of the participants are given in Table 4.1.

Table 4.1 Entire participant sample by age, social class and gender (N=90)

	Total	Middle class		Working class	
		Male	Female	Male	Female
Young (<35)	42	7	15	10	10
Middle-aged (35-65)	25	5	5	7	8
Old (>65)	23	5	5	8	5
Total	90	17	25	25	23

Participants were well-balanced for gender and social class. Although young speakers still outnumber middle-aged and old speakers⁴, each cell by age, social class and gender has at least five participants.

4.1.3.1 Age

The cutoff age between middle-aged and older speakers of 65 years is not arbitrary. In the investigation, I noticed that there seemed to be a gap in linguistic behavior between speakers in their earlier 60s and late 60s, especially in the application of (53-53). People in their early sixties were born after 1949. As explained in Chapter 3, in 1956, the Chinese government adopted the policy of spreading and extending the use of SC. According to this policy, all students and teachers were required to speak SC in school and were not supposed to speak any other dialect or language in school. It took several years for the policy to be carried out in all levels of school. People in their early sixties now are the first generation who went to school after the policy had been adopted

⁴ The dissertation is developed from a previous study (Wang 2016) which was mainly about the applications of (53-21) and (53-53) among young people, with the middle-aged and old people groups were included for comparison in apparent time, to confirm trends observed in the earlier literature. In order to better understand the social motivations for these trends, young people were deliberately over-sampled in that previous study.

and carried out. By contrast, people in their late 60s at the time of data collection would have been unlikely to have received such an immersive education in SC. Henceforward I refer to these as the Old group; the others are Young and Middle-aged. Although the eligibility for young age group is < 35, most young speakers in my speaker sample are aged between 18-22, and I do not have anybody aged 23-34, except #38⁵, who was 30 years old when she was interviewed in 2014. This is not a deliberate decision. It is because I do not have many individuals between 22 and 35 in my personal social network.

4.1.3.2 Social class

China's Communist history and the present-day speed of its economic transformation means that it was difficult to determine speaker social status for this study. I chose to classify participants based on their occupation, education and income, which is in harmony with Labov's sociological formula for measuring social class by combining indicators of occupation (of the family 'bread-winner'), education of the speaker and income of the family (Labov, 1972).

All speakers are classified into two classes: Working class and Middle class. I set criteria for occupation, education and income for Middle class and Working class, mainly based on my intuition, also consulting the classification system of Li (2013). I put factory workers, skilled workers, hairdressers, kindergarten teachers, drivers and vendors into the Working class. Office workers, government employees, technical personnel, nurses, primary school teachers, middle-school teachers, administrative personnel, doctors, lawyers, professors, businessmen and

⁵ #38 (participant 38), was 30 years old when she was interviewed in 2014. Her speech does not show any difference in the use of (53-53), (53-21) and (21-21) from other young speakers. Henceforward each participant in the present study is identified by a unique number, preceded by the pound sign #.

government officials are put into the Middle class. Since all the Young participants are undergraduates and not yet in the workforce, they were classified by their parents' social class.

However, a participant was classified as belonging to the Middle class only when he or she satisfied criteria for education and income as well as occupation. For example, if a participant's income was high enough to qualify him for the middle class, but his/her education was below junior high level, that participant would not be classified as belonging to the middle class. The basic education requirement for middle class categorization, following Li (2013), was graduation from junior high school. In comparison with US social class characterizations, this might seem like very little education for classification as middle class. However, in China, some very successful businessmen have not had much education. In the present study, one participant's father is a very successful businessman, but the father had only completed his junior high school. These successful businessmen are considered to be middle class (Li, 2013: 69). I give businessmen special description because unlike other occupations, this occupation is not so greatly affected by people's education. People have to have a good education before they can become doctors, lawyers, professors and government employees. These occupations can reflect people's education. But anyone in China can start a business, and there are no minimum educational requirements for businessmen.

As for income, Li (2013) classifies as middle class any participant whose income is higher than the average income of the residential area. In this study, I classify as middle class any participant whose income is higher than the average income of Tianjin, so long as that participant also had an education at the junior high level or above. For example, the yearly average income of Tianjin is about 52000 yuan per person (Tianjin Municipal Bureau of Statistics & Survey Office

of the National Bureau of Statistics in Tianjin, 2018), which is about 8200 US dollars. The income requirement for middle class was therefore set as higher than 52000 yuan.

Although people can talk quite freely about their income in China, I did not have to ask every participant their income because the income of most occupations is quite clear to most locals (and to me). I only asked or directed the conversation towards this topic when I was not sure whether his or her income was higher than the average income and could not decide whether he or she belonged to the middle class.

4.1.4 Attitude to migrants and attitude to Tianjin dialect

Since one hypothesis of the current dissertation is that local identity triggered by migrants is the main force that maintains local sandhi features, while negative attitude to Tianjin dialect accelerates the standardization process of the local features, two socio-psychological factors, attitude to migrants and attitude to Tianjin dialect, are added in addition to traditional demographic factors such as age, social class and gender.

4.1.4.1 Attitude to migrants

My own personal experiences and my fieldwork interviews with Tianjin natives have revealed that they have varying attitudes towards migrants. Some think migrants are necessary to the development of the city. Some do not like the migrants because they believe that the migrants make the city more crowded, dirty and unsafe. They also think the migrants have taken some opportunities away from the city natives.

In every interview, I asked: “Do you like migrants?” “Why or Why not?” Based on the answers to these questions, I made a general subjective, qualitative assessment of each participants’ attitudes to migrants. I placed every participant into one of two categories: negative--expresses overall negative attitudes towards migrants; positive—expresses overall positive attitudes towards

migrants. Forty-eight participants are classified as having negative attitudes towards migrants, and forty-two participants are judged as having positive attitudes towards migrants. The distributions of participants by attitude to migrants and age group, attitude to migrants and social class, attitude to migrants and gender, and attitude to migrants and attitude to Tianjin dialect are given in Tables 4.2-4.5.

Table 4.2 The distribution of participants by attitude to migrants and age group

Attitude Age group	Negative	Positive
Young (42)	29	13
Middle (25)	9	16
Old (23)	10	13
Total (90)	48	42

Table 4.3 The distribution of participants by attitude to migrants and social class

Attitude Social class	Negative	Positive
Middle class (42)	22	20
Working class (48)	26	22
Total (90)	48	42

Table 4.4 The distribution of participants by attitude to migrants and gender

Attitude Gender	Negative	Positive
Female (48)	31	17
Male (42)	17	25
Total (90)	48	42

Table 4.5 The distribution of participants by attitude to migrants and attitude to Tianjin

Tianjin Migrants	Negative	Positive
Negative to Tianjin (31)	13	18
Positive to Tianjin (59)	35	24
Total (90)	48	42

Although all participants' attitudes to migrants are classified into two categories, quite a lot of participants talked about both the positive and negative aspects of migrants.

Wailai renkou gongtong jianshe zanmen Tianjin, zheshi hao de yi mian.

'Migrants came to build our Tianjin together with our locals, which is the good aspect'.

Danshi ye you bu hao de. Louxi ye dai dao le Tianjin. Erqie, huanjing chengzai li, Tianjin jiu zheme da dir, ranhou lai le zheme duo ren. Zhe doushi wenti. Huanjing ye hui you yali.

'But there are bad aspects. Together with migrants, their bad habits were brought to Tianjin too. Besides, the pressure on environment. Tianjin area is fixed. It cannot expand although so many migrants came. These are all problems. Environments will be under pressure'.

Bu yao zhaxie ren ye buxing, tamen dou lai le ba, ye buxing.

'It is not good not to allow any of them in, but it is not good either to have them all come'.

(Quotes from #29, a young working class girl)

Her words seem to be unbiased towards migrants, but immediately after this, she mentioned migrants with Blueprint Hukou were banned to take the entrance exam for colleges and universities at Tianjin, which is a local policy very unfavorable to migrants. Her excitement at this policy betrayed her and she said directly this policy was good to Tianjin locals, not good to migrants. She welcomed the policy. So I classified her into the negative category.

4.1.4.2 Attitude to Tianjin dialect

Participants are quantitatively classified as having positive attitudes to Tianjin dialect or negative attitudes to Tianjin dialect, based on the language attitude questions given at the end of each interview, summarized in (2).

(2) Language attitude questions given at the end of each interview

- a. Is it pleasant to hear Tianjin dialect?
- b. Do you prefer to speak Tianjin dialect at home?
- c. Do you think people speaking Tianjin dialect are less-educated or/and from a lower social class?
- d. Do/Will you allow your children to speak Tianjin dialect?
- e. Do/Will you allow your grandchildren to speak Tianjin dialect?

Table 4.6 displays the guide I used to score each of the responses to questions (2a-e) for each participant. I did not ask Young speakers questions d and e, so they are irrelevant to these questions. If the participant cannot give an answer to a question, it is not clear. Scores were summed, giving an overall score for attitudes to Tianjin dialect for each individual.

Table 4.6 Scoring guide for the questions about Tianjin dialect

	Yes	No	Irrelevant	Not clear
a	1	-1	0	0
b	1	-1	0	0
c	-1	1	0	0
d	1	-1	0	0
e	1	-1	0	0

If a participant's score is higher than 0, that participant is considered as having positive attitude to Tianjin dialect. If a participant's score is lower than 0, that participant is considered as having negative attitude to Tianjin dialect. The scores for every participant are provided in Appendix III.

Fifty-nine participants have a positive attitude to Tianjin dialect, while thirty-one participants have a negative attitude to Tianjin dialect. The proportion of participants who feel positively about Tianjin is almost double the proportion of those holding a negative attitude to Tianjin, but this does not imply that most Tianjin natives have positive attitude to Tianjin, or that the Tianjin speech community generally views Tianjin dialect positively. Because the primary eligible condition for all participants was that they were a native speaker of Tianjin dialect, those Tianjin natives who do not speak Tianjin dialect were not recruited, and therefore their attitudes to Tianjin dialect are not included. So the attitude investigation here is not balanced. But on the other hand, the purpose of studying language attitude here is to relate a speaker's linguistic variability uses to his/her language attitude, so in this sense, the attitude study here fits the purpose of the study.

The distributions of participants by attitude to Tianjin and age group, attitude to Tianjin and social class, and attitude to Tianjin and gender, are given in Tables 4.7-4.9.

Table 4.7 The distribution of participants by attitude to Tianjin and age group

Attitude Age group	Negative to Tianjin	Positive to Tianjin
Young (42)	15	27
Middle (25)	11	14
Old (23)	5	18
Total (90)	31	59

Table 4.8 The distribution of participants by attitude to Tianjin and social class

Attitude Social class	Negative to Tianjin	Positive to Tianjin
Middle class (42)	23	19
Working class (48)	8	40
Total (90)	31	59

Table 4.9 The distribution of participants by attitude to Tianjin and gender

Attitude Gender	Negative to Tianjin	Positive to Tianjin
Female (48)	18	30
Male (42)	13	29
Total (90)	31	59

4.2 Interview and word list

The field method adopted in this study is the sociolinguistic interview, which included a word list reading, with a purpose of obtaining different styles of speech so as to study intraspeaker variation as well as interspeaker variation (Labov, 1972). Stylistic variation is an important factor

in confirming whether a change is from below the public awareness or above the public awareness (Labov, 1994). People tend to use more prestigious variants more frequently when they pay more attention to their speech, as a result, if the more prestigious variant appears more frequently in the more careful speech, it indicates that the speaker are conscious of the change, and therefore it is a change from above. If more prestigious variant does not appear more frequently in the more careful speech, the speaker may not be conscious of the change, and therefore it is a change from below.

Previous sociolinguistic studies have observed stylistic variation in various interview situations (Labov, 1972; Trudgill, 1974). Attention to speech is one contributor to stylistic variation, and this contributor that can be captured fairly easily and quantified fairly easily by varying the tasks in the interview. Speakers became more aware of their own ways of speaking as the interview activities moved along a notional scale towards greater formality. They were less attentive to their speech in ‘casual’ style, more attentive in ‘careful’ style, and more attentive in more ‘careful’ word list reading style. Speakers tend to use more prestigious variant in their more ‘careful’ word list reading style than in less ‘careful’ interview style, while they tend to avoid less prestigious variant in their more ‘careful’ style. As a result, we can obtain information as to which variation is prestigious or which is stereotyped and avoided by speakers and whether a change is from above or from below by comparing less ‘careful’ interview style with more ‘careful’ word list reading style.

The interviews were conducted in a quiet room, and the interviewee and I were the only people present in the room during each interview. A Zoom H1 Handy Recorder (version 2.10) with an Audio-technic Pro-70 lavalier cardioid condenser microphone was used for audio-recording.

4.2.1 The interview

The interviews lasted about 45-65 minutes for each speaker. The participants and I all spoke Tianjin dialect. I gave up control over the conversation at the first opportunity and followed the subject's lead, with a purpose of eliciting maximally unmonitored speech from the subject. All interviews started with demographic questions and ended with questions about Tianjin dialect and attitudes towards Tianjin dialect and SC.

Using the method introduced in Labov (1984), the interview questions are divided into different modules and the modules were combined into conversational networks by me. Before the beginning of the interview, I introduced myself as a Tianjin native and also a researcher from Michigan State University. I said I had observed big changes in the development of Tianjin and I wanted to know how the locals viewed the development of the city. I said that I also wanted to know about the changes to Tianjin dialect. After demographic questions like 'How old are you?', 'Where do you live?', I moved on to 'peers' for young speakers since they communicate with their classmates and dormmates every day and their high school memories are fresh. I usually started by asking how they typically communicated with their friends, classmates or dormmates, and then I asked them how they got along with their peers. The order of the following modules such 'fights', 'dating', 'family', 'school' and 'work' was different depending on the interviewee, but every interview ended with 'migrants' and 'language'.

I selected 'peers' as the module immediately after the 'demographic' module for each interview because it is not too private. It is comparatively easy to talk about classmates, dormmates or friends with a person just acquainted than to talk about family, boyfriends or girlfriends. 'Fights' and 'dating' are the two most 'dangerous' modules with regard to participant sensitivity, but are also the modules that could have unexpected gains. Not all young speakers wanted to talk about

‘fights’ or ‘dating’. It could be that they classified ‘fights’ as something negative and ‘dating’ as too private to share with the interviewer. Some of them did not have a boyfriend or a girlfriend. But on the other hand, if the young speaker was willing to talk about at least one of the two modules, the results sometimes were unexpectedly good. For example, # 21 talked about her conflicts with one of her dormmates. Six girls lived in one room and they each had a bed, a cabinet and a table. The dormmates were studying different subjects, and therefore they had different class schedules. P21 said one of her dormmates was very bossy. If that girl was going to go to bed, she would turn off all the lights in the dorm no matter whether other girls were using them.

*Ta buguan neitou de ren yong bu yong dixia de jingzi a, yong bu yong nei liangr,
ta jiu bu guan, ta jiu gei ni guan le.*

‘She did not care whether people on that side were using the mirror or they need the light. She just turned it off’.

(Quotes from # 21)

If other girls got up early on the morning, she would not allow them to turn on the lights in the room and make any noise to disturb her. But if the bossy girl had class in the morning, she was free to talk loudly and turn on the lights in the dorm!

*Danshi, tamen yaoshi zaoqi, women sa shuijiao, ta he nage ninghe xiao nvhai ta
gen ta shuohua shi gai duoda yin you duo da yin, buguan womensa shuijiao
zenmeyang.*

‘But, if they got up early while we three were asleep, she and the girl from Ninghe talked as loudly as they like. They did not care whether their talking would disturb our sleep’

(Quotes from # 21)

This was just one example of the bad behaviors of the bossy girl. When # 21 listed the bossy girl's misdeeds, her speech became much faster and her face turned red. She talked about for a long time without any pause. She seemed to be so agitated and lost in her anger that she somewhat ignored my presence.

For the Middle-aged and Old speakers, I usually directed the talk to their jobs or former jobs after demographic questions. Most participants are willing to share their work experiences. Then the talk moves to the next module family, which starts from children or grandchildren since all participants are talkative in talking about their children or grandchildren. The modules and questions for every age group are given in Appendix I.

During the interviews, I initiated topics by asking questions, but I would follow the participant's interests and ideas wherever they went. I spoke as little as possible. I did not utter a word unless it was necessary. Most of the interviewees were willing to respond to interview questions and they talked freely. If the interviewee was not willing to speak about a specific topic, I usually tried a different topic. In this way, I could eventually find some topic in which the speaker was interested. All speakers were willing to recount their personal experiences. It was not very difficult to record a continuous speech for each interview.

4.2.2 The word list

After the interview and a ten-minute break, all participants were asked to read the word list in Tianjin dialect. All the words were written on cards, one card per word. Each word was read one time by each participant. For each speaker, the order of the cards was randomized by hand. The word list is given in Appendix II.

The word list contains 124 words. Tianjin has four lexical tones, and therefore it has a total of sixteen disyllabic combination patterns. All the four lexical tones and the sixteen possible

disyllabic patterns are included in the word list. The word list contains two parts: (i) twelve monosyllabic words, three for each of the four lexical tones; (ii) 112 disyllabic sequences, twenty for each of the three variables and four for each of the other thirteen disyllabic patterns respectively. All test words were common words based on frequency counts in the Modern Chinese Frequency Dictionary (1986)⁶. Although the intention is to obtain the outputs of the three variables, the lexical tones and other thirteen disyllabic patterns were included to avoid persistence effect (Woods, 1986) and to prevent participants from paying special attention to the three target variables.

Although the studies on Tianjin tone sandhi have not shown morpho-syntactic influences on disyllabic tone sandhi patterns, morpho-syntactic effects were taken into consideration when test words were selected. All test patterns include different parts of speech. For example, (21-21) has *fei ji* ‘airplane’ which is a noun, and it also has *kai che* ‘drive a vehicle’ which is a verb.

20 words on the list contained the variable context for (53-21), 20 words contained the variable context for (53-53) and 20 words contained the variable context for (21-21). *Sheng yin* (21-21) ‘sound’ and *xin xian* (21-21) ‘fresh’ in the word list end with neutral tone in Tianjin. Disyllabic tone sandhi does not apply if the second syllable carries a neutral tone (J. Wang, 2002). So words for (21-21) for each participant are effectively 18, instead of 20. When I designed the word list, I did not realize that the two words end in neutral tone in Tianjin. Otherwise they would not have been included in the word list. All legitimate tokens of (21-21), (53-21) and (53-53) from the wordlist were included in the analysis, for a total of 58 tokens per speaker.

⁶ Since there are no frequency dictionary or corpora based on spoken Tianjin, I use the frequency dictionary based on Standard Chinese to approximate the frequency of Tianjin words. The justifications for this are: (i) there is a considerable overlap in vocabulary between SC and Tianjin; (ii) Chinese dialects have the same orthographic tradition. Zhang & Liu (2011) use similar points to defend their use of a SC corpus for a Tianjin study. Although the dictionary was published in 1986, it is the only Chinese frequency dictionary published and still commonly used in Chinese linguistic studies.

4.3 Methodology

4.3.1 Impressionistic coding and inter-rater reliability

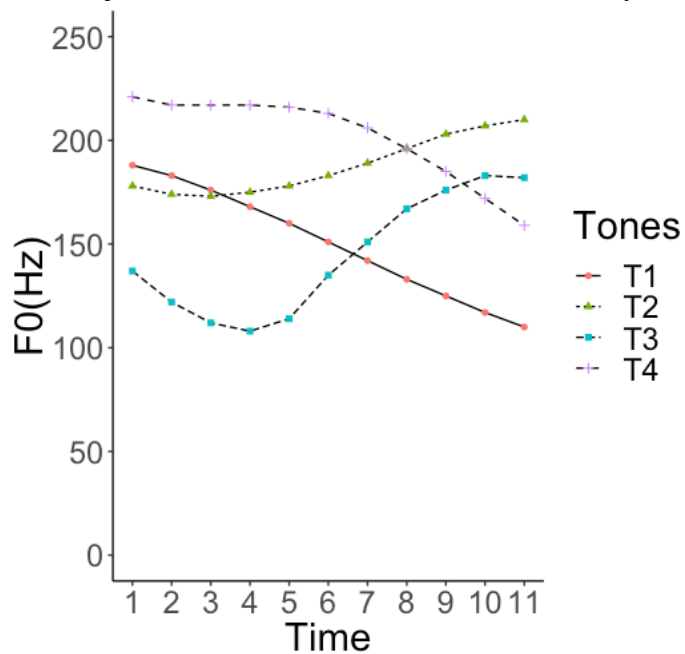
The data were coded impressionistically for application or non-application of (53-53), (53-21) and (21-21) tone sandhi. Both interview and word list recording were transcribed in ELAN 5.0 (Max Planck Institute for Psycholinguistics, 2016), with the transcription and the coding on the same tier. The entire interview for every participant is transcribed.

After impressionistically coding 20 word list tokens of each variable in ELAN, I exported the transcription as a Praat textgrid file, and then opened the textgrid file together with the corresponding sound file in Praat (Boersma & Weenink, 2014).

A script was used for F0 extraction. For each token, the F0 at every 10% of the rhyme duration was extracted, giving eleven F0 measurements for each syllable. All the tokens for each test tone pattern were averaged.

Here I use the young group as an example to show how the acoustic analysis was conducted to the check the impressionistic coding. I selected 16 participants (male = 8, female = 8) randomly from the young group and conducted acoustic analysis of their word list. In order to show the difference between the changed tone and lexical tones, I first analyzed their lexical tones and the results are given in Figure 4.1.

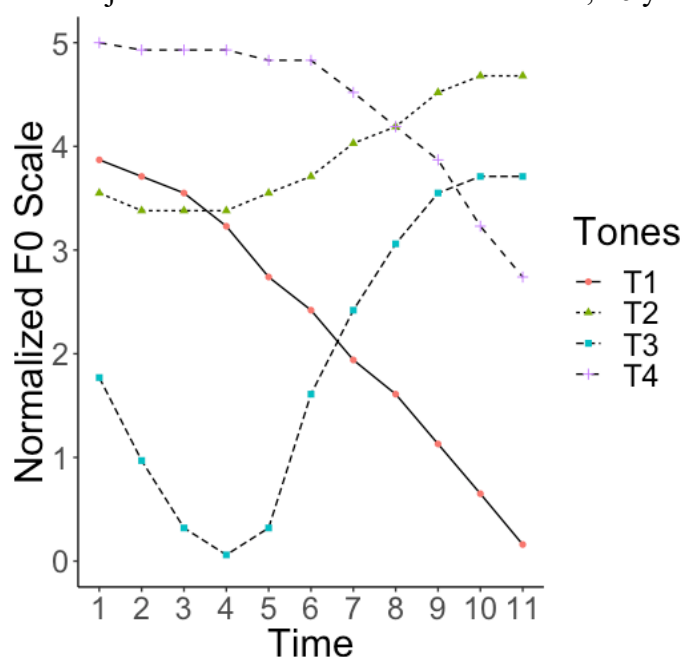
Figure 4.1 Tianjin lexical tones in raw Hertz values, 16 young speakers



For each tone, the raw F0 of each of the eleven points is normalized into Chao number by using the formula put forward in Shi (1986), as shown in Formula (1). Log_x is the testing point. $\text{Log}_{\max}$ is the highest value of the average value of all points and $\text{log}_{\min}$ is the lowest value of the average value of all points. The normalized results are shown in Figure 4.2.

$$\text{Formula (1)} \quad F0_{\text{normalized}} = \{[\text{Log}_x - \text{Log}_{\min}] / [\text{Log}_{\max} - \text{Log}_{\min}]\} \times 5$$

Figure 4.2 Tianjin lexical tones in normalized results, 16 young speakers



The normalized results indicate that T1 is 41, T2 is 45, T3 is 213 and T4 is 53 in Chao number in the Young subsample. T1 was 21/11 in previous studies on Tianjin tones (Li & Liu, 1985; Shi, 1988), but its starting point is 41 among young speakers now, as shown in Figure 4.2 above.

In the present study, I still use LL to represent T1 because it is still a low falling tone comparing to the high falling T4(53). Besides, the coding and selected acoustic analysis show consistently that the starting point of T1 remains low among old speakers, but T1 in old group is 31.

The raw and normalized F0 results show approximately the same pattern, which echoes Zhang & Liu (2011) that the normalized pattern with this formula is similar to the pattern of F0 results. The normalized results are in Chao numbers, so I can check whether the acoustic results are consistent with the impressionistic coding results.

Since the raw hertz results show the same pattern with normalized results, I will only present the raw hertz results for the four traditional tone sandhi patterns, which are given in Figures 4.3-4.6.

Figure 4.3 The output of T4+T4 → T1+T4 in raw Hertz

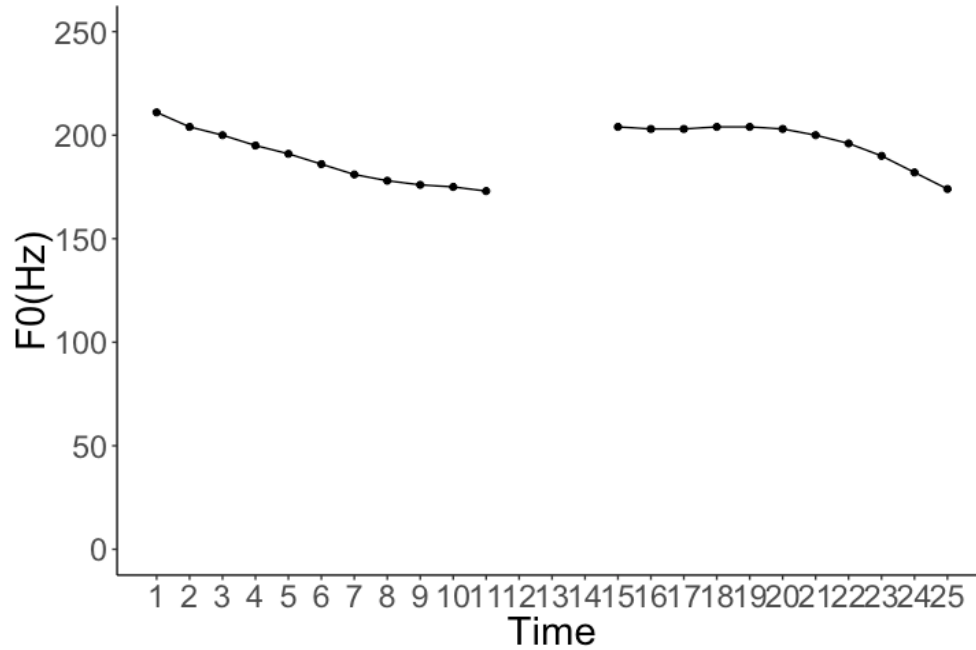


Figure 4.4 The output of T4+T1 → T2+T1 in raw Hertz

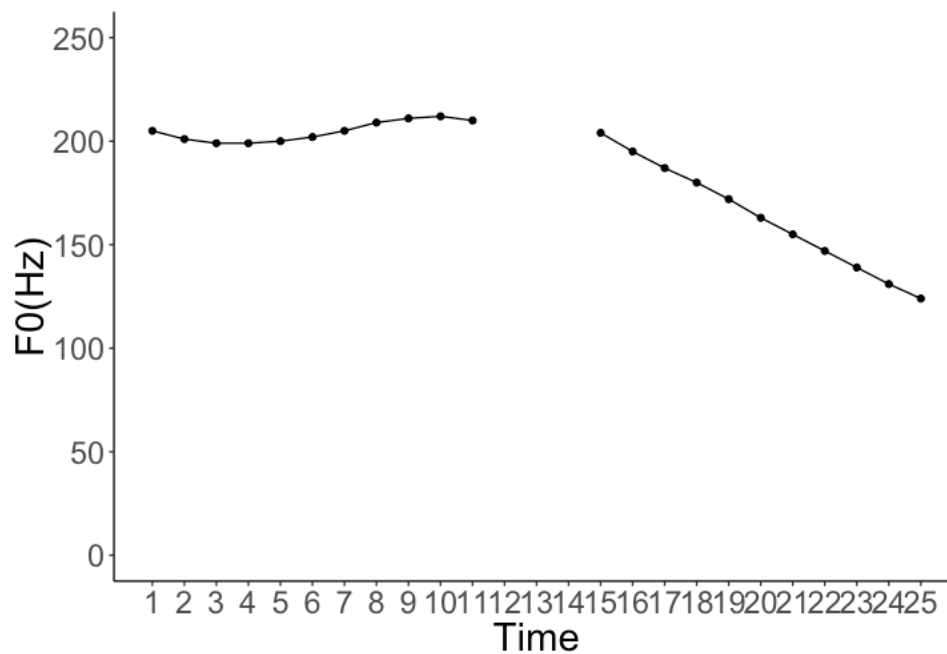


Figure 4.5 The output of $T1+T1 \rightarrow T2+T1$ in raw Hertz

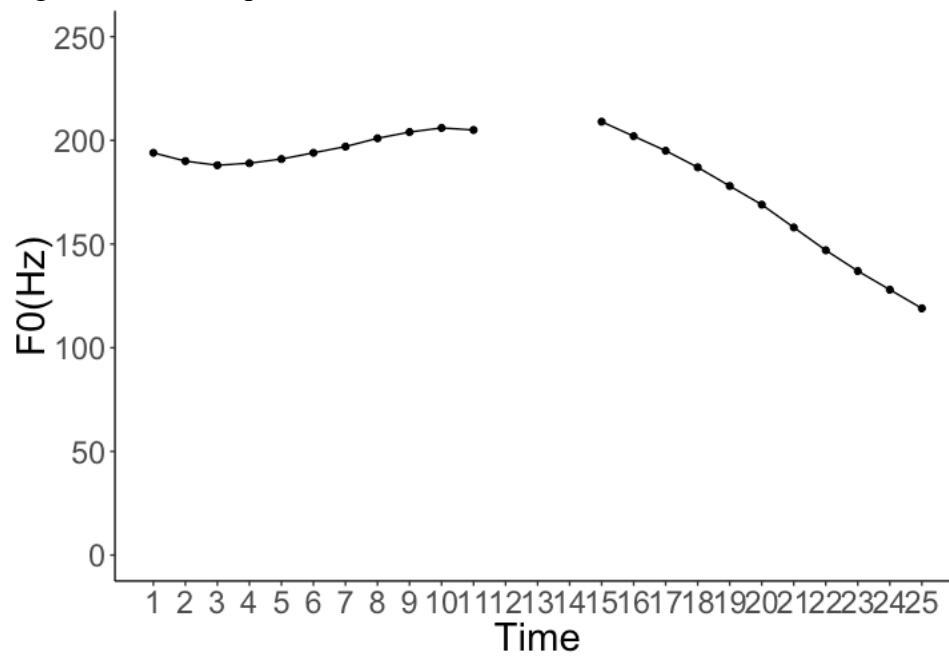
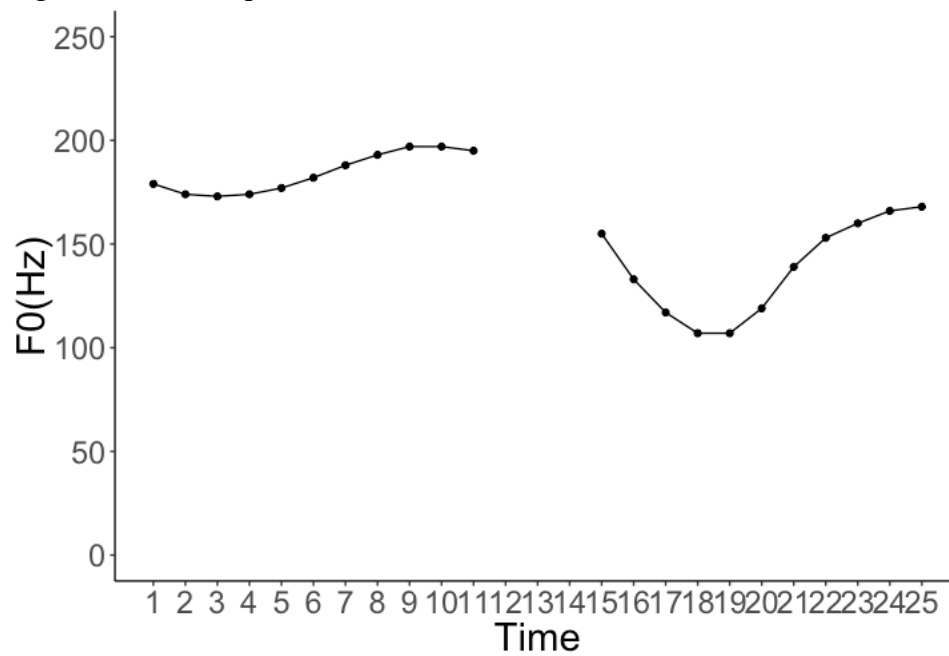


Figure 4.6 The output of $T3+T3 \rightarrow T2+T3$ in raw Hertz



Overall, there was a high degree of consistency between my impressionistic coding and the instrumental analysis. The proportion of accuracy was very high: (20/20) 100% for (53-21), (19/20)

95% for (53-53) and (17/18)⁷ 94% for (21-21) on average for each participant. I therefore coded all of the rest of the data impressionistically. A subsample of the coding of the word list and interview were double-checked by an experienced fieldworker.

Impressionistic coding was also reliable because of the tones involved. The differences between the variants of each tone variable are self-evident. In (53-21), the underlying tone of the first syllable is HL (53), which is a high falling tone. But the derived tone after tone sandhi application is HH (35), which is a high rising tone. One is rising, while the other is falling. It is not difficult to tell rising and falling apart. In (53-53), the underlying tone of the first syllable is HL (53), which is a high falling tone. But the derived tone after tone sandhi application is LL (11/21), which is a low level tone. When I coded (53-53), I paid attention to the starting point of the first syllable. If it was high, I took it as HL (53), and no tone sandhi applied. If the starting point of the first syllable was not so high, I took it as LL (21), and tone sandhi applied. The two application variants of (21-21) are the most difficult. One application variant is HH (35), while the other application variant is LH (213) or (13). HH (35) is a rising tone in high register, while LH (213) or (13) is a rising tone in low register. If the starting point is high, the tone is coded as HH (35). If the starting point is low and usually a turning point perceived, the tone is coded as LH (213) or (13).

4.3.2 Extraction of tokens

For the sociolinguistic interview, the audio recording was exhaustively mined for all of the linguistic environments in which the relevant three Tianjin sandhi patterns would be expected to

⁷ *Shengyin* ‘sound’ and *xinxian* ‘fresh’ in the word list end with neutral tone in Tianjin. So effective words for (21-21) for each participant are 18, instead of 20.

obtain. A total of 3876 tokens of (53-21), 5662 tokens of (53-53) and 2497 tokens of (21-21) were extracted, for a total of 12035 effective tokens in interview style.

For the word list reading, a total of 1800 tokens of (53-21), 1800 tokens of (53-53) and 1800 tokens of (21-21) were extracted, but as previously mentioned, a total of 181 tokens had to be excluded because of neutral tone or mispronunciation, reducing the overall total to 1800 tokens of (53-21), 1799 tokens of (53-53) and 1622 tokens of (21-21).

4.3.3 Statistical analysis

Mixed model logistic regression was carried out to get a comprehensive picture showing the simultaneous effect of sex, class, age and style on the tonal changes. Whether attitudes to migrants and to Tianjin dialects affect the use of language is an important part of the dissertation, so attitudes to migrants and Tianjin dialects are also independent variables in the study. The results of the word list and sociolinguistic interview analysis will be presented in Chapter 5 and Chapter 6 respectively. Chapter 7 provides further interpretation and comparison of the two sets of results.

Chapter 5 Word list Results

This chapter reports on the distribution of the variants of (53-53), (53-21) and (21-21) produced by the participants when reading the word list. Previous acoustic studies on Tianjin tone sandhi (Shi & P. Wang, 2004; Zhang & Liu, 2011; Li & Chen, 2016) are all based on data from word list/sentence reading style. By studying data from word list reading style, the present dissertation confirms the generational changes in Tianjin tone sandhi observed in previous real time studies. Additionally, the analysis of word list style in this chapter lays the foundation for a comparison with the less careful interview style (Chapter 6), to be discussed in Chapter 7. The intra-speaker variation analyzed in Chapter 7 will reveal speakers' attitudes towards the local dialect. For now, the present chapter provides a descriptive overview of the inter-speaker distribution of (53-53), (53-21) and (21-21) variants by age, social class and gender. Regression analyses reveal which of these social factors, including overt attitudes to migrants and to Tianjin dialect, significantly affect speakers' variant selection. A general introduction to the results of the three variables is given in §5.1. Regression model construction is presented in §5.2, followed by a detailed report of the three variables (§5.3-§5.5).

5.1 A general introduction to word list results

To begin, let us recap the main findings of the relevant studies of Tianjin tone sandhi, all of which were derived from word list data (see Chapter 3 for a previous discussion, with more detail). Traditionally Tianjin has four disyllabic tone sandhi patterns (Li, 1956; Li & Liu, 1985; Shi, 1988), but recent acoustic studies have reported variations of these tone sandhi patterns. The (53-53) sandhi pattern has been described as obsolescing or obsolete for young people (Zhang & Liu, 2011; Li & Chen, 2016). The (53-21) sandhi pattern was the least frequent among the four Tianjin sandhi changes observed in the 1980s (Shi, 1988), but it is reportedly applied almost

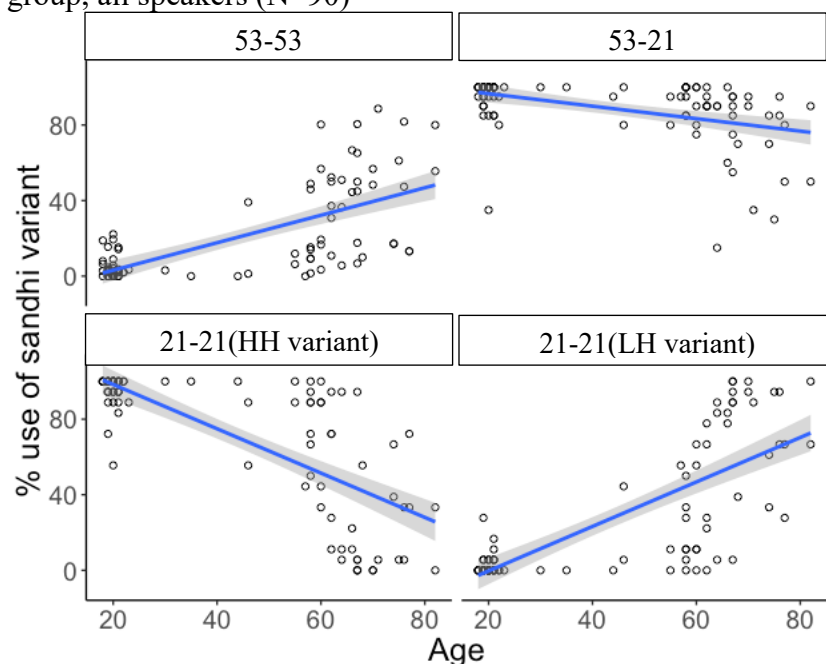
without exception among young Tianjin speakers (Shi & Wang, 2004). A third sandhi pattern, (21-21), has a new variant, HH, among young female speakers (Lu, 1997) in addition to the old variant LH reported in previous studies (Li, 1956; Li & Liu, 1985; Shi, 1988). It is not clear whether the reported variation is related to changes in these variables over the generations in apparent time. For this reason I present an overview of the distribution of variation for the word list results by age group.

The overall rates of use of traditional Tianjin sandhi patterns for (53-53), (53-21) and (21-21) are given in Table 5.1 and Figure 5.1 below. ‘App’ is an abbreviation for ‘application’, i.e. application of the tone sandhi rule (hereafter the same). Overall, as shown in Table 5.1, tone sandhi rules are applied most frequently for (21-21), at 98.6%, but the traditional local LH output variant accounts for only 29.3% of rule applications, versus 69.3% for the new variant. The local Tianjin output for (53-21) is next, at 88.1%, followed far behind by the traditional Tianjin variant of (53-53) at only 11.5%.

Table 5.1 Rates of sandhi rule application for (53-53), (53-21) and (21-21) in word list style, by age group, all speakers (N=90)

Age group	53-53		53-21		21-21 (HH variant)		21-21 (LH variant)	
	App/N	%	App/N	%	App/N	%	App/N	%
Young	28/839	3.3	801/840	95.0	733/758	96.7	18/758	2.4
Middle	39/500	7.8	372/500	89.4	298/450	66.2	145/450	32.2
Old	140/460	30.4	340/460	73.9	92/414	22.2	313/414	75.6
Total	207/1799	11.5	1586/1800	88.1	1123/1621	69.3	475/1621	29.3
					LH+HH			
					App/N	1598/1621	98.6%	

Figure 5.1 Rate of use by variant for (53-53), (53-21) and (21-21) in word list style, by age group, all speakers (N=90)



Looking at both Table 5.1 and Figure 5.1, we can see that the word list results for young speakers echo recent acoustic studies on Tianjin tone sandhi (Zhang & Liu, 2011; Li & Chen, 2016), which are mainly based on data from young and middle-aged speakers. The very low application rate of (53-53) of the Young group (3.3%) and the very high application rate of (53-21) of the Young group (95%) is consistent with these acoustic studies. We also see in Table 5.1 and Figure 5.1 that the application variant of (21-21) sandhi rule among young speakers is mainly HH (96.7%), which is consistent with Zhang and Liu (2011) and earlier Lu (1997).

But the results of the Old group are different from the Young group and Middle-aged group, and also different from recent acoustic studies (Zhang & Liu, 2011; Li & Chen, 2016). The (53-53) tone sandhi rule still exists among old speakers (application rate 30.4%), but it has practically disappeared in the Young group (3.3%). And although (53-21) applies categorically in the Young group (95%), its application rate is only 73.9% in the Old group. For (21-21), Old speakers still use the LH variant (75.6%). The Middle-aged group is closer to Young group in application

frequency in (53-53) and (53-21) variables, at 7.8% for (53-53) and 89.4% for (53-21), but its frequencies of the two variants of (21-21) are almost in the middle between Young group and Old group, at 66.2% for HH variant and 32.2% for LH variant. The age group differences in the present apparent study confirmed our prediction that the three variables have changed over time.

Although there are clear relationships between speaker age group and tone sandhi use, as shown in Table 5.1 and Figure 5.1, it is necessary to investigate the relationships between tone sandhi use and the other linguistic and social variables introduced in Chapter 4. In the following sections, I report on a series of mixed effects logistic regression analyses that I conducted on the sandhi variables separately. The goal of these analyses was to determine which linguistic and social factors contribute to the likelihood of the traditional Tianjin sandhi patterns for (53-53), (53-21) and (21-21) being applied in word list style. This in turn will contribute to the discussion, in Chapter 7, of what is motivating these generational changes to Tianjin tone sandhi.

5.2 Regression model construction

Mixed effects logistic regression analyses were conducted in R (R Core Team, 2018) using the lme4 package (Bates, Macchler, Bolker, & Walker, 2015) to check whether social factors have significant effects on (53-21), (53-53) and (21-21) respectively in wordlist style. The input data comprised all 1799 tokens of (53-21), 1800 tokens of (53-53) and 1621 tokens of (21-21) produced by all of the 90 speakers in the word list style. For the analysis of (53-21) and (53-53), the dependent variable is binary: application versus non-application of the (53-21) and (53-53) sandhi rule, with the variant of interest being its application. For the analysis of (21-21), the dependent variable is also binary: new application variant HH versus the old application variant LH plus non-application variant LL, with the variant of interest being the new variant HH.

Five fixed independent variables were included in the analyses: AGE GROUP (Old, Middle + Young) or (Old, Middle, Young), SOCIAL CLASS (working class, middle class), GENDER (male, female), ATTITUDE TO MIGRANTS (positive, negative), ATTITUDE TO TIANJIN (positive, negative). AS shown in Table 5.1 and Figure 5.1, the Middle-aged group is very close to Young people in (53-53) and (53-21), so these two groups are merged in the analyses of these two variables. But for (21-21), the difference between the Middle-aged group and either the Old group or the Young group is too big to merge it with either of them. So the age group for (21-21) is a three-group predictor, and I use dummy or default comparisons in which Young group and Old group are compared to the Middle-aged group respectively.

Reference levels were set at Young (+Middle) group for (53-53) and (53-21), Middle-aged group for (21-21), Working social class, Male gender, positive Attitude to Migrants, and positive Attitude to Tianjin. Two random effects were included: PARTICIPANT, and the output tone of the first word in each token, henceforth W1⁸. P-values are obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question.

The specific fixed independent variables entered into a given regression analysis varies depending on the initial checking of the data and relevant previous studies.

The fixed independent variables for the regression analysis of (53-53) are AGE GROUP (Young+Middle, Old), SOCIAL CLASS, GENDER, and ATTITUDE TO TIANJIN. Previous studies (Gao

⁸ I have not found any word to be more variable than others in the wordlists in (53-53) and (53-21), but found *xinche* ‘new vehicle’ or *xince* (a variation of *xinche* in Tianjin) seems to be an outlier in (21-21) wordlist, which takes up 10 out of the 23 tokens in the word list dataset in which (21-21) sandhi rule does not apply. The difference is not really caused by differences in words or syllables, but by whether *che* or *ce* ‘vehicle’ is realized as T1 or neutral tone. *Che* or *ce* can be freely pronounced either as tone 1 or neutral tone in the second syllable of a disyllabic sequence. When the second syllable *che* or *ce* is pronounced as a neutral tone, the tone sandhi does not apply. Although (21-21) had some skews because of neutral tone, mixed model regression can handle this imbalance through the inclusion of lexical word as a random effect, which was also a reason why I subsequently ran this kind of analysis. The random effects of W1 support my observations.

& Lu, 2003; X. Wang, 2016) have shown that the change in (53-53) may be related to age group, social class, and gender. As for attitudes to Tianjin, Feng (2007/2015) reports that positive attitude to Tianjin cannot preserve (53-53) because it only has effects on more salient features like ‘special consonants’. One proposal of the current dissertation is to check the effects of language attitude on salient features like (53-53), so I include it in the regression. But attitude to migrants is not included in the regression model for (53-53) for three reasons. First, it is not mentioned in any previous studies on (53-53). Second, the decrease in (53-53) seems to have occurred earlier than large-scale in-migration to Tianjin. Third, the initial checking of the data shows that the application frequency difference between speakers with positive attitude to migrants and negative attitude is smaller than any of the other four fixed predictors.

The fixed independent variables for the regression analysis of (53-21) are AGE GROUP (Young+Middle, Old), SOCIAL CLASS, GENDER and ATTITUDE TO MIGRANTS. After initial checking of the data, I found the application frequency difference by attitude to Tianjin is only 1.4%, smaller than the difference by any other independent variables.

The fixed independent variables for the regression analysis of (21-21) are AGE GROUP (Old, Middle, Young), SOCIAL CLASS, GENDER, ATTITUDE TO MIGRANTS, and ATTITUDE TO TIANJIN. Initial checking of the data shows that the variance by gender is smaller than the variance by any other predictor (10.3%), but Lu (1997) reported that young female seemed to lead the change towards the new local variant HH, so I include gender as an independent variable in the regression of (21-21). People with negative attitude to migrants apply the new local variant at 76.9%, 16.3% higher than people with positive attitude to migrants (60.6%). The variance is quite big, so I keep attitude to migrants in the regression.

Model reduction method or backward stepwise selection is used to select the best model. The practice of model reduction starts with a full model which includes all candidate predictors and possible two-way interactions in the model, and then systematically removing the term with the highest p-value one-by-one until only significant predictors are left. The significant p-value in the current study is <0.5 .

Variance inflation factor (hereafter VIF) and tolerance statistics (hereafter $1/VIF$) were run to test whether there is a strong correlation between two or more predictors on all models reported, and that no collinearity was found for any of them.

5.3 The word list results of (53-53)

A mixed model regression indicates that there are significant main effects of age group, $p<0.001$, and social class, $p<0.05$, on the application of (53-53) in word list style. However, gender, and attitudes to Tianjin dialect, do not have significant main effects. No significant two-way interaction effect between any predictor of (53-53) sandhi rule was found for word list style, and therefore I will only report main effects. The parameter estimates for the generalized mixed model fit by maximum likelihood are given in Table 5.2.

The variances for random effects, PARTICIPANT and W1, are 6.743 and 1.167 respectively, not close to 0, which means it is reasonable to treat them as random effects and mixed effect model is good fit for my data.

Table 5.2 Fixed effects of generalized mixed model of (53-53) in word list

Fixed effects	Estimate	SE	Z	P
(intercept)	-2.883	0.799	-3.609	<0.001
Age group (Young+Middle)	-4.103	0.802	-5.118	<0.001
Social class (Working)	2.189	0.728	3.005	<0.05

To be more specific, the Young group (including Middle-aged group) are 98.3% ($1 - \exp(-4.103) = 0.983$) less likely to apply the (53-53) variable than the Old group. This confirms the overall trend shown in Table 5.1 and Figure 5.1 for (53-53), where there was a positive correlation between age and (53-53) application. Additionally, working class speakers are almost 9 times ($\exp(2.189) = 8.926$) more likely to use the (53-53) sandhi rule than middle class speakers, which is in line with the previous studies (Gao & Lu, 2003; X. Wang, 2016).

In the next four subsections, I first explore the two significant effects in more detail, starting with the main effect of age group. I then provide some descriptive statistics and brief discussion for the non-significant effects.

The VIF and tolerance statistics were run to test whether there was a strong correlation between the two predictors in the regression model. VIF results of both Age group and social class are 1.004, and the 1/VIF of these two groups are 0.995.

For VIF, Myers (1990) suggests that 10 is the value at which to worry. For 1/VIF, Menard (1995) suggests that values below 0.2 need to worry. The VIF values are well below 10, and the tolerance statistics all well above 0.2. Based on these measures, I conclude that there is no collinearity between the predictors.

5.3.1 The effect of age group

As we saw, the output of mixed regression analysis indicates that age group has a significant overall effect on the application of (53-53) word list, $p < 0.001$. This sandhi rule has two variants: application variant LL and non-application variant HL. The rates of application by age group is given in Table 5.3, and displayed in the boxplot in Figure 5.2. Although Young and Middle-aged groups are merged in regression, here they are provided separately to give more details of the distribution of (53-53) by age group. The same method of display is used for the other fixed predictor variables as well.

Table 5.3 The application frequency of LL variant in word list by age group

Age groups	LL variant in word list		SD
	App / Tokens	%	
Young	28/839	3.3	1.6
Middle	39/500	7.8	3.85
Old	140/460	30.4	4.82
Total	207/1799	11.5	4.01

Figure 5.2 Application frequency of LL variant in word list by age group

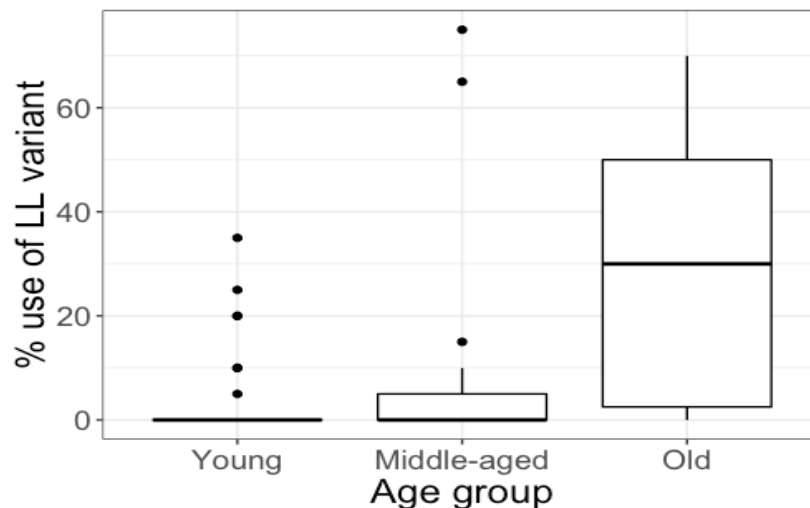


Table 5.3 and Figure 5.2 indicate the (53-53) sandhi rule has almost disappeared in the Young and Middle-aged groups, but it still exists among old speakers. Its application rate is 0 for most young speakers, and the SD of the Young group is 1.6, indicating the loss of the traditional LL variant in this age group seems to be complete. Although the application rate of (53-53) in the Middle-aged group is 7.8%, a little higher than the 3.3% application rate in the Young group, as shown in Figure 5.2, the median of the application of (53-53) in the Middle-aged group is almost 0, which is very close to the median in the Young group. Besides, mixed regression indicates that the difference between the Young group and Middle-aged group is not significant, suggesting this sandhi rule has become virtually obsolete in the Middle-aged group as well.

The application rate of the (53-53) rule in the Old group is 30.4%, much higher than those of the other two age groups. The median of the Old group shown in Figure 5.2 is also much higher than those of the other two age groups. Besides, the bigger SD of the Old group shown in Table 5.3 and the bigger box of the Old group in Figure 5.2 suggest bigger in-group variation among old speakers. It seems that some members of this age cohort are more advanced with regard to the loss of (53-53) than others.

As shown in Figure 5.2, there are several outliers in both Young and Middle-aged groups. Their positive attitude to Tianjin dialect might be a reason that makes their application rates much higher than other speakers of their age groups. I discuss attitudes, and these outlier individuals, in more detail in subsection 5.3.4.

5.3.2 The effect of social class

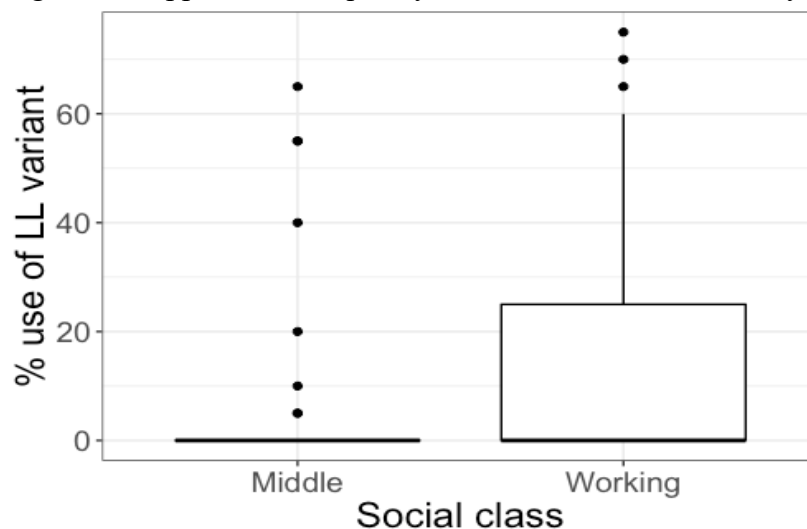
Mixed logistic regression indicated that social class plays a significant role in the application of (53-53) in word list reading, $p < 0.05$. As shown Table 5.4, overall the working class

have a higher application frequency of this sandhi rule than the middle class have, and the tendency is consistent in every age group.

Table 5.4 The application frequency of LL variant by age and social class

	Middle class		Working class		Total	
	App/Tokens	%	App/Tokens	%	App/Tokens	%
Young	8/440	1.8	20/399	5	28/839	3.3
Middle	2/200	1	37/300	12.3	39/500	7.8
Old	44/200	22	96/260	36.9	140/460	30.4
Total	54/840	6.4	153/959	16	207/1799	11.5

Figure 5.3 Application frequency of LL variant in word list by social class



As shown in Table 5.4 and Figure 5.3, the application rate of (53-53) in the middle class is very low, which suggests that the rule practically does not exist in this social class, despite the existence of a handful of outlier speakers. Table 5.4 and Figure 5.3 show that in contrast, working class speakers retain the (53-53) rule. Thus the change towards the standard and prestigious non-application HH variant seems to be led by middle class speakers.

The result that middle class have higher application frequency of the standard and prestigious variant is consistent with previous studies on Tianjin (Gao & Lu, 2003), and also consistent with the general tendency between social class and language change from above observed in previous sociolinguistic studies (Labov, 2000).

5.3.3 The effect of gender

The role of gender in language change and variation has been extensively studied in sociolinguistics (e.g. Trudgill, 1972; Labov, 1990 & 2001; Fought, 1999; Carvalho, 2004; Zhang, 2005; among others). In a survey of many studies, Labov (1990) observed that women tend to use the standard variant more than men do for diachronically stable variables and women tend to lead language change no matter whether it is a change from above or a change from below. Labov (1990) also emphasizes the importance of the interaction between social class and gender. Besides, a previous sociolinguistic study on Tianjin (Gao & Lu, 2003) indicates that men are more likely to retain the local variant LL than women. For these reasons, I provide some descriptive statistics on the application of the traditional LL variant, even though the mixed effect regression analysis did not indicate that gender is a significant predictor of the application of (53-53) sandhi rule. Table 5.5 shows the application frequency of the LL variant by gender and age group.

Table 5.5 The application frequency of LL variant in word list by gender and age group

	Female		Male		Total	
	App / Tokens	%	App / Tokens	%	App / Tokens	
Young	7/500	1.4	21/339	6.2	29/839	3.3
Middle	22/260	8.5	17/240	7.1	39/500	7.8
Old	44/200	22	96/260	36.9	140/460	30.4
Total	73/960	7.6	134/839	16	207/1799	11.5

As shown in Table 5.5, overall men have a higher application frequency of the local variant LL than women do, and the tendency is confirmed in the Young and Old groups. The only exception is the Middle-aged group, in which women have a slightly higher application rate than men do. The difference in the Middle-aged group might be caused by the outlier speaker #79, who produced 15 of all the 22 applications of the sandhi rule by middle-aged females. Without the data from #79, the application frequency of LL variant among female middle-aged speakers is only 2.9%, much lower than the application rate of middle-aged males.

Table 5.6 is the application frequency of the LL variant by gender and social class. Looking first at the row for female speakers, we can see that the overall application rate of the local variant is 7.6% in the female group. Looking next at the row for male speakers, we can see that males are overall more likely to use the local variant LL than females (16%), and this tendency is consistent in both middle class and working class groups. In both the female group and the male group, middle class speakers are less likely to use the local variant LL than working class speakers.

Among all subgroups by social class and gender, working class males have the highest application frequency of the local variant, while middle class females have the lowest application frequency, indicating that working class males better preserve the local feature and the levelling of this local feature might be led by middle class women.

Table 5.6 The distribution of LL variant in word list by social class and gender

	Middle class		Working class		Total	
	App/Tokens	%	App/Tokens	%	App/Tokens	%
Female	20/500	4	53/460	11.5	73/960	7.6
Male	34/340	10	100/499	20	134/839	16
Total	54/840	6.4	153/959	16	207/1799	11.5

5.3.4 The effect of attitudes to Tianjin dialect

One prediction of the present dissertation is that the drop in the application of (53-53) variable over apparent time is related to changing attitudes to Tianjin dialect. Since the application variant LL is a salient feature of Tianjin dialect, we would expect that people with a positive attitude to Tianjin dialect should have a higher application frequency of the local non-prestigious variant LL than people with a negative attitude to Tianjin dialect. For this reason, the application frequency of (53-53) by attitudes to Tianjin dialect is reported here, even though the mixed regression analysis did not indicate that attitudes to Tianjin dialect is a significant predictor of the application of (53-53) sandhi rule.

Cross-tabulations for attitude to Tianjin dialect and the two significant factors influencing the application of (53-53) sandhi rule, age group and social class, are provided in Table 5.7 and Table 5.8 respectively to check whether attitudes to Tianjin are changing over apparent time and/or varying with social class.

Table 5.7 The application frequency of LL variant in word list by attitude to Tianjin and age group

	Negative attitude		Positive attitude		Total	
	App/Tokens	%	App/Tokens	%	App/Tokens	%
Young	0/300	0	28/539 (27)	5.2	28/839	3.3
Middle	7/220	3.2	32/280 (14)	11.4	39/500	7.8
Old	24/100	24	116/360 (18)	32.2	140/460	30.4
Total	31/620	5	176/1179 (59)	14.9	207/1799	11.5

As shown in Table 5.7, overall people with positive attitude to Tianjin dialect have higher application frequency of the local variant LL (14.9%) than people with negative attitude (5%), and the tendency is consistent in every age group. Speakers with a negative attitude to Tianjin dialect

in the Young group do not even have any local variant at all (0%). All the local variant tokens in the Young group are from speakers with a positive attitude to Tianjin dialect (5.2%). The differences between speakers with positive attitude and negative attitude in the other two age groups are even bigger, suggesting the change towards the standard variant is led, as predicted, by speakers with negative attitudes to Tianjin dialect.

Recall that in subsection 5.3.1, Figure 5.2 included several outliers in both Young and Middle-aged groups and that I proposed that a positive attitude to Tianjin dialect might be the reason that makes their application rates much higher than other speakers of their age groups. In the Young group, the 28 application tokens are from speakers #5 (5 tokens), #9 (4 tokens), #13 (7 tokens), #16 (2 tokens), #29 (1 tokens), #31 (3 tokens), #35 (2 tokens), and #37 (4 tokens). These speakers include males and females from both middle class and working class. #5, #13, #31, #35, #37 are males, while #9, #16, #29 are females. #9, #29, #31 are middle class speakers, while #5, #13, #16, #35 and #37 are working class speakers. Their common point is they all have positive attitude to Tianjin dialect. In the Middle-aged group, #74 contributes 15 local tokens. She has a strong positive attitude to Tianjin dialect. The following words are from her.

Wo bu juede Tianjinhua buru Putonghua.

‘I don’t think Tianjin dialect is inferior to Putonghua’.

Tianjin ren jiu gai shuo Tianjinhua, zhe shi ni de yuyan.

‘Tianjin people should speak Tianjin dialect because it is your language’

Tianjin ren shuo Putonghua Zhuang.

‘Tianjin natives are snobbish when they speak Putonghua’.

(Quotes from #79, a 60 year old working class female)

These seem to support that the outliers' linguistic behavior with respect to (53-53) is related to their positive attitude to Tianjin dialect.

After studying attitudes to Tianjin dialect over apparent time, I now report on the distribution of (53-53) by attitudes and social class. As shown in subsection 5.3.2, the middle class leads the change in the standardization process of (53-53); here I want to see whether the social class distribution of (53-53) varies with attitude to Tianjin dialect (Table 5.8).

Table 5.8 The application frequency of LL variant in word list by social class and attitude to Tianjin

	Negative attitude		Positive attitude		Total	
	App/Tokens	%	App/Tokens	%	App/Tokens	%
Middle	22/460	4.8	32/380	8.4	54/840	6.4
Working	9/160	5.6	144/800	18	153/959	16
Total	31/620	5	176/1179	14.9	207/1799	11.5

Table 5.8 above shows that people with positive attitudes to Tianjin are more likely to use the local variant in every social class. Working class speakers with positive attitudes towards Tianjin have the highest application rate among the subgroups by social class and attitudes to Tianjin, which suggests that this group of people best keeps the local variant. Middle class speakers with negative attitudes have the lowest application among the subgroups by social class and attitudes to Tianjin, which suggests that this group of people are leading the change towards the non-local variant.

The percentages of people with positive attitude to Tianjin and negative attitude to Tianjin are different in middle class and working class. The working class tends to have positive attitudes towards Tianjin dialect while the middle class tends to have negative attitudes. Thirty-nine out of forty-eight working class speakers have positive attitudes to Tianjin, which takes up 81.3% of all

working class speakers. Twenty-three out of 42 middle class speakers have negative attitudes towards Tianjin dialect, but only nine out of 48 working class speakers have negative attitudes towards Tianjin dialect.

Although gender does not have a significant effect on the application of (53-53), a previous study has indicated the influence of gender on the application of (53-53) variable (Gao & Lu 2003). As a result, the cross-tabulation for attitude to Tianjin dialect and gender is provided in Table 5.9.

Table 5.9 The distribution of LL variant in word list by gender and attitude to Tianjin

	Negative attitude		Positive attitude		Total	
	App / Tokens	%	App / Tokens	%	App / Tokens	%
Female	17/360	4.7	56/600	9.3	73/960	7.6
Male	14/260	5.4	120/579	20.7	134/839	16
Total	31/620	5	176/1179	14.9	207/1799	11.5

As shown in Table 5.9, male speakers with positive attitudes have the highest application frequency of the local variant (20.7%), while female speakers with negative attitudes have the lowest application rates of the local variant (4.7%).

Although mixed regression analysis does not indicate that attitude to Tianjin is a significant predictor of the application of (53-53), people with positive attitude to Tianjin dialect tend to use it more than people with negative attitude to Tianjin dialect, and the tendency is consistent in every age group, social class and gender.

Overall, the distribution of (53-53) application by gender, social class, attitudes to Tianjin and their interactions is consistent with the pattern we would expect to see for a change from above (Labov, 1990). I discuss this further in Chapter 7. First, however, I turn in the next section to the (53-21) variable, for which the sociolinguistic patterns are very different.

5.4 The word list results of (53-21)

The fixed effects of generalized mixed model fit by maximum likelihood are given in Table 5.10 for (53-21). Mixed model regression analysis indicates that age group, $p < 0.001$, and attitude to migrants, $p < 0.05$, play a significant role in the application of (53-21) in wordlist style. However, social class $p > 0.05$, and gender $P > 0.05$, in general, do not have significant main effects. No significant two-way interaction effect between any predictor is observed in (53-21) word list.

The variances for random effects, PARTICIPANT and W1, are 2.2516 and 0.6426 respectively, not close to 0, which means it is reasonable to treat them as random effects and mixed effect model is good fit for my data.

Table 5.10 Fixed effects of generalized mixed model of (53-21) word list

Fixed effects	Estimate	SE	Z	P
(Intercept)	2.08	0.46	4.54	<0.001
Age group (Young)	1.94	0.44	4.4	<0.001
Attitude to Migrants (Positive)	-1.03	0.41	-2.54	<0.05

To be more specific, the Young group (including Middle-aged group) are almost 7 (exp (1.94) = 6.95) times more likely to apply the (53-21) variable than the Old group. This confirms the overall trend shown in Table 5.1 and Figure 5.1 for (53-21), where there was a negative correlation between age and (53-21) application. Additionally, people with positive attitude to migrants are 64.3% ($1 - \exp(-1.03) = 0.643$) less likely to use the HH local variant than people with negative attitudes to migrants, which is in line with the prediction that the rise in (53-21) is related to a local resistance to migration.

The VIF and tolerance statistics were run to test whether there was a strong correlation between the two predictors in the regression model. VIF results of both Age group and Attitude to

migrants are 1.005, and the 1/VIF of these two groups are 0.995. The VIF values are well below 10, and the tolerance statistics all well above 0.2. Based on these measures, I conclude that there is no collinearity between the predictors.

As I did for (53-53), I discuss the two significant effects, age group and attitude to migrants, in subsections 5.4.1 and 5.4.2. I do not, however, discuss the non-significant effects of attitude to Tianjin, social class or gender because none of them have ever been reported to have relation with the application of (53-21) or proposed to be a cause of the rise in the application frequency of (53-21).

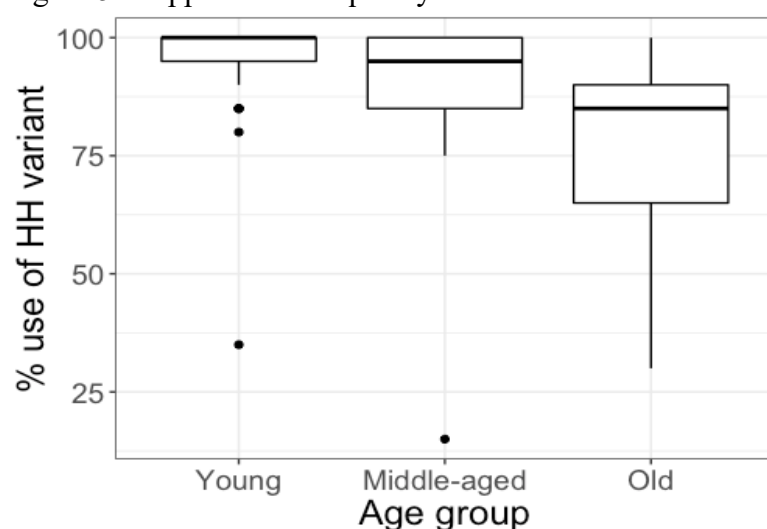
5.4.1 The effect of age group

The output of mixed regression analysis indicated that age group has a significant overall effect on the application of (53-21) word list, $p < 0.001$. The rates by age group are given in Table 5.11 and presented visually as boxplots in Figure 5.4. Recall that the traditional Tianjin output of the (53-21) variable is a HH tone sequence.

Table 5.11 The application frequency of the HH variant by age group

Age group	FL word list		
	App/Tokens	%	SD
Young	801/840	95	2.2
Middle-aged	372/500	89.4	3.47
Old	340/460	73.9	4.08
Total	1586/1800	88.1	3.46

Figure 5.4 Application frequency of HH variant in word list by age group



As shown in Table 5.11 and Figure 5.4, the application frequency of the (53-21) sandhi rule exhibits a clear negative correlation with age group. Table 5.11 and Figure 5.4 also show that the level of within-group variation for (53-21) is very low for the Young, fairly low for the Middle, but comparatively high in the Old, which is similar to the levels of within-age-group variation that we saw for the (53-53) variable in Table 5.1. The largest box of the Old group in Figure 5.4 indicates that this age group has the biggest variability, which is supported by the biggest SD among the three age groups in Table 5.11. The Young and Middle-aged groups have smaller in-group variations, indicated by their smaller SD in Table 5.11 and the smaller, more compressed boxes in Figure 5.4. In fact, the traditional (53-21) pattern is used near categorically in the Young group (SD=2.2.) and Middle-aged group (SD=3.47), which indicates that the change in these two groups toward renewed high frequency use of the traditional (53-21) sandhi rule is almost complete.

5.4.2 The effect of attitudes to migrants

According to the results of mixed regression analysis reported in Table 5.10, attitudes to migrants have a significant effect on the application frequency of (53-21) in word list style, $p < 0.05$. Table 5.12 reports the application rates of (53-21) in word list for speakers with positive or

negative views of migrants. Figure 5.5 displays the distribution of (53-21) application by attitudes to migrants visually.

Table 5.12 The application means of the HH variant by attitude to migrants

	FL Interview		
	App / Tokens	%	SD
Negative attitudes	894/959	93.2	1.94
Positive attitudes	701/841	83.4	4.38
Total	1595/1800	88.6	3.46

Figure 5.5 Application frequency of HH variant in word list by attitude to migrants

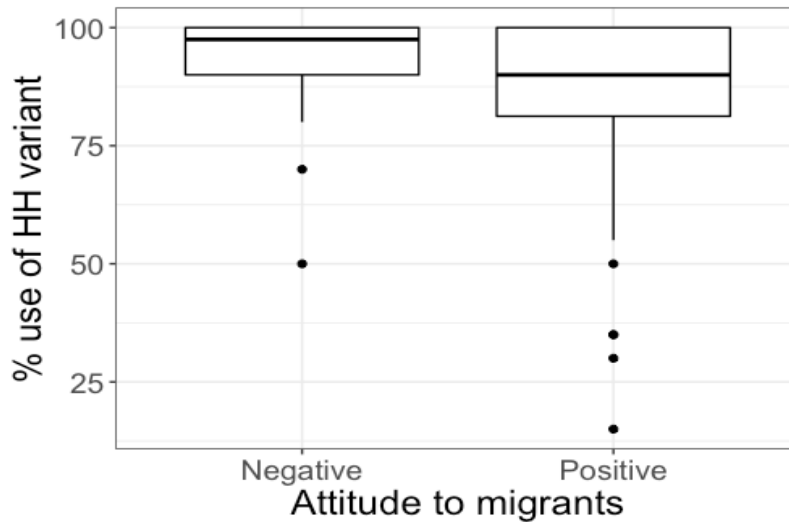


Table 5.12 and Figure 5.5 show that people with negative attitudes to migrants tend to have higher application frequency of the local variant HH, which is consistent with the hypothesis of the present dissertation. The distribution of the HH variant in each age group by attitudes to migrants is given in Table 5.13.

Table 5.13 Distribution of HH variant in word list by attitude to migrants and age group

	Negative		Positive		Total	
	App / Tokens	%	App / Tokens	%	App/Tokens	%
Young	562/580	96.9	236/260	90.8	798/840	95
Middle	165/180	91.7	282/320	88.1	448/500	89.4
Old	167/199	83.9	183/261	70.1	350/460	76.1
Total	894/959	93.2	701/841	83.4	1595/1800	88.6

Table 5.13 indicates that people with negative attitude to migrants are more likely to use the local variant HH than people with positive attitude to migrants in every age group, although this tendency is most pronounced among the Old speakers. This may be because the changes in the Young and Middle-aged groups are near complete, while the change in the Old group is ongoing.

5.5 The word list results of (21-21)

Mixed model logistic regression for the (21-21) variable was carried out in R (R Core Team 2018) using the packages, procedures and predictors introduced in §5.2. The fixed effects for the generalized mixed model fit by maximum likelihood are given in Table 5.14.

There are significant main effects of age group, $p < 0.001$, and attitude to Tianjin, $p < 0.001$, on the application of (5221) wordlist. However, social class, $p > 0.05$, gender, $p > 0.05$, and attitude to migrants, $p > 0.05$, do not have significant main effects. No significant two-way interaction effects have been found in the application of (21-21) word list.

The variances for random effects, PARTICIPANT and W1, are 3.2638 and 0.2501 respectively, not close to 0, which means it is reasonable to treat them as random effects and mixed effect model is good fit for my data.

Table 5.14 The fixed effects of generalized mixed model of (21-21) in word list

Fixed effects	Estimate	SE	Z	P
(intercept)	2.215	0.525	4.22	<0.001
Age group (Old)	-2.93	0.605	-4.841	<0.001
Age group (Young)	3.721	0.6	6.199	<0.001
Attitude to Tianjin (Positive)	-1.753	0.528	-3.322	<0.001

To be more specific, the Old group are 94.7% ($1 - \exp(-2.93) = 0.947$) less likely to use the new local HH variant, than the reference group, Middle-aged group, while the Young group are about 41 ($\exp(3.721) = 41.306$) times more likely than the reference group to apply the HH variant. This confirms the overall trend shown in Figure 5.1 for (21-21), where there is a negative correlation between age and the application of the new local variant HH. Additionally, people with positive attitude to Tianjin dialect are 82.7% ($1 - \exp(-1.753) = 0.827$) less likely to use the new local variant than people with negative attitude to Tianjin dialect.

The VIF and tolerance statistics were run to test whether there was a strong correlation between the two predictors in the regression model. VIF results of both Age group and Social class are 1.033, and the $1/\text{VIF}$ of these two groups are 0.968. The VIF values are well below 10, and the tolerance statistics all well above 0.2. Based on these measures, I conclude that there is no collinearity between the predictors.

Although the remaining predictors were not significant, male speakers are less likely to use the new local variant than female speakers, working class speakers are less likely to use the new local variant than middle class speakers, and people with positive attitudes to migrants are less likely to use the new local variant than people with negative attitudes to migrants. Next I introduce the effects of the two significant predictors of (21-21) variable.

5.5.1 The effect of age group

The output of mixed regression analysis indicated that age group has a significant overall effect on the application of (21-21) word list, $p < 0.001$. Different from the other variables of the present study, (21-21) has two application variants (HH and LH). Young speakers tend to use the HH variant, while old speakers tend to use the LH variant, as shown in Table 5.1 and Figure 5.1. Since the LH variant is reported in comparatively early studies on Tianjin tone sandhi, and is used mostly by old and middle-aged speakers, I take LH variant as the old variant and assume that the change in (21-21) is from LH to HH.

When I calculate the application frequency of (21-21), the frequency of HH variant is the HH variant vs the total number of tokens (HH variant + LH variant + non-application variant), while the LH frequency is the LH variant vs the total number of tokens (HH variant + LH variant + non-application variant).

Tables 5.15-5.16 are the application frequencies of HH variant and LH variant respectively, Figures 5.6-5.7 are the boxplots of the application frequencies of HH variant and LH variant respectively.

Table 5.15 The application frequency of new HH variant in word list by age group

Age group	LL wordlist		
	App/Tokens	%	SD
Young	733/758	96.7	1.54
Middle	298/450	66.2	6.06
Old	92/414	22.2	4.19
Total	1123/1621	69.3	3.63

Table 5.16 The application frequency of old LH variant in word list by age group

Age group	LL word list		
	App/Tokens	%	SD
Young	18/758	2.4	0.96
Middle	145/450	32.2	5.9
Old	313/414	75.6	4.14
Total	475/1621	29.3	3.55

Figure 5.6 Application frequency of HH variant in word list by age group

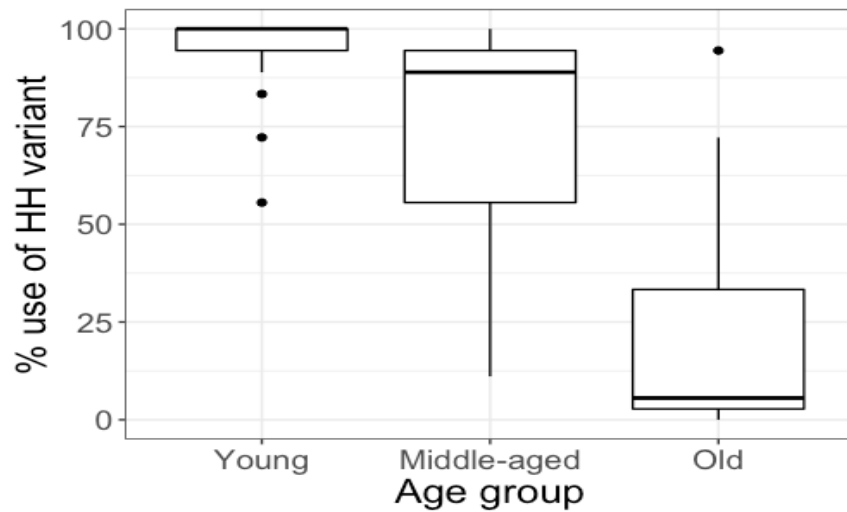
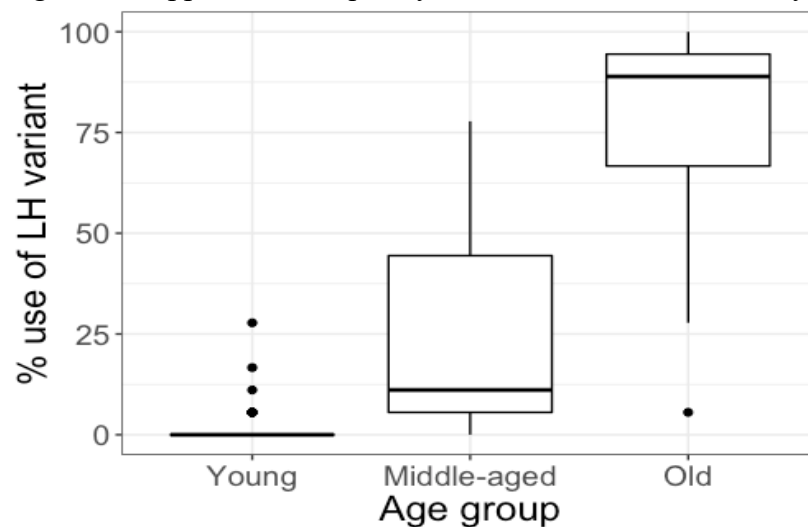


Figure 5.7 Application frequency of LH variant in word list by age group



The application of (21-21) tone sandhi changes clearly with age, as shown in Tables 5.15-5.16 and Figures 5.6-5.7. Young people use the HH variant almost categorically, while old speakers tend to use the LH variant. The Standard deviation of the Young group is 1.54 for the HH variant and 0.96 for the LH variant, which indicates that the variation within this age group is small and the change has completed in this age group. The Standard deviation of the Middle-aged group is 6.06 for the HH variant and 5.9 for the LH variant. The Standard deviation of the Old group is 4.19 for the HH variant and 4.14 for the LH variant. The bigger Standard deviations in the Middle-aged and Old groups indicate that the distribution of (21-21) outputs are spread out over a wide range of values and the change in this two groups is not complete.

The new application variant HH was first noticed by Lu (1997), who reported that HH seemed to be preferred by young female speakers in their twenties. Those young speakers in the 1990s would be in the Middle-aged group in the present dissertation, and these speakers might be the first generation who preferred the new HH variant to the old LH.

5.5.2 The effect of attitudes to Tianjin dialect

Mixed model regression results indicate that attitude to Tianjin dialect is a significant predictor of the application of the (21-21) sandhi rule in word list style, $p < 0.001$. People with negative attitude to Tianjin dialect have a higher application frequency of the new local application variant HH than people with positive attitude to Tianjin dialect. People with negative attitude to Tianjin dialect have a lower application frequency of the old local application variant LH than people with positive attitude to Tianjin dialect.

Table 5.17 The application frequency of HH variant in word list by attitude to Tianjin and age group

Age group	Negative attitude		Positive attitude		Total	
	App / Tokens	%	App / Tokens	%	App / Tokens	%
Young	285/288	99	448/469	95.5	733/758	96.7
Middle	136/180	75.6	162/270	60	298/450	66.2
Old	43/90	47.8	49/324	15.1	92/414	22.2
Total	437/558	83.2	679/1063	62	1123/1621	69.3

Table 5.18 The application frequency of LH variant in word list by attitude to Tianjin and age group

Age group	Negative attitude		Positive attitude		Total	
	App/Tokens	%	App/Tokens	%	App/Tokens	%
Young	1/288	0.3	16/469	3.4	18/758	2.4
Middle	41/180	22.8	104/270	38.5	145/450	32.2
Old	45/90	50	268/324	82.7	313/414	75.6
Total	87/558	15.6	388/1063	36.5	475/1621	29.3

Figure 5.8 Application frequency of HH variant in word list by attitude to Tianjin

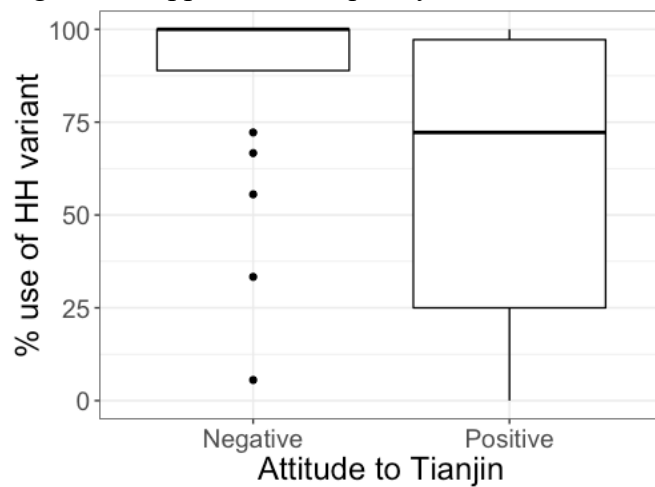
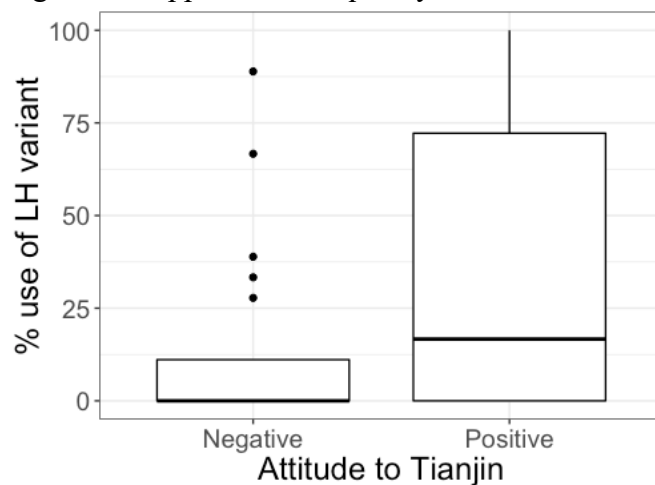


Figure 5.9 Application frequency of LH variant in word list by attitude to Tianjin



As shown in Tables 5.17-5.18 and Figures 5.8-5.9, people having negative attitude to Tianjin dialect tend to have higher HH variant than people having positive attitudes to Tianjin dialect, while people having positive attitude to Tianjin dialect are more likely to use the old local application variant LH. These tendencies are consistent in every age group.

5.6 Summary

The word list results confirm that the three variables, (53-53), (53-21) and (21-21), have undergone changes over time. Age group is the only predictor that has significant effects on the applications of all three variables. Social class is significant in the application of (53-53). The

change in (53-53) seems to be a change towards the standard prestigious variant HH led by young middle class speakers. Attitude to migrants is the other significant predictor influencing the application frequency of (53-21), in addition to age group. The arrival of large number of migrants may have made Tianjin natives feel that their local identity and culture are in danger of disappearance. As a result, they recycle some unmarked old feature of Tianjin dialect to keep their local identity and culture. The word list results –that people with negative attitude to migrants are more likely to use the local application variant HH—seem to support this view. (21-21) sandhi rule has two application local variants: HH and LH. In addition to age group, attitude to Tianjin dialect is a significant predictor. People with a positive attitude to Tianjin dialect are more likely to keep the old local variant, so the change towards the new local variant seems to be led by people with negative attitude to Tianjin dialect.

Since word list reading style is a very careful style and people pay much attention to their language, the social class difference might be minimized in this style, but the difference might be shown in the less careful interview style, which will be presented in the next chapter (Chapter 6). Besides, by comparing the more careful word list results with less careful interview results, we could find out style-shifting which itself is revealing of speakers' attitudes towards the local dialect (Chapter 7).

Chapter 6 Interview Results

This chapter reports on the distribution of (53-53), (53-21) and (21-21) produced by the participants in interview style, which is a near natural style in which speakers pay less attention to their speech than in word list reading style, and therefore interview style is closer to speakers' natural speech than word list reading style. Interview style can also reveal inter-speaker variation, such as social class differences, more systematically and more clearly than word list reading style can (Trudgill, 1972).

As noted in Chapter 5, although age was the only significant predictor for the three variables in word list style, other predictors might have significant effects on them in interview style. For example, social class is only significant in (53-53) in word list style, but does it have significant effects on other variables in interview? Gender is not significant in any variable in word list style, but it is unclear whether it has significant effects in less careful interview style. Speakers having positive attitudes to Tianjin consistently have higher application frequency of the local variant of (53-53) than speakers with negative attitudes, but the difference is not significant in word list reading style, so it is worthwhile to check whether the difference is significant in interview style. In short, it is interesting to check whether the predictors have significant effects in a less careful style when speakers pay less attention to their speech. This chapter first provides a general introduction to interview results in § 6.1, then in § 6.2 the regression model is introduced, followed by a detailed report of the three variables from § 6.3 § 6.5.

6.1 A general introduction to interview results

The overall rates of use of traditional Tianjin sandhi patterns for (53-53), (53-21) and (21-21) in interview style are given in the bottom two rows of Table 6.1 below. Although tone sandhi rules are applied most frequently for (21-21), at 92.3%, the traditional local LH output variant

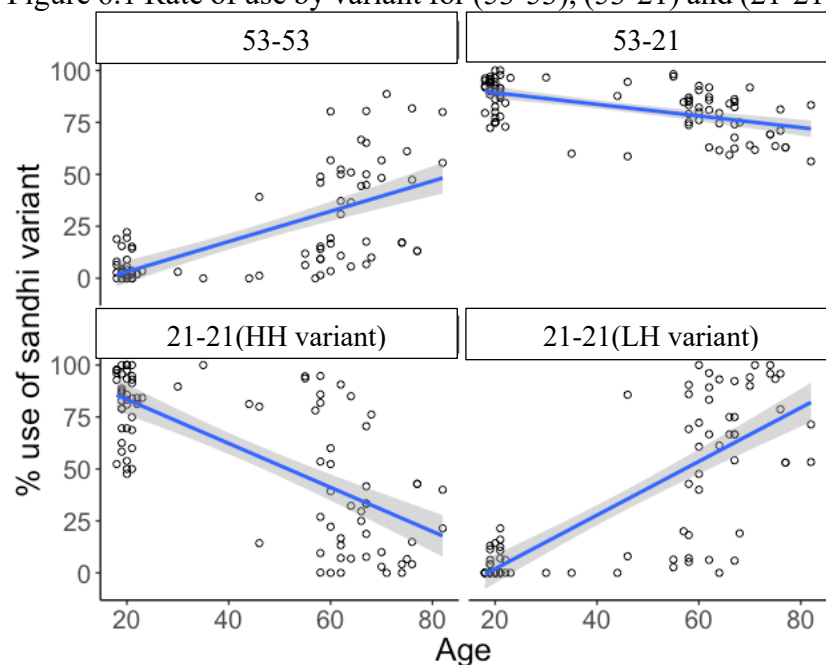
accounts for only 28.7% of rule applications. The local Tianjin output pattern for (53-21) is next, at 84.7%, followed far behind by the traditional variant of (53-53) at only 22.2% respectively. Overall, the ranking order is the same as the ordering in word list style that we saw in Chapter 5. When reading the word list, speakers applied tone sandhi rules most frequently for (21-21), at 98.6%, with the traditional local LH output variant accounting for only 29.3% of rule applications. The next is the local Tianjin variant for (53-21), at 88.1%, much higher than the traditional Tianjin variant of (53-53) at only 11.5%. Although the overall application ranking order of the three variables is the same across both styles, the application rate for (53-53) is about 11% higher in interview style than in word list style, suggesting speakers avoid the local application variant in more careful word list style.

As was the case for word list style, a mixed regression analysis using R, which will be reported in detail in the following subsections in this chapter, indicated that age is the most important social factor influencing the application of the three variables in interview style. For this reason I present an overview of the distribution of variation for the interview results by age group in both Table 6.1 and Figure 6.1.

Table 6.1 Rates of sandhi rule applications of (53-53), (53-21) and (21-21) in interview style, by age group, all speakers (N=90)

Age group	FF		FL		LL (HH variant)		LL (LH variant)	
	App/N	%	App/N	%	App/N	%	App/N	%
Young	94/2260	4.2	1751/1943	90.1	1026/1199	85.6	31/1199	2.6
Middle	488/1861	26.2	930/1111	83.7	420/740	56.8	290/740	39.2
Old	673/1541	43.7	607/829	73.2	143/558	25.6	395/558	70.8
Total	1255/ 5662	22.2	3288/3883	84.7	1589/2497	63.6	716/2497	28.7
					LH+HH			
					App/N	2305/2497	92.3%	

Figure 6.1 Rate of use by variant for (53-53), (53-21) and (21-21) in interview style by age (N=90)



As shown in Table 6.1 and Figure 6.1, the applications of the three variables show a correlation with age. The application frequency of the traditional Tianjin (53-53) variant increases as age increases. The application frequency of the traditional Tianjin (53-21) variant decreases as age increases. The two application variants of (21-21) show different application patterns. The

new HH variant drops as age increases, while the old LH variant rises as age increases. Although the application frequencies of both the Tianjin variant of (53-53) and the old LH variant rises as age increases, the steepness of the old LH variant is more abrupt than that of the local variant of (53-53). Both variants are almost obsolete among young speakers, but application frequency of the LH variant is much higher than the local LL variant of (53-53) among old speakers. The application frequencies of both the Tianjin (53-21) variant and the new HH variant decrease as age increases, but the (53-21) line is fairly shallow, while the LL one is very steep. The application frequencies of both Tianjin (53-21) variant and new HH variant of (21-21) are very high among young speakers, but Tianjin (53-21) variant is much higher than HH variant of (21-21) among old speakers. Figure 6.1 also indicates that young speakers have smaller in-group variation than middle-aged and old speakers for all three variables, suggesting that the changes may be close to completion among young people.

In the following sections, I report on a series of mixed effects logistic regression analyses that I conducted on the sandhi variables separately. The goal of these analyses was to determine which linguistic and social factors contribute to the likelihood of the traditional Tianjin sandhi patterns for (53-53), (53-21) and (21-21) being applied in interview style, and whether they are the same as for word list style. This in turn contributes to the discussion, in Chapter 6, of what is motivating these generational changes to Tianjin tone sandhi.

6.2 Regression model construction

Mixed effects logistic regression analyses were conducted in R (R Core Team 2018) using the lme4 package (Bates, Maechler, Bolker & Walker 2015) to check whether social factors have significant effects on (53-21), (53-53) and (21-21) respectively. The input data comprised all 3883 tokens of (53-21), 5662 tokens of (53-53) and 2497 tokens of (21-21) produced by all of the 90

speakers in the interview style. For the analysis of (53-21) and (53-53), the dependent variable is binary: application versus non-application of the (53-21) and (53-53) sandhi rule, with the variant of interest being its application. The dependent variable, (21-21), has three variants: the new application variant HH, the old variant LH, and non-application of the (21-21) rule. For the purposes of the regression analysis, this was reduced to a binary variable, with the variant of interest being the new variant HH, in opposition to the two other variants combined.

As for the word list analysis, five fixed independent variables were included in the analysis: AGE GROUP (Old, Middle, Young), SOCIAL CLASS (working class and middle class), GENDER (male, female), ATTITUDE TO MIGRANTS (positive, negative), ATTITUDE TO TIANJIN (positive, negative). For the interview data, Middle-aged group is not merged with Young group because its application is not so biased towards the Young group as in word list style, as shown in Table 6.1 and Figure 6.1. Therefore age group is a three-level predictor, and I use dummy or default comparisons in which Young group and Old group are compared to Middle-aged group respectively. Reference levels were set at Middle age group, Working social class, Male gender, positive Attitude to Migrants, and positive Attitude to Tianjin. Two random effects were included: PARTICIPANT, and the output tone of the first word in each token, W1. P-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question. The significance threshold was set at <0.5 .

The specific fixed independent variables that were entered in a given regression analysis varied depending on the initial checking of the data, and relevant previous studies.

The fixed independent variables for the regression on (53-53) are AGE GROUP, SOCIAL CLASS, GENDER, and ATTITUDE TO TIANJIN. Attitude to migrants is not included for the same reasons introduced in section 5.2.

The fixed independent variables for the regression on (53-21) are AGE GROUP, ATTITUDE TO MIGRANTS, SOCIAL CLASS and ATTITUDE TO TIANJIN. After initial checking of the data, I found the application frequency difference by gender is only 2%, smaller than the difference by any other independent variables, so I do not include gender as an independent variable in the regression.

The fixed independent variables for (21-21) are AGE GROUP, ATTITUDE TO MIGRANTS, SOCIAL CLASS, GENDER and ATTITUDE TO TIANJIN, the same as the fixed predictors for (21-21) in the word list.

The VIF and tolerance statistics were run to test whether there was a strong correlation between two or more predictors on all models reported, and that no collinearity was found for any of them.

6.3 Factors predicting the application of the (53-53) sandhi rule

Mixed model regression indicates there are significant main effects of age group, $p < 0.001$, social class, $p < 0.001$ and gender on the application of (53-53) in interviews. The p-value of gender is 0.054, but I treat it as significant since its value is so close to the significance threshold. However, attitudes to Tianjin dialect $p > 0.05$, do not have significant main effects. The variances for random effects, PARTICIPANT and W1, are 1.451 and 1.427 respectively, not close to 0, which means it is reasonable to treat them as random effects and mixed effect model is good fit for my data.

The fixed effects of the generalized mixed model fit by maximum likelihood are given in Table 6.2.

Table 6.2 The parameter estimates for the generalized mixed model of (53-53) interview

Fixed effects	Estimate	SE	Z	P
(intercept)	-3.971	0.386	-10.276	<0.001
Age group (Old)	1.616	0.376	4.298	<0.001
Age group (Young)	-1.76	0.36	-4.896	<0.001
Social class (Working)	1.807	0.304	5.949	<0.001
Gender (Male)	0.572	0.297	1.924	=0.054

As shown in Table 6.2, Old speakers were 5 ($\exp(1.616) = 5.033$) times more likely to use the (53-53) rule than the reference group, Middle speakers, and Young speakers were 82.8% ($1 - \exp(-1.76) = 0.828$) less likely than the reference group to use the (53-53) rule. This confirms the overall trend shown in Figure 6.1 for (53-53), where there was a positive correlation between age and (53-53) application. Working class are 6 times ($\exp(1.807) = 6.092$) more likely to use this variable than middle class, and males are 1.77 times ($\exp(0.572) = 1.772$) more likely to apply the sandhi rule than females, consistent with the tendency observed in wordlist style. People with positive attitudes to Tianjin dialect are more likely to use the sandhi than people with negative attitudes to Tianjin dialect, which is the same in word list style too, but the difference is much smaller in interview style than in word list style.

The VIF and tolerance statistics were run to test whether there was a strong correlation between the two predictors in the regression model. VIF results of both Age group and social class are 1.009, and gender is 1.005. The $1/\text{VIF}$ of age group and social class is 0.99, and gender is 0.995.

The VIF values are well below 10, and the tolerance statistics all well above 0.2. Based on these measures, I conclude that there is no collinearity between the predictors.

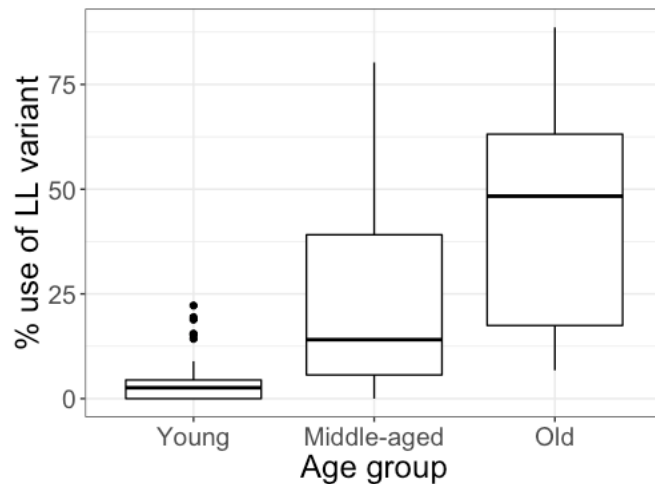
6.3.1 The effect of age group

The output of mixed regression analysis indicates that age group has a significant overall effect on the application of (53-53) interview, $p < 0.001$. The difference in rates of application across age groups is given in Table 6.3, and displayed in the boxplot in Figure 6.2.

Table 6.3 The application frequency of LL variant in interview by age group

Age groups	LL variant in interview		
	App / Tokens	%	SD
Young	94/ 2260	4.2%	2.19
Middle	488/ 1861	26.7%	21.19
Old	673/ 1541	43.7%	24.91
Total	1255/ 5662	22.2%	20.27

Figure 6.2 Application frequency of LL variant in interview by age group



First, the overall application frequency of (53-53) is 22.2% in interview, higher than that in word list (11.5%), and the tendency is consistent in every age group. Second, the standard deviation (SD) values in Table 6.3 indicate that there is a lot of interspeaker variation in the Old, while the Young age group exhibited almost no within-group variation, which is similar to the

trend shown in (53-53) word list style. The difference is that the Middle-aged group in interview style exhibits a lot of within-group variation in (53-53) application, but in the word list style, this variation was very small. Since speakers pay less attention to their speech in less careful interview style, the bigger variation within the Middle-aged group suggests that the local application variation may still exist in this age group when they are less careful about their speech.

6.3.2 The effect of social class

According to the results of mixed regression analysis reported in Table 2, social class has a significant overall effect on the application of (53-53) interview, $p < 0.001$. Table 6.4 is the application frequency of the local LL variant of (53-53) by social class, and the difference is plotted in Figure 6.3.

Table 6.4 The application frequency of LL variant by social class

Social class	LL variant in interview		
	App/Tokens	%	SD
Middle class	259/ 2531	10.2	9.68
Working class	996/ 3131	31.8%	24.41
Total	1255/ 5662	22.2%	20.27

Figure 6.3 Application frequency of LL variant in interview by social class

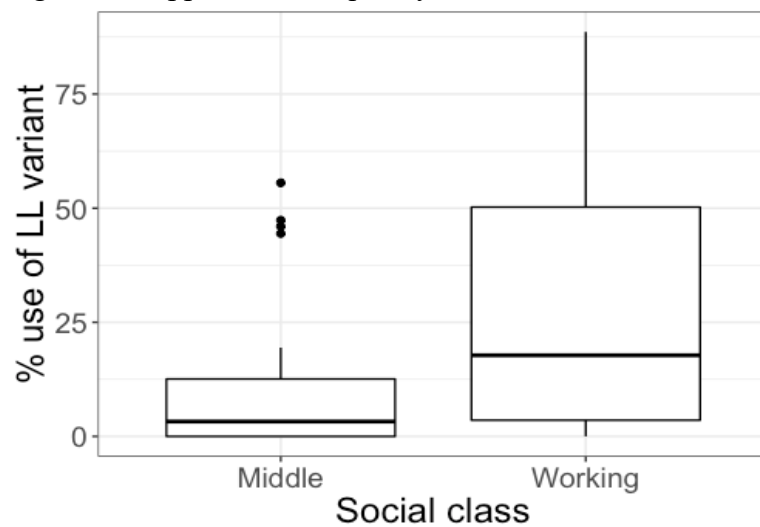


Table 6.4 and Figure 6.3 show that the application rates of (53-53) are much higher in the working class than in the middle class, which is similar to the significant tendency observed in (53-53) word list. Since the non-application variant is the new, standard and prestigious form, the higher application rates of the application variant indicate that the working class is retaining the old, local and non-prestigious form more than the middle class. In other words, the change towards the new, standard and prestigious form is led by middle class.

Although overall the working class has a higher application frequency of the old Tianjin variant LL, the social class distribution in every age group is not clear. Table 6.5 displays the application rates of the LL variant in interview by social class and age group.

Table 6.5 The application frequency of LL variant by social class and age group

	Middle class		Working class		Total	
	App/Tokens	%	App/Tokens	%	App/Tokens	%
Young	35/1165	3	59/1095	5.4	94/2260	4.2
Middle	72/666	10.8	416/1195	34.8	488/1861	26.2
Old	152/700	21.7	521/841	62	673/1541	43.7
Total	259/2531	10.2	996/3131	31.8	1255/5662	22.2

As indicated in Table 6.5, the working class has higher application frequency of LL, the local and non-prestigious variant, than the middle class consistently in every age group, which is consistent with the trend observed in word list style. Old working class speakers have the highest application rates in the six subgroups by age group and social class, suggesting that this subgroup retains the local variant more than any other subgroups by age group and social class. Young middle class speakers have the lowest application rates in the six subgroups, suggesting that this subgroup leads the change toward the standard and prestigious variant⁹.

6.3.3 The effect of gender

Mixed regression analysis introduced in section 5.2 indicates that gender has an approximate significance on the application of (53-53) in interview, $p=0.054$. Table 6.6 is the application rates of the traditional LL variant by gender. The boxplot of the application rates of (53-53) by gender is given in Figure 6.4.

⁹ Although the non-application variant, T4 (53), is a Tianjin tone, its pitch value difference from T4 in SC (51) is not noticed by Tianjin speakers. Many participants said that T4 in Tianjin is the same as T4 in SC. Besides, SC does not have 53-53 sandhi rule, so the non-application variant is more like SC in this sense.

Table 6.6 The application frequency of LL variant by gender

Gender	LL variant in interview		
	App / Tokens	%	SD
Female	609/ 3154	19.3	22.04
Male	646/ 2508	25.8	18.21
Total	1255/ 5662	22.2	20.27

Figure 6.4 The application frequency of LL variant in interview by gender

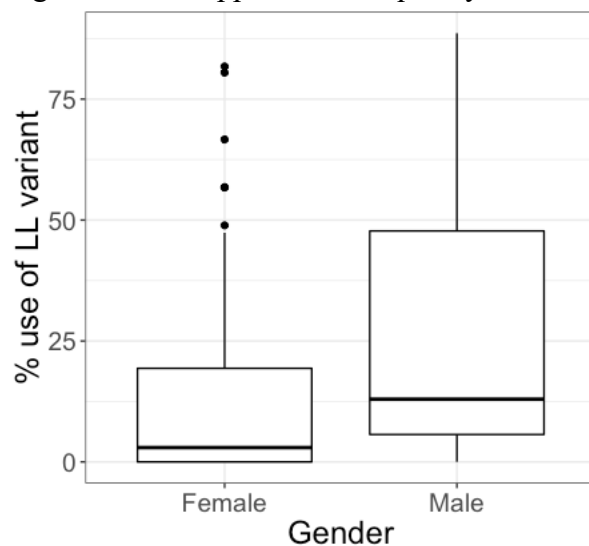


Table 6.6 and Figure 6.4 show that men have higher application rates of the local variant LL in interview style than women, which is similar to the non-significant tendency shown in word list style. That men have higher application rates of the old, local and non-prestigious variant indicates that it is women who lead the change towards the new, standard and prestigious variant.

It is unclear whether the gender difference in the Tianjin data is related to social classes: an interaction was observed in many previous sociolinguistic studies of language change (Labov, 1990). The contrasts between the application rates of LL variant in interview by social class and gender are given in Table 6.7.

Table 6.7 Application frequency of LL variant in interview by social class and gender

Gender Social class	Female		Male		Total	
	App/Tokens	%	App / Tokens	%	App / Tokens	%
Middle class	108/1547	7	151/984	15.3	259/2531	10.2
Working class	501/1607	31.2	495/1524	32.5	996/3131	31.8
Total	609/3154	19.3	646/2508	25.8	1255/5662	22.2

Table 6.7 indicates that men have higher application frequency of the local and non-prestigious variant LL than women have in every social class. Working class men have the highest application rates in the subgroups by social class and gender, while middle class women have the lowest application rates. The results suggest that the change towards the standard variant might be led by middle class women, and the local variant is best preserved among working class men. Furthermore, the gender difference is bigger in the middle class than in the working class.

6.4 Factors predicting the application of the (53-21) sandhi rule

The fixed effects for the generalized mixed model fit by maximum likelihood are given in Table 6.8. There are significant main effects of age group, $p < 0.001$, migrants, $p < 0.05$, and social class, $p < 0.05$, on the application of the traditional (53-21) sandhi rule in interview style. However, attitudes to Tianjin dialect, $p > 0.05$, do not have significant main effects.

The variances for random effects, PARTICIPANT and W1, are 0.222 and 1.091 respectively, not close to 0, which means it is reasonable to treat them as random effects and mixed effect model is good fit for my data.

Table 6.8 The parameter estimates for the generalized mixed model of (53-21) in interview style

Fixed effects	Estimate	SE	Z	P
(intercept)	1.971	0.225	8.767	<0.001
Age group (Old)	-0.801	0.194	-4.123	<0.001
Age group (Young)	0.704	0.189	3.73	<0.001
Social class (Working)	0.291	0.147	1.975	<0.05
Attitude to migrants (Positive)	-0.434	0.155	-2.792	<0.05

Old speakers are 55.1% ($1 - \exp(-0.801) = 0.551$) less likely to use the (53-21) rule than the reference group, Middle-aged speakers, and Young speakers are 2 times ($\exp(0.704) = 2.022$) more likely than the reference group to use the (53-21) rule. This confirms the overall trend shown in Figure 6.1 for (53-21), where there is a negative correlation between age and (53-21) application. Working class speakers are 1.338 times ($\exp(0.291) = 1.338$) more likely to use the (53-21) rule than the middle class. People with a positive attitude to migrants are 35.2% ($1 - \exp(-0.434) = 0.352$) less likely to use the (53-21) rule than people with a negative attitude to migrants.

The VIF and tolerance statistics were run to test whether there was a strong correlation between the two predictors in the regression model. VIF results of age group, social class and attitude to migrants are 1.127, 1.013, and 1.113 respectively. The $1/\text{VIF}$ of age group, social class and attitude to migrants are 0.888, 0.987 and 0.898 respectively.

The VIF values are well below 10, and the tolerance statistics all well above 0.2. Based on these measures, I conclude that there is no collinearity between the predictors.

6.4.1 The effect of age group

The output of mixed regression analysis indicates that age group has a significant overall effect on the application of (53-21) interview, $p < 0.001$. The rates by age group are given in Table 6.9 and presented visually as boxplots in Figure 6.5.

Table 6.9 The application frequency of HH variant by age group

Age group	HH variant in interview		
	App / Tokens	%	SD
Young	1751/1943	90.1	7.91
Middle-aged	930/1111	83.7	11.27
Old	607/829	73.2	10.07
Total	3288/3883	84.7	11.59

Figure 6.5 Application frequency of HH variant in interview by age group

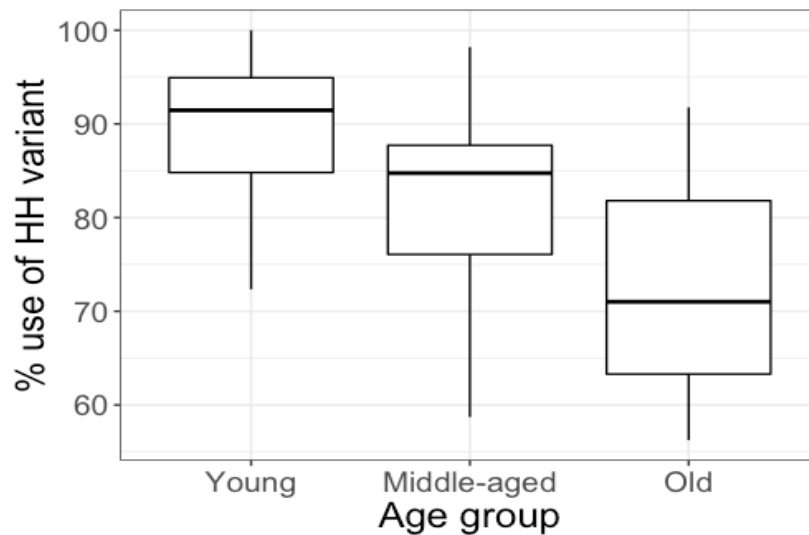


Table 6.9 and Figure 6.5 show age differences clearly. The Young group has the highest application frequency for (53-21), while the Old group has the lowest frequency, with the Middle-aged group in the middle. The largest box of the Old group in Figure 6.5 indicates that this age

group has the biggest variability within the interquartile range, suggesting the biggest variation within the middle 50% of observations. Although the box of the Middle-aged group in Figure 6.5 is smaller than that of the Old group, the Middle-aged group has the longest whisker, indicating the biggest variability outside the upper and lower quartiles, which is reflected in the biggest SD in Table 6.9. The Young group has smaller in-group variation, indicated by their measures of variance in Table 6.9 and the smaller, more compressed boxes in Figure 6.5. Overall, the results are similar to those for (53-21) in wordlist style, except that in the wordlist style, the Young and Middle groups had almost no in-group variation.

6.4.2 The effect of attitudes to migrants

According to the results of mixed regression analysis reported in Table 8, attitudes to migrants have a significant effect on the application frequency of (53-21) in interviews, $p < 0.05$. Table 6.10 reports the application rates of (53-21) in interviews for speakers with positive or negative views of migrants. Figure 6.6 displays this distribution visually.

Table 6.10 The application frequency of HH variant by attitude to migrants

Age group	HH variant in interview		
	App/Tokens	%	SD
Negative attitudes	2030/2324	87.3	9.47
Positive attitudes	1258/1559	80.7	12.12
Total	3288/3883	84.7	11.59

Figure 6.6 Application frequency of HH variant in interview by attitude to migrants

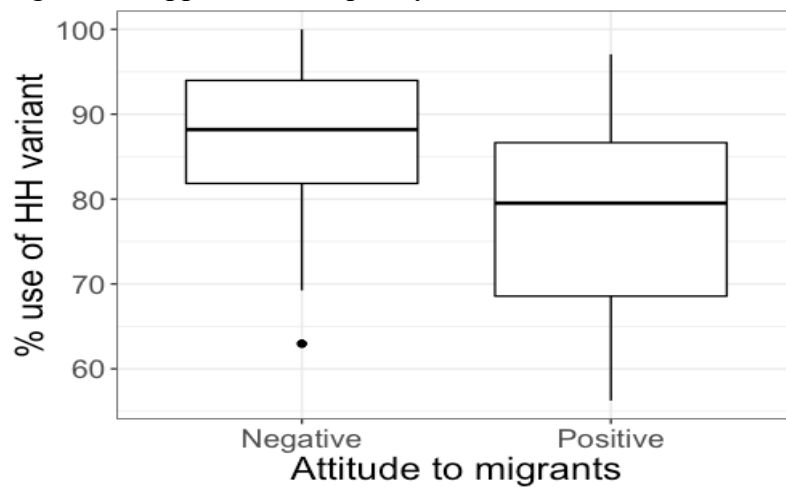


Table 6.10 and Figure 6.6 indicate that people having negative attitudes towards migrants have significantly higher application rates of the traditional Tianjin (53-21) pattern in interview style, which is consistent with the tendency shown in the word list. But it is unclear whether attitudes to migrants are related to the other two significant predictors of (53-21) interview: age group and social class. The distribution of attitudes to migrants in every age group is reported in Table 6.11. The distribution of attitude to migrants in every social class is provided in subsection 6.4.3.

Table 6.11 The application frequency of HH variant by attitude to migrants and age group

Attitude Age group	Negative		Positive		Total	
	App/Tokens	%	App / Tokens	%	App / Tokens	%
Young	1310/1452	90.2	441/491	89.8	1751/1943	90.1
Middle	371/417	89	559/694	80.5	930/1111	83.7
Old	349/455	76.7	258/374	69	607/829	73.2
Total	2030/2324	87.3	1258/1559	80.7	3288/3883	84.7

As shown in Table 6.11, the difference in (53-21) application rate between speakers having negative attitudes to migrants and positive attitudes to migrants is very minimal within Young group, but the differences by negative or positive attitudes to migrants are big in both the Middle-aged and Old groups, with Middle-aged group having the biggest difference. Besides, the proportions of speakers with negative attitudes to migrants (29 negative, 13 positive) are much higher in the Young group than in the other two age groups. About two thirds young speakers have negative attitudes to migrants, while less than half speakers have negative attitudes to migrants in the Middle-aged and Old groups. The Middle-aged speakers seem to be most friendly to migrants, judging by the proportion of speakers with positive attitudes to migrants. Why are the Young the most negative about migrants, and why doesn't this have an effect on their rate of (53-21) application? Why are the Middle group the most positive about migrants, and why is this affecting their use of (53-21)? These questions will be discussed in Chapter 7.

6.4.3 The effect of social class

According to the results of mixed regression analysis reported in Table 6.8, social class has a significant effect on the application frequency of (53-21) in interviews, $p < 0.05$. Table 6.12 and Figure 6.7 display the application frequency of (53-21) in interview by social class.

Table 6.12 The application frequency of HH variant by social class

Social class	HH variant in interview		
	App / Tokens	%	SD
Middle class	1489/1805	82.5	12
Working class	1799/ 2078	86.6	11
Total	3288/3883	84.7	11.59

Figure 6.7 Application frequency of HH variant in interview by social class

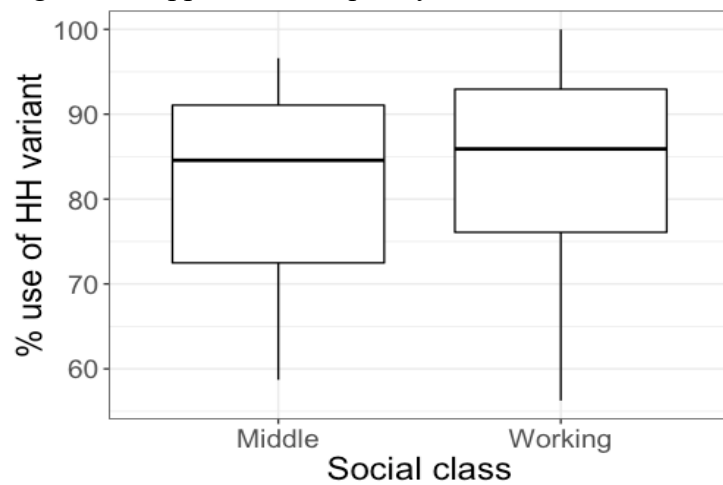


Table 6.12 and Figure 6.7 indicate that overall working class has higher application rates than middle class, which is similar to the tendency shown in word list style, but it is unclear whether this overall trend is consistent in every age group. Table 6.13 is the distribution of the application of (53-21) sandhi rule by social class and age group, indicating that working class has higher application frequency of the local variant than middle class consistently in every age group. But the degree of difference in every age group is different. The social class difference within the Young group is the smallest, while the social class difference in the Old group is the biggest.

Table 6.13 The application frequency of HH variant by social class and age group

	Middle class		Working class		Total	
	App/Tokens	%	App/Tokens	%	App/Tokens	%
Young	917/1038	88.3	834/905	92.2	1751/1943	90.1
Middle	345/428	80.6	585/683	85.7	930/1111	83.7
Old	227/339	67	380/490	77.6	607/829	73.2
Total	1489/1805	82.5	1799/ 2078	86.6	3288/3883	84.7

Community changes towards increasing adoption of a non-standard variant are often led by the interior or lower classes (see e.g. Labov, 1966; Holmes, 1995). The local Tianjin application variant, HH, is the non-standard and non-prestigious variant. The results of the present study indicate that the renewed rise of the old, local, non-standard and non-prestigious variant seems to be led by the working class, which is consistent with the findings of previous sociolinguistic studies.

We saw that working class speakers have a higher application frequency of the (53-21) than middle class speakers, and the tendency is consistent in every age group. We also saw that a negative attitude to migrants predicted higher rates of (53-21) sandhi rule application than a positive attitude to migrants, and the tendency is consistent in every age group in §6.4.2. Now we want to check the application frequency of (53-21) sandhi rule by social class and migrants.

Table 6.14 The application frequency of HH variant by social class and attitude to migrants

Attitude Social class	Negative attitude		Positive attitude		Total	
	App / Tokens	%	App/ Tokens	%	App/Tokens	%
Middle class	950/1124	84.5	539/681	79.1	1489/1805	82.5
Working class	1080/1200	90	719/878	81.9	1800/2078	86.6
Total	2030/2324	87.3	1258/1559	80.7	3288/3883	84.7

Table 6.14 indicates that people with negative attitudes to migrants have higher application rates than those with positive attitudes to migrants, and this tendency is true in both middle class and working class. The difference between different attitudes in working class is bigger than the difference in middle class. Working class speakers with negative attitudes have the highest application rates of (53-21) tone sandhi in the subgroups by social class and attitudes to migrant, while middle class speakers with positive attitudes to migrants have the lowest application rates in

these subgroups. This tendency suggests that the increase in the application of (53-21) might be led by working class speakers with negative attitudes to migrants. What we have found in the results for (53-21) in interviews is consistent with the results for (53-21) in wordlist style.

6.5 Factors predicting the application of the (21-21) sandhi rule

Mixed model logistic regression was carried out in R (R Core Team 2018) using the packages, procedures and predictors introduced in §6.2. In this model, the dependent variable was HH vs LH plus non-application. The fixed effects for the generalized mixed model fit by maximum likelihood are given in Table 6.15.

There are significant main effects of age group, $p < 0.001$, and social class, $P < 0.05$, on the application of (21-21) in interview style. However, gender, $p > 0.05$, migrants, $p > 0.05$, and attitudes to Tianjin dialect, $p > 0.05$, in general, do not have significant main effects.

The variances for random effects, PARTICIPANT and W1, are 2.205 and 0.391 respectively, not close to 0, which means it is reasonable to treat them as random effects and mixed effect model is good fit for my data.

Table 6.15 The fixed effects for the generalized mixed model of (21-21) interview

Fixed effects	Estimate	SE	Z	P
(intercept)	0.335	0.381	0.877	n.s.
Age group (Old)	-1.454	0.467	-3.112	<0.05
Age group (Young)	2.051	0.412	4.983	<0.001
Social class (Working)	-0.679	0.345	-1.967	<0.05

Old speakers are 76.6% ($1 - \exp(-1.454) = 0.766$) less likely to use HH, the new local variant than the reference group, Middle-aged speakers, and Young speakers are 7.776 times ($\exp(2.051) = 7.776$) more likely than the reference group to use the new local variant HH. This confirms the

overall trend shown in Figure 6.1 for (21-21), where there is a negative correlation between age and the application of the new local variant HH. Working class speakers are 49.3% ($1 - \exp(-0.679) = 0.493$) less likely to use the new local variant than the middle class.

The VIF and tolerance statistics were run to test whether there was a strong correlation between the two predictors in the regression model. VIF results of age group and social class are 1.008. The $1/\text{VIF}$ of these two groups are 0.992.

The VIF values are well below 10, and the tolerance statistics all well above 0.2. Based on these measures, I conclude that there is no collinearity between the predictors.

6.5.1 The main effects of age group

The output of mixed regression analysis indicates that age group has a significant overall effect on the application of (21-21) interview, $p < 0.001$.

Tables 6.16-6.17 display the application frequencies of the HH variant and the LH variant respectively. Recall that, different from the other variables of the present study, (21-21) has two application variants (HH and LH). When I calculate the application frequency of (21-21), the frequency of HH variant is HH vs the total number of tokens (HH variant + LH variant + non-application variant), while the LH frequency is the LH variant vs the total number of tokens (HH variant + LH variant + non-application variant). Figures 6.8-6.9 are the boxplots of the application frequencies of HH variant and LH variant respectively.

Table 6.16 The application frequency of HH variant by age group

Age group	HH variant in interview		
	App/Tokens	%	SD
Young	1026/1199	85.57	2.62
Middle	420/740	56.7	36.05
Old	143/558	25.63	21.3
Total	1589/2497	63.64	33.97

Table 6.17 The application frequency of LH variant by age group

Age group	LH variant in interview		
	App/Tokens	%	SD
Young	31/1199	2.59	5.56
Middle	290/740	39.19	37.15
Old	395/558	70.79	24.9
Total	716/2497	28.67	37.28

Figure 6.8 Application frequency of HH variant in interview by age group

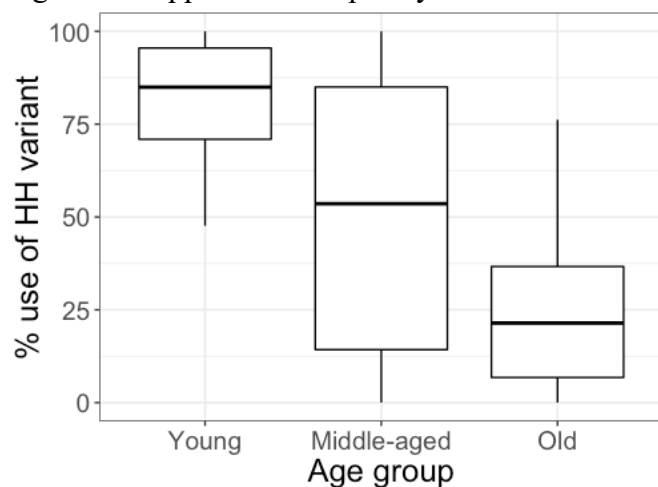
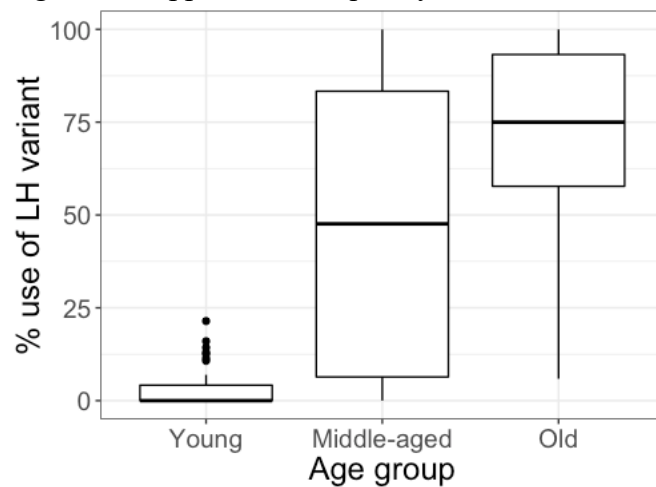


Figure 6.9 Application frequency of LH variant in interview by age group



The application of (21-21) tone sandhi changes clearly with age groups, as shown in Tables 6.16-6.17 and Figures 6.8-6.9. It seems that the local variant is changing from LH to HH, consistent to the tendency observed in wordlist style. The highest SD in Table 6.16-6.17 and the biggest box in Figures 6.8-6.9 indicate that the in-group varieties of Middle-aged group is the biggest among the three age groups.

6.5.2 The main effects of social class

According to the results of mixed regression analysis reported in Table 6.8, social class has a significant overall effect on the application of (21-21) sandhi rule with the HH output variant in interview style, $p < 0.05$. The application frequencies of new and old local variants by social class are given in Table 6.18 and Table 6.19 respectively. Figures 6.10-6.11 are the boxplots of the applications of the new and old variants respectively.

Table 6.18 The application frequency of HH variant by social class

Social class	HH variant in interview		
	App/Tokens	%	SD
Middle class	854/1217	70.17	30.74
Working class	735/1280	57.42	34.96
Total	1589/2497	63.64	33.97

Table 6.19 The application frequency of LH variant by social class

Social class	LH variant in interview		
	App/Tokens	%	SD
Middle class	253/1217	20.8	34.42
Working class	522/1280	40.8	40.8
Total	775/2497	32.06	37.28

Table 6.18 shows that the change towards the new HH variant seems to be led by middle class, since the application frequency of the HH variant is 70.17% in the middle class, 12.75% higher than that of working class. Additionally, Table 6.19 shows that the application mean of the LH variant is only 20.8% in the middle class, 20% lower than that of the working class. The tendency is consistent with the results of word list results, though the social class difference in word list style is not significant.

Figure 6.10 Application frequency of HH variant in interview by social class

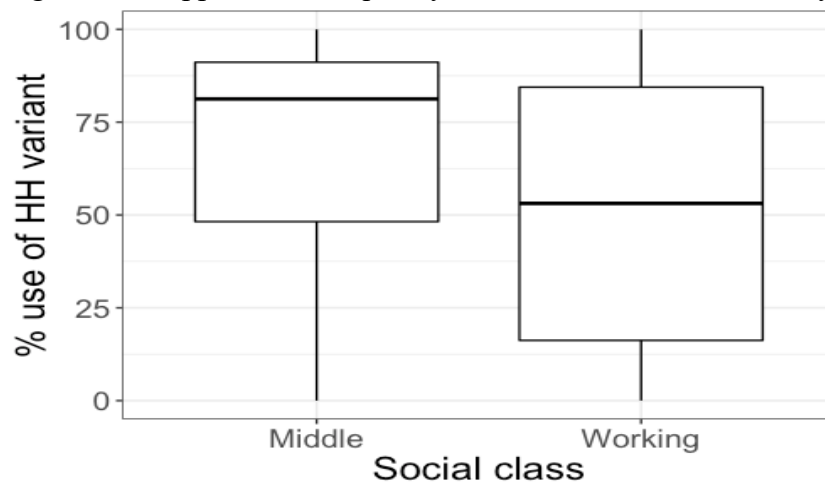
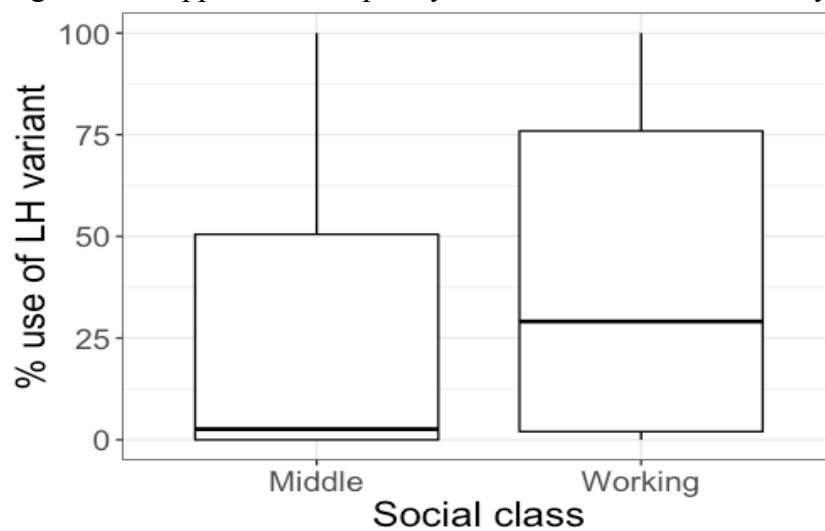


Figure 6.11 Application frequency of LH variant in interview by social class



We have seen that overall the middle class seems to be leading in the adoption of the HH variant of (21-21), but it is unclear whether the tendency is consistent in every age group. Table 6.20 displays the application frequency of the HH variant by age group and social class.

Table 6.20 Distribution of HH variant by age group and social class

Age group	Middle class		Working class	
	App/tokens	%	App/Tokens	%
Young	538/617	87.2	488/582	83.8
Middle	233/326	71.5	187/414	45.2
Old	83/274	30.3	60/284	21.1
Total	854/1217	70.2	735/1280	57.4

The application frequencies of HH variant among middle class speakers are shown to be consistently higher than those of working class speakers in every age group, further supporting the proposal that the change towards the new local variant is led by the middle class.

6.6 Summary

Social factors have shown complicated effects on the applications of the three traditional Tianjin sandhi rules in interview style. Age and social class have significant effects on all three variables. With regard to age, the application frequency of the non-prestigious Tianjin variant of (53-53) sandhi rule increases as age increases, while the application frequency of the non-prestigious Tianjin variant of the (53-21) sandhi rule decreases as age increases. The (21-21) variable has two Tianjin application variants, HH and LH. HH is the new Tianjin variant, preferred by young speakers, while LH is the old Tianjin variant, preferred by old speakers. The old variant LH shows a positive correlation with age, but the new variant HH exhibits a negative correlation with age.

With regard to social class, middle class speakers have a higher application frequency of the local non-prestigious variant of the (53-53) sandhi rule than working class speakers, and the tendency is consistent in both men and women in every age group, suggesting the change towards

the standard and prestigious variant of (53-53) might be led by the middle class. As for (53-21), working class speakers have a higher application frequency of the local and non-prestigious variant HH than middle class speakers, and the trend is consistent in every age group. The application frequency of the new local variant HH for the (21-21) sandhi rule is higher among middle class speakers is higher than that among working class speakers.

Social class does not have significant effects in the application of (53-21) and (21-21) in word list style, but it is a significant predictor of (53-21) and (21-21) in interview style. This echoes the observation that social class difference tends to be minimized in more careful style (Trudgill 1972). This is also a reason why interview style should be studied in addition to wordlist style.

Consistent with the results of (53-21) word list, speakers with negative attitudes to migrants have significantly higher application frequency of the local variant HH. Speakers with negative attitudes to migrants may be more eager to mark themselves as different from migrants than speakers with positive attitudes to migrants, and therefore they have higher application frequency of the local variant HH. The increase in (53-21) application frequency since the beginning of the 1990s coincides with the in-migration history of Tianjin.

After analyzing wordlist results and interview results separately, the stylistic variation will be studied by comparing the two styles, with a purpose of revealing speakers' attitudes to Tianjin dialect. The stylistic variation will be reported in the next chapter.

Chapter 7 Analysis and discussion

In Chapters 5 and 6, Tianjin tone sandhi variation was analyzed separately within word list and interview speech. In the present chapter, I provide a summary of the interspeaker variation patterns observed so far. The chapter additionally considers intraspeaker variation, also known as style: How do speakers vary their proportional application of traditional Tianjin sandhi rules in formal vs more informal speech? Collectively, the inter- and intra-speaker variation patterns provide diagnostics for some discussion of the possible social motivations for changes to the (53-53), (53-21) and (21-21) sandhi rules. Some linguistic motivations are also discussed. In particular, I consider whether the three sandhi variables are undergoing ‘change from above’ or ‘change from below’. Section 7.1 gives a brief overview of sociolinguistic style, while in section 7.2, I return to the concepts of ‘change from above’ and ‘change from below’, which were briefly introduced and defined in Chapter 1. In section 7.3, I return to the concepts of ‘indicator’, ‘marker’ and ‘stereotype’, which were briefly introduced in Chapter 1. A combination of intra-speaker variation and inter-speaker variation could bring light to the motivations of language changes, so the chapter will discuss the three variables one by one from both intra-speaker variation and inter-speaker variation.

7.1 Style and sociolinguistic variation

Intra-speaker stylistic shifts are affected by the attention the speaker paid to the speech (Labov, 1972) and also the addressee(s) and other audiences (Bell, 1984). More recently, sociolinguists have also integrated ‘speaker design’ considerations, i.e. the degree to which speakers agentively construct their identity via sociolinguistic style (Coupland, 1985; Schilling-Estes, 1998; Podesva, 2008). In the current study, all participants were recorded speaking to a single interlocutor (the author), and so audience effects are not examined. Speaker identity is

beyond the scope of this dissertation. Instead, I adopt the attention-to-speech model, which has been used successfully for decades in sociolinguistic research. In his study of New York City, Labov (1972: 79) employed four experimentally controlled contexts, designed to incrementally increase the speaker's attention to their own speech: Context A, "for those situations which escape the social constraints of the interview situation"; Context B, the interview situation; Context C, reading style, and Context D and D' (word list and minimal pairs respectively). The two styles in this study are Context B, the interview situation and Context D, word list situation¹⁰ The expectation is that speakers are more careful about their speech in word list reading style than in less careful interview style (Labov, 1966). Because they are more careful in the word list reading style, the variant they use more frequently in the word list is usually the variant they consider (consciously or unconsciously) to be more prestigious.

Stylistic variation is a window into the speech community's social evaluation of a sociolinguistic variant. And that is a more useful tool than e.g. asking people for their opinions of a variant because people could hide their real opinions for various reasons if they are asked directly, but they cannot totally control their speech especially in less careful interview style.

By comparing interview and wordlist speech, we can find out whether the application rates of a given Tianjin variant are different in different styles. If the prestigious non-application variant appears more frequently in the more careful style consistently, we have good reason to propose that it is a change from above (Section 7.2 below). If the application variant appears more frequently in the more careful word list style, or it appears less frequently in the less careful interview style, we tend to suggest that it is a change from below (Section 7.2 below). Of course,

¹⁰ I did not distinguish, for the interview speech, between more or less casual styles, as has been done in other studies (Eckert & Rickford 2002:85-108).

if a variant is also undergoing a change of evaluation in the speech community over time, the youngest age group may treat a formerly prestigious variant as non-prestigious, or vice versa (Dubois & Horvath, 2000). Thus it will also be important to examine style-shifting patterns in apparent time for each sandhi variable.

7.2 ‘Change from above’ and ‘Change from below’

Labov (1966: 328) introduced the terms ‘change from above’ and ‘change from below’ principally to distinguish changes that are above public awareness from those that are below awareness. “Change from above” (Labov, 1966 & 1994) is a language change above the level of social awareness. A well-known example is Labov’s (1966) characterization of the increasing frequency of post-vocalic [r] in New York City as a change from above. “Change from below” (Labov, 1966 & 1994) is a language change below the level of social awareness at first, whereby the winning variant typically evolves from within the community. The deletion of (l) in words like *il* ‘he’ and *elle(s)* ‘she (them, fem)’ in Montreal French (Poplack & Walker, 1986) is an example of a change from below.

Speaker gender¹¹, social class and style provide clues as to whether a change can be considered ‘from above’ or ‘from below’ (Labov, 1990). It has been evident from the earliest sociolinguistic studies that gender plays an important role in language change. Labov (1966), Trudgill (1972, 1995) and Labov (1990, 2001a), for example, have all discussed gender influences

¹¹ The differences between male and female linguistic preferences have been explored extensively in the literature. There are some linguists who attribute gender preferences to relations with power (Coates & Cameron 1988, Uchida 1992, West & Zimmerman 1983, Matus-Mendoza 2004). Matus-Mendoza (2004) argues that “female preference for the innovative and prestigious variant as a way to acquire power in a society where women do not have a voice to claim or seek social mobility and power”. Others analyze “social forces that attach to women by virtue of their place in the economy” (Eckert 1989:225). Labov (1990) and Chambers (1992) explained how sociolinguistic factors intertwine with linguistic factors. I do not go into further detail regarding *why* gender differences obtain in cases of language change, but refer the reader to these sources.

on language. Labov (1990, 2001a) has drawn some well-known generalizations about gender and language variation and change. For diachronically stable sociolinguistic variables such as (ing) and word-final consonant-cluster reduction, women are likely to use proportionally more of the standard variant than men (Fischer, 1958; Labov, 1966; Wolfram, 1969; Cofer, 1972; Trudgill, 1974; Macaulay, 1977). For a change from above in progress, women tend to use more of the overtly prestigious variant than men do (Labov, 1966; Lennig, 1978; Milroy & Milroy, 1985). For change from below in progress, women tend to use more of the incoming variant than men (Trudgill, 1974; Luthin, 1987).

There is a common—but not predictable—relationship between change from above and social class. When an incoming sound change is above the level of awareness, people from higher social classes typically tend to use the more prestigious form more frequently than lower social classes (e.g. Labov, 1966; Carvalho, 2004). For example, the prestige rhotic pattern was introduced into New York City’s vernacular by upper-middle-class New Yorkers, and the change was (and still is) above public awareness (Labov, 1966). Similarly, Baranowski (2008) reports on the diffusion of Southern (ow)-fronting to the conservative city of Charleston, in a study of 32 speakers aged 15-65. He found that the Upper Middle class had the highest mean F2 of (ow), while the Working class had the lowest mean F2 of (ow). The (ow)-fronting change is above the level of social awareness (Baranowski 2008:548). When a sound change is below the level of awareness, by contrast, people from lower social class typically, though not predictably, tend to use the less prestigious form more frequently than higher social classes. The deletion of (l) in Montreal French (Poplack & Walker, 1986) seems to be led by lower social classes, and that is a change from below.

Stylistic variation is also helpful in revealing change patterns, as change from above tends to have proportionally more of the innovative form in the more careful style, indicating that it

carries normative or standard prestige. Post-vocalic /r/ in New York city stands as a well-known example (Labov, 1972). In the department store study, post-vocalic /r/ is related to speakers from higher social class, indicating the presence of post-vocalic /r/ is a prestigious variant. And post-vocalic /r/ tends to drop in casual speech in the rapid and anonymous studies, suggesting the more prestigious variant tends to appear more frequently in the more careful speech. The point is further supported in other production studies (Labov, 1972), which found that the dropping of post-vocalic /r/ appears more frequently in casual speech, while in careful speech the rates were lower.

Changes from below exhibit different stylistic shifts from changes from above. As Trudgill (1972) put it, there is either no stylistic variation or the prestigious variant does not appear more frequently in the more careful style in changes from below. Which type of stylistic variation is shown in a specific change from below depends on the level of social awareness of the innovative linguistic form. This is discussed in the next section.

7.3 ‘Indicator’, ‘marker’ and ‘stereotype’

It is a common practice to assume a speech feature has group-salient or situation-salient meaning if the feature is used more frequently by one group rather than another group, or in one situation more than another. This relation between social meaning and variable is formalized in Labov’s use of ‘indicator’, ‘marker’ and ‘stereotype’ (Labov, 1972: 237), differing in the degree of social awareness they bear.

There are no stylistic shifts at the so-called ‘indicator’ stage (Labov, 1972), during which only social class variation is observed. Indicators “show a regular distribution over socioeconomic, ethnic, or age groups, but are used by each individual in more or less the same way in any context” (Labov, 1972:237). Changes from below at ‘indicator’ stage does not show stylistic variations.

Some phonological mergers such as the one affecting the back vowels of ‘cot’ and ‘caught’ operate as indicators.

But there are stylistic variations once they’ve progressed from ‘indicator’ stage to ‘marker’ stage, during which both social and stylistic variation can be observed, but no metalinguistic mentions of the variable. Markers “not only show social distribution, but also stylistic differentiation” (Labov, 1972:237). Markers will produce regular responses on subjective reaction tests although they might lie below the level of consciousness (Labov, 1972: 314).

Changes from below at ‘marker’ stage usually show stylistic variation, in the prestigious variant does not appear more frequently in the more careful style. The raising of /æ/ in New York City is a marker, as it was stratified by social class, with raising more common at the lower end of the socioeconomic scale, and subject to style-shifting, with the highest rates of raising in casual speech.

Stereotypes are variables that are highly salient to speakers and are subject to overt comment and control. “Stereotypes are socially marked forms, prominently labelled by society” (Labov, 1972: 314). Stereotypes may have a general label or a characteristic phrase serving to identify them. And of course at the ‘stereotype’ stage, a former change from below will exhibit stylistic variation (Labov, 1972:237). Some stereotyped features could be heavily stigmatized, which usually led to total extinction, but some stigmatized features could be resistant and enduring (Labov, 1972:315). One example of stereotype is (ing) in English: the presence or absence of a final velar for unstressed /ing/ can signal social and stylistic stratifications, and English speakers have metalinguistic knowledge of it (Labov, 1972).

‘Indicator’, ‘marker’ and ‘stereotype’ represent different social awareness of linguistic changes. Changes from above and changes from below depend on social awareness too. So

‘indicator’, ‘marker’ and ‘stereotype’ are indispensable in determining changes from above and changes from below. Whether a change is from above or from below reveals social meanings of variants, which is one main purpose of the current study.

7.4 The change in (53-53)

7.4.1 (53-53): a decrease from above

As shown in Chapter 5 and Chapter 6, age group and social class significantly predict the application of (53-53) in both word list and interview styles, and gender is only significant in interview style. We saw that application of the traditional rule with its LL variant is used by men more than women, by older people more than younger people, and by the working class more than the middle class. All of this suggests that the LL variant is being replaced by an incoming change from above: the non-application of the sandhi rule, which I have also referred to as “the HH variant”. Although social class and gender can provide some hints as to whether a change is from above or not, the relation between them is not predictable because changes from above (Labov, 1966) refer to changes above speakers’ consciousness, not from higher social class. So it is necessary to study stylistic variations of LL variant to confirm whether the change is from above or not.

In this section, I provide an overview of the distribution of the LL variant in word list and interview styles, across these three significant predictors, beginning with age (Figure 7.1). Since the application frequency of LL variant by age group in word list and interview styles are reported in Chapter 5 and Chapter 6 respectively, here I only give the figure to illustrate the application frequencies in two styles. Hereafter the application frequencies by social group and gender will not be repeated, but only figures are provided to demonstrate stylistic variations.

Figure 7.1 Application rates of LL variant in word list and interview styles by age group

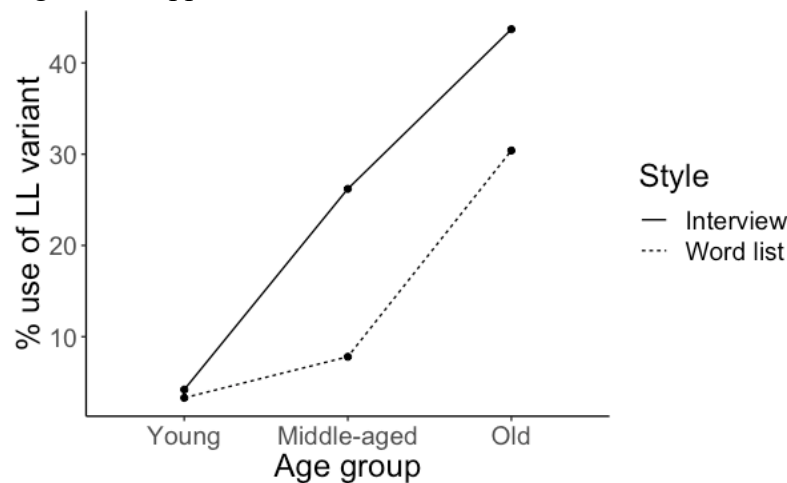


Figure 7.1 shows that the traditional LL variant is avoided in the more formal word list style by both the Old and Middle age groups. For the Young group, rates of (53-53) application are so low that style-shifting is irrelevant. But note that there is an increase in the degree of shift from the Old to the Middle generation: The widening gap between styles with decreasing age suggests that negative social evaluation of the traditional variant increased over the generations, at the same time that it was decreasing in frequency. The avoidance of LL in word list style suggests speakers are aware (at some level) of the differences between the two variants, and therefore that (53-53) is at least somewhat above public consciousness.

The distribution of the LL variant by social class and style is given in Figure 7.2, and by gender and style in Figure 7.3.

Figure 7.2 Application rates of LL variant by social class and style

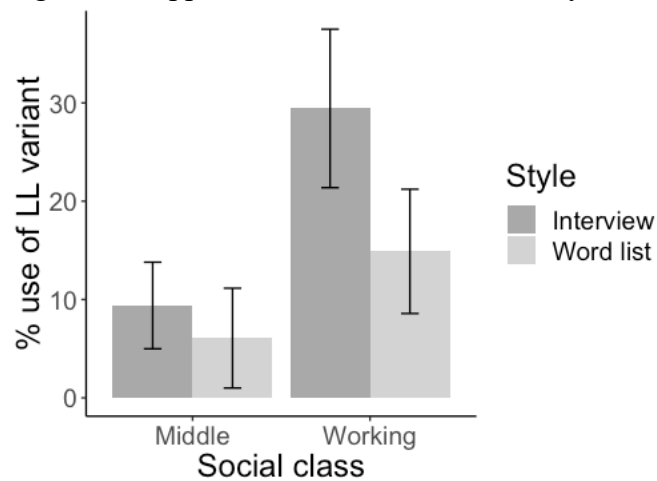
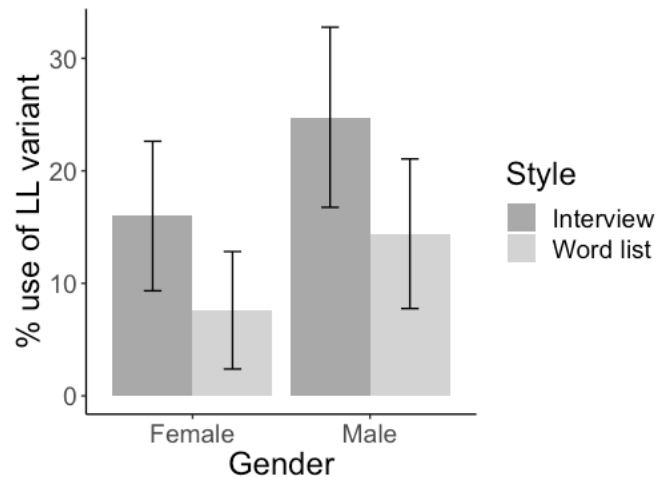


Figure 7.3 Application rates of LL variant by gender and style



Figures 7.2-7.3 show that the application frequencies of LL variant, the local and non-prestigious variant, are higher in interview style than in word list style not only in each social class, but also in each gender. The decrease for each social class is almost 50%, which is a very steep style shift. This tells us that (53-53) must be very socially salient. Notice too the very steep decrease for women of more than 50% from interview to word list style. They seem to be more sensitive to the social value of (53-53) than men do, which again is what we would expect for a socially salient variable with a stigmatized variant. The big drop of the application frequency from

interview to word list style by age, social class and gender suggest that (53-53) has been evaluated negatively by the whole Tianjin speech community for some time.

Altogether, the social and stylistic distribution of (53-53) supports the hypothesis that the ongoing disappearance of the (53-53) rule is a case of “change from above” (Labov 1966). Changes from above typically also originate outside of the speech community, which is presumably the case here, where the Tianjin tone sandhi pattern is being replaced by the SC non-application of the pattern. Indeed, the present study confirms Gao and Lu’s (2003) report that (53-53) is disappearing from Tianjin, with the young and those of higher socioeconomic status leading the disappearance.

We have now determined that (53-53) appears to be undergoing change from above, suggesting its application variant is unfavored and highly salient in the whole speech community. In the next section, I consider some of the possible social and linguistic motivations for this decrease.

7.4.2 LL: a stereotype avoided in both tone sandhi and lexical tone

The (53-53) sandhi rule is a stereotype because its application variant LL carries the publicly mentioned ‘low’ tone: the most salient feature of Tianjin dialect. The brief language attitude investigation at the end of every interview indicates that LL, the application variant of (53-53), is a non-prestigious stereotype of Tianjin speech. Since LL (T1) is the application variant of (53-53), Tianjin speakers do not like the connotations like ‘rude’ and ‘less educated’ that are connected with Tianjin dialect (previously mentioned in Chapter 3), so they try to avoid the bad connotation which is related to the stereotyped LL, the application and Tianjin variant of (53-53). This is perhaps the main reason why (53-53) is becoming obsolete. The change in lexical T1 or LL seems to support that T1 or LL is avoided.

The starting pitch of LL as a lexical tone has had a consistent tendency to rise, as shown in Chapter 3 and restated here in Table 7.1, supporting that as a stereotyped feature of Tianjin dialect, the low T1 or LL is avoided by the whole Tianjin speech community, no matter as a sandhi tone or lexical tone.

Since the starting point of T1 has changed from 11/21 to 41 among young speakers, it is also possible that (53-53) applies but the output T1 is a falling tone, which is similar to T4. In other words, T4+T4 applies but it seems to be unapplied because the starting point of T1 is high (41), which is very close to the starting point of T4 (53). X. Wang (2015) checked T4+T4 (53-53) and T1+T4 patterns. Since T1+T4 has never been reported to have undergone any tone sandhi changes, the first syllable in the T1+T4 pattern is supposed to be T1. According to X. Wang (2015), the raised T1 is still much lower than T4, so the first syllable is T4+T4 is T4.

But X. Wang (2015) is based on acoustic analysis of production data. It may be possible the difference between T4 and raised T1 cannot be perceived because of the discrepancy between production and perception. What I want to emphasize here is that it does not matter to the current dissertation whether the falling output of (53-53) is the applied variant (raised T1) or the unapplied variant (T4) because they both result in a falling output: the stereotyped LL variant is avoided.

Table 7.1 Tonal values of Tianjin lexical tones in previous and present study

Lexical tones Studies	T1	T2	T3	T4
Li (1956)	11	55	214	51
Luo & Wang (1956)	11	55	24	42
Guo (1981)	21	45	213	53
D.S.Davison (1982)	21、 11	45	24	53
Li & Liu (1985)	21	34	213	53
Shi (1987 & 1988)	21	45	213	53
Zhang & Liu (2011)	41	34	12	52
X. Wang & Lin (2016)	41	45	214	53
The present study (Young group)	41	45	214	53
The present study (Old group)	31	45	214	53

The last two rows of Table 7.1 provide the results of an acoustic analysis of a subsample of 16 Young speakers (8m, 8f) and 16 Old speakers (8m, 8f) selected randomly from database of the present dissertation. They indicate that the tonal values of T1 are different in the Young and Old groups, while the other three lexical tones have not exhibited an age group difference.

The acoustic analysis and normalization methods were introduced in Chapter 4, so I will not repeat them here. The normalized results of both young and old groups are given in Figure 7.4.

Figure 7.4 The normalized F0 results of T1 in young and old groups

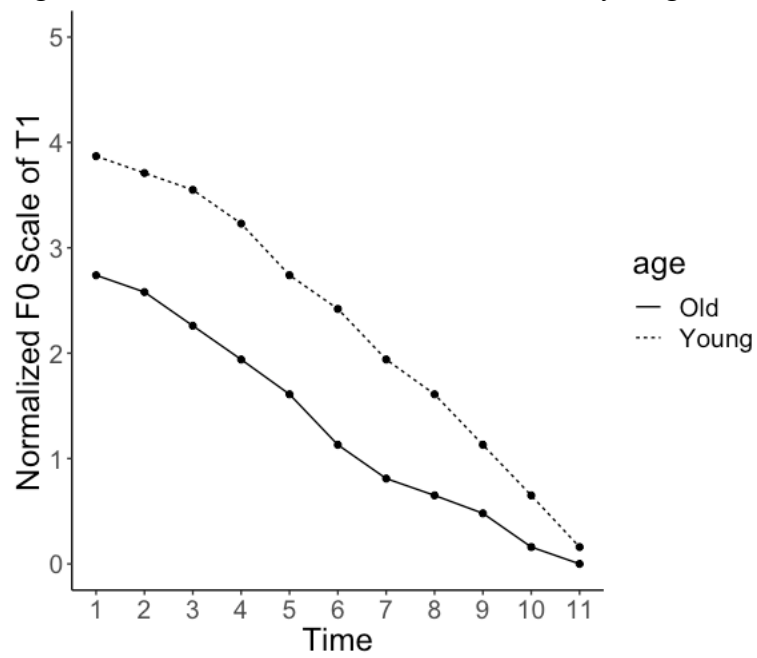


Figure 7.5 The normalized F0 results in young and old groups

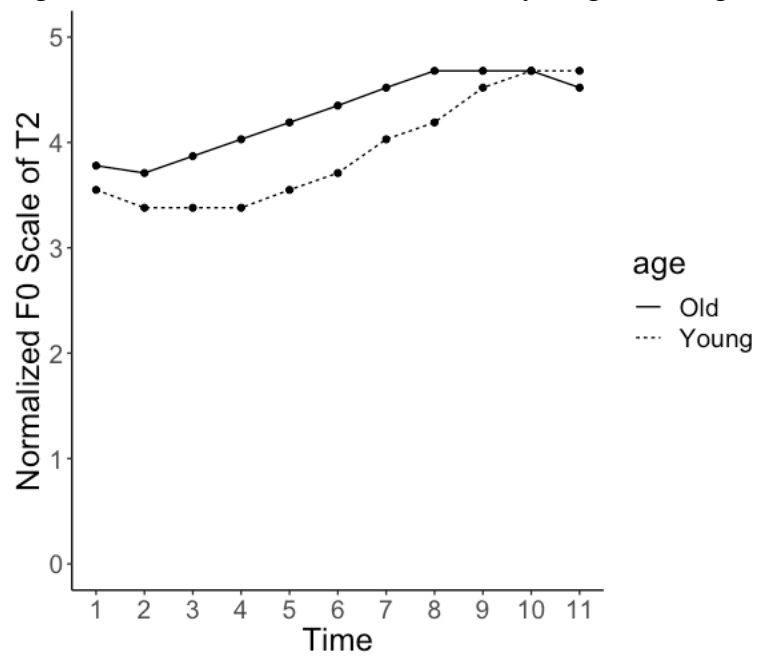


Figure 7.6 The normalized F0 results of T3 in young and old groups

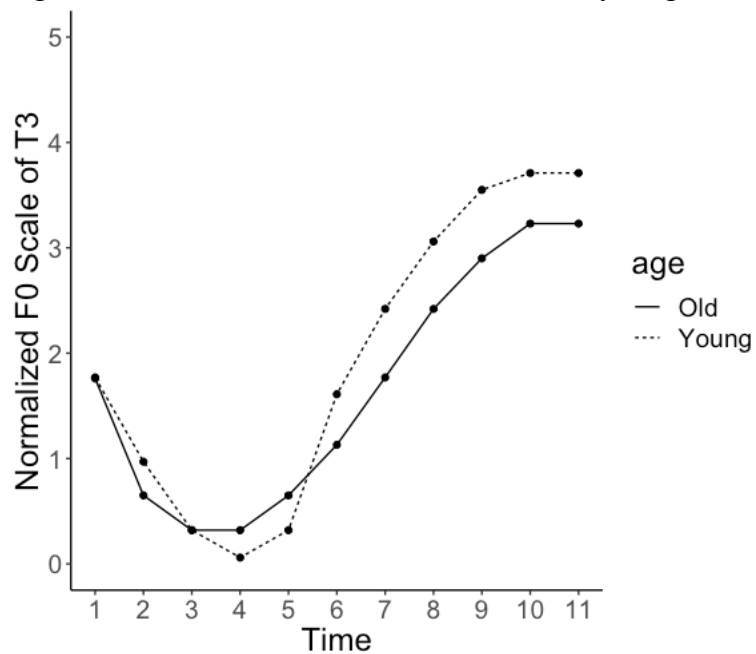
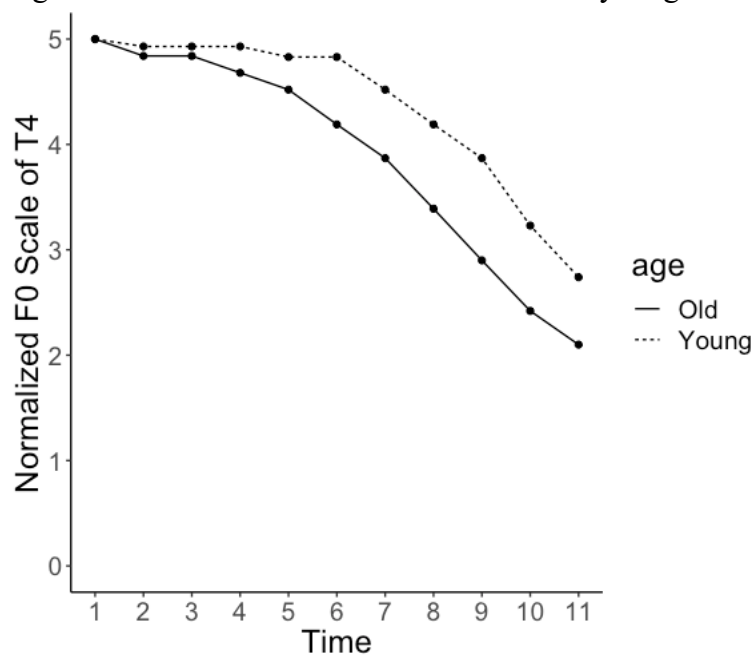


Figure 7.7 The normalized F0 results of T4 in young and old groups



Previous dialectal and acoustic studies indicate that the starting point of T1 is on the rise steadily, which is confirmed by the acoustic analyses of the present dissertation, indicating that T1 or LL, not only as the application output of (53-53), but also as a lexical tone, is avoided by the whole Tianjin speech community.

The change in lexical T1 supports the view that the decrease in (53-53) is caused by its stigmatized LL variant.

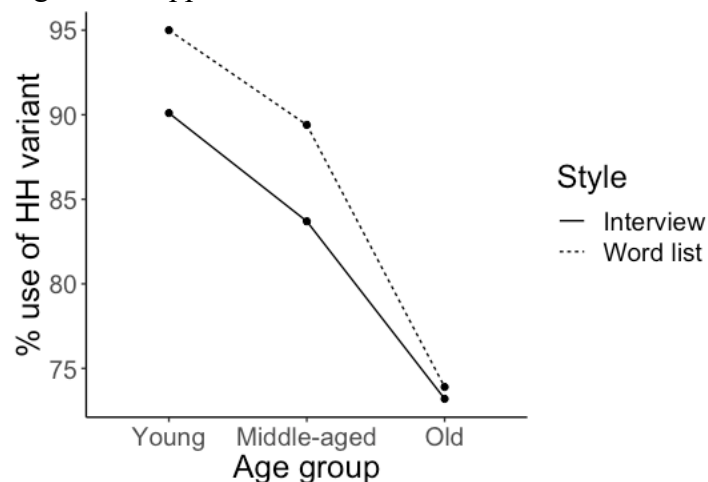
7.5 *The change in (53-21)*

7.5.1 (53-21): an increase from below

As shown in Chapter 5 and Chapter 6, age group and attitude to migrants significantly predict the application of (53-21) in both word list and interview styles, while social class is significant in interview style alone. We saw that application of the traditional rule with its HH variant is used by younger people more than older people, by people with negative attitude to migrants more than people with positive attitude to migrants and by the working class more than the middle class. The social class variation suggests that the local HH variant might be a change from below, which is further confirmed by the style analysis shown in this section.

In this section, I provide an overview of the distribution of the local HH variant in word list and interview styles, across age (Figure 7.8) and social class (Figure 7.9). The other significant predictor, attitude to migrants, will be discussed in the next section.

Figure 7.8 Application rates of HH variant in word list and interview styles by age group



As shown in Figure 7.8, the application rates of the local HH variant are higher in word list style than in interview style in every age group, although this difference is much smaller than that for (53-21), indicating it is barely salient in the speech community.

The stylistic variation suggests that the change in (53-21) tone sandhi might be a change from below, because the less prestigious Tianjin variant HH seems to be more frequent in word list style than in interview style. It is clear that speakers do not actively avoid the Tianjin variant of (53-21) in their most formal (i.e. word list) style. If they were aware of the change toward the local variant and considered the local (53-21) sandhi pattern to be undesirable, the frequency of the more prestigious non-application variant would be higher in the word list reading style than in interview. This is not the case.

Figure 7.9 Application rates of HH variant by social class and style

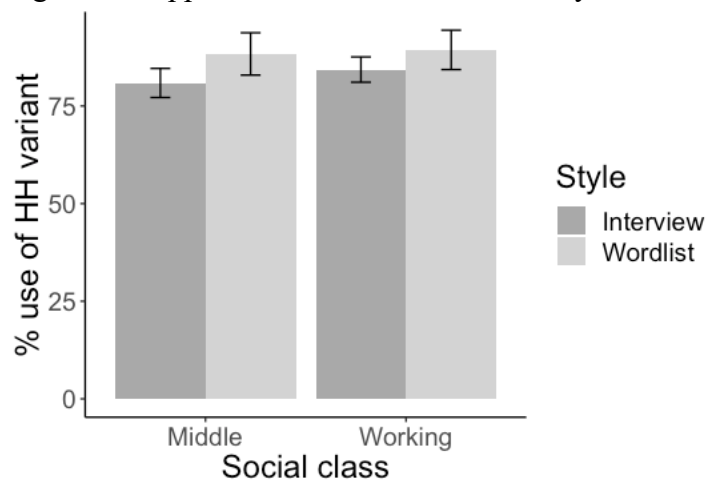


Figure 7.9 above shows that the application frequencies of HH variant, the local and non-prestigious variant, are higher in word list style than in interview style in every social class. But the stylistic difference is very small in every social class, suggesting its social salience is very low.

Altogether, the social and stylistic distribution of (53-21) supports the hypothesis that the ongoing increase of the (53-21) rule by application frequency is a case of “change from below” (Labov, 1966). Changes from below typically originate from within the speech community. That

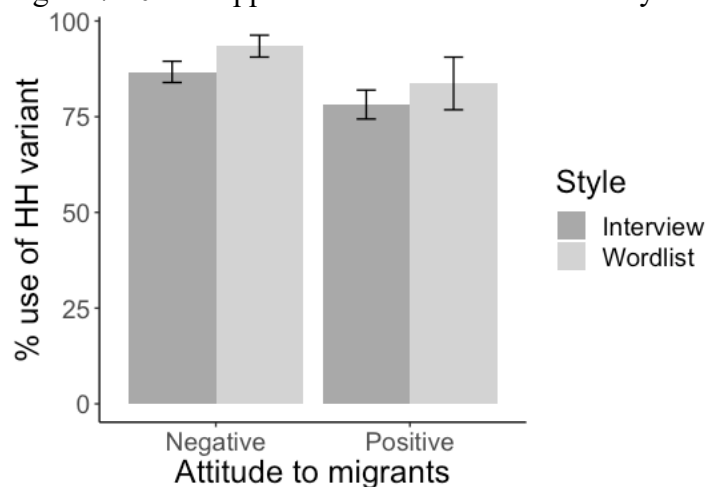
is presumably the case here, where the winning variant of the (53-21) is HH, evolved from within the Tianjin speech community. Although stylistic variations of changes from below are different at “indicator” stage and “marker” stage (Labov, 1966), one common point of changes from below is that more prestigious variant does not appear more frequently in more careful speech, which is exactly the case here.

We have now determined that (53-21) appears to be undergoing change from below. In the next section, I consider some of the possible social and linguistic motivations for this change.

7.5.2 (53-21): an increase in local variant led by people with negative attitude to migrants

As shown in Chapter 5 and Chapter 6, attitude to migrants is a significant predictor of (53-21) in both word list and interview styles. People with negative attitude to migrants are more likely to use the local variant than people with positive attitude. The distribution of the HH variant by attitude to migrants and style is given in Figure 7.10.

Figure 7.10 The application rates of HH variant by attitude to migrants and style



As shown in Figure 7.10, no matter whether people with negative attitude to migrants or positive attitude, their application frequencies of the local variant are lower in interview style than in word list style, and again the stylistic difference is small, which confirms the general trend observed in subsection 7.5.1. An interesting point is that the stylistic difference in negative group

is bigger than the stylistic difference in positive group, suggesting people with negative attitude to migrants are more conscious of the variable than those with positive attitude.

Although multivariate analysis could show significance of a predictor, the hidden interactions of the predictor and other factors such as age, gender and attitudes are irretrievably lost (Labov, 1990:221). So here cross-tabulations are given to show the interaction of attitudes to migrants and other significant predictors of (53-21). The distribution of (53-21) by attitude to migrants and age group in both styles is given in Table 7.2.

Table 7.2 The distribution of HH variant by attitude to migrants and age group in both styles

Age group	Attitude to migrants	Word list		Interview	
		App/Tokens	%	App/Tokens	%
Young	Negative	562/580	96.9	1310/1452	90.2
	Positive	236/260	90.8	441/491	89.8
Middle	Negative	165/180	91.7	371/417	89
	Positive	282/320	88.1	559/694	80.5
Old	Negative	167/199	83.9	349/455	76.7
	Positive	183/261	70.1	258/374	69
Total	Negative	894/959	93.2	2030/2324	87.3
	Positive	701/841	83.4	1258/1559	80.7

Table 7.2 shows that people with negative attitudes to migrants have higher application frequency of (53-21) in both styles, and the tendency is consistent in every subgroup by age and attitude to migrants. It seems that the change towards the application variant of (53-21) is led by people with negative attitudes to migrants, which supports the proposal that Tianjin natives want to mark the difference from migrants by ‘recycling’ the non-stigmatized local features.

As shown in Chapter 4, working class speakers are more likely to have negative attitude to migrants than middle class are. For this reason, I wanted to check the interaction between social class and attitude to migrants. The distribution of (53-21) by attitude to migrants and social class is given in Table 7.3.

Table 7.3 The distribution of (53-21) by attitude to migrants and social class in both styles

Social class	Attitude to migrants	Word list		Interview	
		App/Tokens	%	App/Tokens	%
Middle	Negative	397/439	90.4	950/1124	84.5
	Positive	331/401	82.5	539/681	79.1
Working	Negative	497/520	95.6	1080/1200	90
	Positive	370/440	84.1	719/878	81.9
Total	Negative	894/959	93.2	2030/2324	87.3
	Positive	701/841	83.4	1258/1559	80.7

As shown in Table 7.3, stylistic differences exist in every subgroup by social class and attitude to migrants, with the local variant appears more frequently in word list style than in interview style, but the style difference is small, compared to the style difference of (53-53). This style distribution of (53-21) by attitude to migrant and social class not only echoes the tendency shown by attitude to migrant and age, but also the trend observed in subsection 7.5.1.

7.5.3 (53-21): recycling of an unmarked local variant to mark Tianjin identity

As we have seen, the Tianjin variant of (53-21) tone sandhi was reported to occur only at a low frequency in the late 20th century (Shi, 1988), but later studies, including the present one, show that it now occurs almost categorically in the speech of young and middle-aged speakers (Shi & P. Wang, 2004). It is natural to ask what has caused this reversal in the application

frequency of (53-21), particularly since Tianjin people never mention the change. Is this change similar to the language change reversals that exhibit a v-shape age distribution in apparent time observed in the well-known Martha's Vineyard study (Labov, 1972) and in the more recent Cajun community study by Dubois and Horvath (2000)?

Labov observed that the characteristic centralization of (ay) and (aw) (/ay/ and /aw/) on Martha's Vineyard was returning after a period of decline. He suggested that the increase in centralization was closely related to people's attitudes towards the island. Because the fishing industry declined, the Vineyarders depended increasingly on summer tourism. But on the other hand, they viewed the summer trade as a threat to their island identity. Labov argued that centralization, subconsciously, means 'Vineyarder'. People who considered themselves to be belonging to the island tended to use centralization more than people who did not. The centralization of (ay) and (aw) marks the speaker's island identity, and therefore can be used to resist the challenges to their native status. The rise in centralization occurred because the challenges had become much sharper in the past two generations.

Similarly, in a study of three generations of ethnic Cajuns in a rural area of Louisiana, Dubois and Horvath (2000) reported on a 'recycling' of French-derived features in the Cajun variety of English. The older generation used many French-derived Cajun features in their English, and the middle-aged generation dramatically decreased their use of all of these features. But some of the Cajun features revived among the young generation. The fall and rise of the Cajun features are all related to the socioeconomic development of the area. Industrialization and urbanization brought some fundamental changes in southern Louisiana. Getting a better job meant speaking English. In addition to this economic reason, the middle-aged speakers were aware of the stigma attached to Cajun English. In contrast, the young speakers were influenced by the so-called "Cajun

Renaissance”, and they were proud to be Cajuns. Besides, they were able to profit from the important economic benefits of display of Cajun culture associated with tourism. Dubois and Horvath argue that the Cajun features increased among younger Cajun speakers because the use of Cajun variants has become a sociolinguistic marker of Cajun identity, just as the use of traditional Martha’s Vineyard variants marked island identity.

The rise in (53-21) also conforms to Ferrer and Sankoff (2003), which examines the factors affecting language choice using sociodemographic, behavioral, and ideological data on a sample of 180 high school students in the city of Valencia, Spain and the neighboring town of Xativa. They declare that the speakers’ feeling of identity, their sense of being members of the group to which a particular way of speaking is thought to belong, determines the way they choose to speak. Using a semimatched guise technique, Ferrer and D. Sankoff (2003) argue that identity is a more important factor behind language choice than the prestige of the varieties. Although SC is a more prestigious language, it cannot mark Tianjin natives’ identity. When Tianjin natives face identity or economic pressures from migrants, they especially want to mark their Tianjin identity and to mark them as a different group.

The reason why (53-21) revived in 1990s instead of other times is proposed to be because Tianjin people, just like Vineyarders, felt the threat to their local features and culture when a huge number of migrants rushed into the city after the “Opening-up and Reform” policy of 1978. Crucially, this in-migration would have coincided with the Tianjin residents born after 1950 entering their 20s and 30s. These Tianjin speakers would have been competing with migrants for jobs. More generally, Tianjin people felt the pressure of keeping their identity more fiercely than any other time.

This is illustrated by comments made by some participants. Participant (#12), a young male undergraduate in Tianjin Normal University, talked about how Tianjin was going to keep its culture and tradition and expressed his worries. He said Tianjin was different from Beijing, which has profound culture background. Although there were much more migrants in Beijing, Beijing people did not have to worry about keeping their culture and identity because Beijing was too rich in its cultural deposits to worry about them being lost. Tianjin was not so rich and profound in its culture, and it was difficult to keep Tianjin identity and culture with so much in-migration, together with their cultures. What he said represents some Tianjin young people's worries, although not all of them expressed this so clearly.

Young speakers like participant (#12) are concerned with losing their identity, while middle-aged speakers have more practical issues to worry about. Participant (#79) had been a state-owned factory worker but lost her job during the reform of state-owned factories. She said that all workers in that factory were fired because the factory was bankrupt. She became a street vendor, selling lunch to students. At first, she could make some money, but later she had to end her small, marginal business because of the cut-throat competition from migrants. She complained that the migrants would rather stand a loss to undercut the price, and finally their competition had her on the ropes. Then she had to work for a relative who ran a small hardware store at a hardware shopping center, but the former migrant employee of the relative started his hardware store at the same shopping center and sold the same commodities at a price always a little bit lower than theirs. Finally, her relative had to end the small store and she became jobless again. She said bitterly: "You should go to the hardware shopping center to see by yourself. All the stores there are owned by migrants now. Tianjin natives were all elbowed out of the market". The reason she gave why migrants could win the competition is that they can give the price so low that Tianjin natives could

not afford to give. In sum, middle-aged people are breadwinners and some working class people, especially street vendors like participant (#79), have to face severe competitions from migrants directly, which affects their attitudes to migrants.

But why was (53-21) selected as a new marker of Tianjin identity and not one of the other non-salient Tianjin tone sandhi patterns? This is probably related to the application variant. The application variant of (53-21) is Tone 2, which is very similar to SC. Tone 2 in Tianjin is 45 (HH), while T2 in SC is 35 (HH). Both are high rising tone, and most Tianjin speakers are unaware of their differences. Since there is no FL tone sandhi in SC, this tone sandhi can mark their Tianjin identity. At the same time, the application variant is not a salient feature of Tianjin dialect, as evidenced both by the low degree of style-shifting for the (53-21) variable and by the lack of metalinguistic commentary about it. Therefore it does not carry the bad connotations like the LL variant of (53-53) does. So Tianjin speakers may be employing (53-21) to mark their Tianjin identity, but doing so in a way that is not markedly linguistically different from the standard variety. This is something of a ‘having their cake and eating it’ scenario, whereby Tianjin speakers can distance themselves from sounding highly local and traditional, while actually preserving a traditional feature. In this sense, the revival of (53-21) among young Tianjin speakers is somewhat like the recycling of old features observed in Martha’s Vineyard and the Louisiana Cajun community. The young members of the communities want to keep their identities.

7.6 The change in (21-21)

7.6.1 (21-21): a change from above led by middle class speakers

As shown in Chapter 5 and Chapter 6, age group significantly predicts the application of (21-21) in both word list and interview, and attitude to Tianjin dialect and social class are significant predictors in word list style and interview style respectively. We saw that application

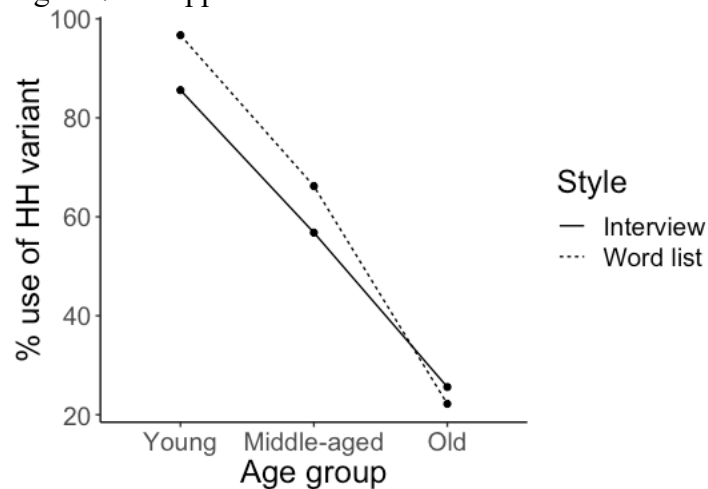
of the traditional LH variant is used by older people more than younger people, by people with positive attitude to Tianjin dialect more than people with negative attitude, and by the working class more than the middle class. All of this suggests that the old local variant LH is being replaced by a new local change from above. Although social class has an association with conscious awareness, the relation is not predictable since changes from above (Labov 1966) refer to changes above speakers' consciousness, not from higher social class. So inter-speaker variations alone cannot confirm whether a change is from above. As a result, stylistic variations of (21-21) tone sandhi are studied to confirm whether Tianjin speakers are conscious of the change.

In this section, I provide an overview of the distribution of the LH variant in word list and interview styles, across these three predictors, beginning with age.

Figure 7.11 Application rates of the old LH variant in word list and interview styles



Figure 7.12 Application rates of the new HH variant in word list and interview styles



The new HH variant is more like SC, and therefore it is the prestigious variant. The old LH variant is more local, and therefore it is the less prestigious variant. As shown in Figure 7.12, the new HH variant is more likely to appear in word list style than in interview style in the Young and Middle-aged groups, suggesting it is a change from above, though its variation pattern is reversed in the Old group, with HH variant appeared more frequently in interview style. Recall in Chapter 5 and Chapter 6, the overall application rates of the new local HH variant are higher, at 69.3% (1123/1621), in word list style than in interview style, at 63.6% (1589/2497). It seems that the young and middle-aged speakers have evaluated the old variant negatively for some time, but the Old group seem to be different.

The change towards HH, the new variant, seems to be very rapid. The middle aged and young people have utterly abandoned the old variant in favor of the new one. Furthermore, they did so without a preceding period of stigmatization of the old variant. There's no period in which, say, old people or middle-aged people exhibit a high degree of style shifting for (21-21), as we saw for (53-53). And once the new HH variant was adopted, it's used in both styles at very similar rates.

Yet there's a clear progression in these three generations from LH to HH. Old people strongly favor LH in both styles; Middle people have more of a 60-40 distribution that is also similar in both styles; Young people very strongly favor HH in both styles. The rapid change pattern of (21-21) might be related to phonological reasons, which will be introduced in subsection 7.6.2.

The distribution of (21-21) tone sandhi by social class and style, given in Figures 7.13 and 7.14, and by attitude to Tianjin dialect and style in Figure 7.15, seems to support that (21-21) is a change from above.

Figure 7.13 Application rates of LH variant by social class and style

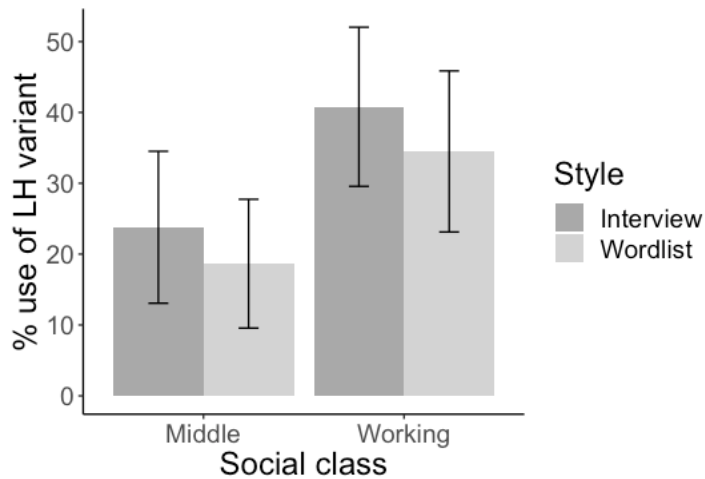
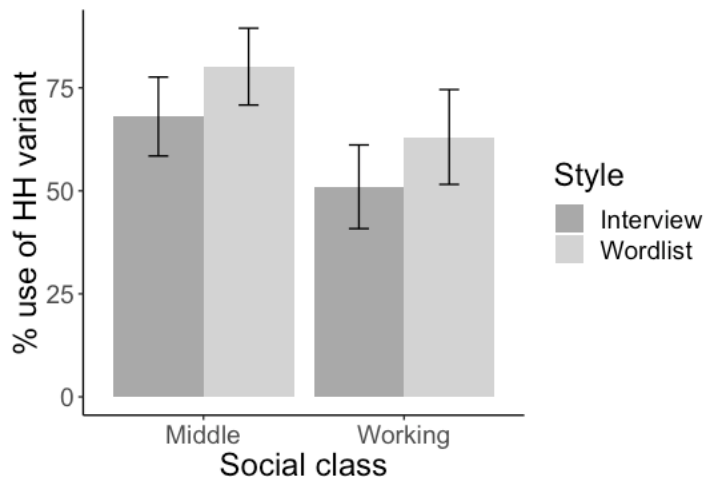


Figure 7.14 Application rates of HH variant by social class and style



As shown in Figure 7.14, the new and prestigious variant HH appears more frequently in the middle class, indicating it is a change from above. Figures 7.13 and 7.14 also show that the new variant appears more frequently in wordlist style, while the old variant appears more frequently in interview style in both social classes. The degrees of style difference in both social class are bigger than the style difference in (53-21), but smaller than the degree in (53-53), indicating (21-21) is more socially salient than (53-21), but less socially salient than (53-53).

Figure 7.15 Application rates of LH variant by attitude to Tianjin and style

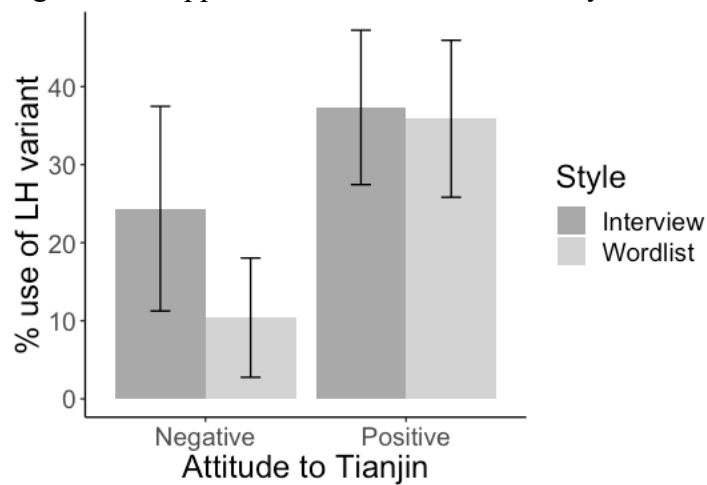
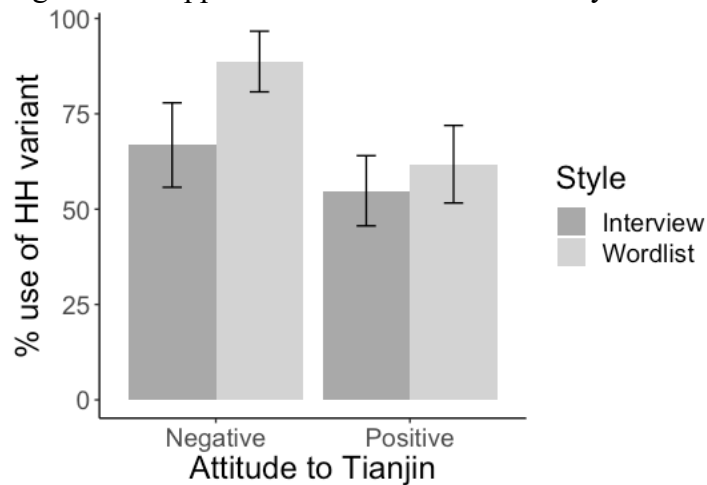


Figure 7.16 Application rates of HH variant by attitude to Tianjin and style



Figures 7.15-7.16 show that the new variant HH appears more frequently in word list style than in interview style, while the old LH appears more frequently in interview style than in word list style. The degree of style difference is much bigger in the negative group than in the positive group, suggesting that people with negative attitude to Tianjin dialect are more conscious of the change.

Figures 7.11-7.12 at the beginning of this subsection show that the stylistic variation in the Old group differs from that in the Young and Middle-age groups, and the seemingly contradictory stylistic variation in the Old group reflects the change in the Old group is in progress. Table 7.2 is the distribution of (21-21) sandhi rule by age group, social class and style, indicating the stylistic variation among old middle class speakers is consistent with that in the Young and Middle-aged groups. Only the distribution of HH variant is reported here since the change is towards the new and more standard variant.

Table 7.4 The distribution of HH variant by age group, social class and style

Style Age group	Wordlist				Interview			
	Middle class		Working class		Middle class		Working class	
	App/N	%	App/N	%	App/N	%	App/N	%
Young	386/396	97.5	347/361	96.1	538/617	87.2	488/582	83.8
Middle	139/180	77.2	159/270	58.9	233/326	71.5	187/414	45.2
Old	64/180	35.6	28/234	12	83/274	30.3	60/284	21.1
Total	589/756	77.91	534/865	61.73	854/1217	70.2	735/1280	57.4

Although Table 7.2 shows that the application frequency of the new HH variant in interview style is lower than it in word list style in the Old group, which is different from the stylistic variation in the Young and Middle-aged groups, Table 7.4 here suggests that stylistic

variations of old middle class speakers are consistent with Young and Middle-aged groups. It seems while the change in the Old group is not complete, the Middle class led the change towards the new variant. The more prestigious new variant HH appears more frequently in word list style than in interview style in all subgroups by age and social class except old working class, indicating only old working class speakers did not participate in the change towards the new variant, or/and that they do not evaluate (21-21)'s prestige difference between the new and old variant in the same way as other subgroups.

Both (53-53) and (21-21) appear to be changes from above, so why is the stylistic variation consistent in every age group in (53-53) but not so in (21-21)? Although both changes are from above, (53-53) is a stereotype while (21-21) is a marker. Different from (53-53), people do not talk about (21-21) openly. Participants talked about the T1 and the low feature of Tianjin dialect at the end of every interview, but they did not talk about T3 openly. The less prestigious variant LH, is not so stereotyped as a lexical tone like LL is. It is a less prestigious because it is not common to have a T3 in the first syllable of a disyllabic sequence. But it is quite common to have a T2 in the first syllable of a disyllabic sequence in both SC and Tianjin. Some participants mentioned that Tianjin dialect is 'lengthened' or 'too heavy', and therefore it is not pleasant to hear it. Different from LL variant of (53-53), none of the participants directly related this "heavy" or 'long' to LH variant of (21-21), but it requires more effort and takes longer duration to pronounce LH variant than HH variant, so in this sense, LH variant is lengthened or heavy but HH variant not.

Tianjin hua wo jue de jiushi youdian la chang de ganjue. Shengdiao di, lachangyin.

'I feel Tianjin dialect is lengthened. Its tone is low and lengthened'.

(Quotes from #27, a young female working class speaker)

Besides, previous studies on Tianjin have not mentioned T3 or LH variant as a salient feature of Tianjin dialect, as T1 or LL did. Although Tianjin speakers are conscious of (53-53) and (21-21) changes, the conscious degrees of the two variables are different.

Besides, it looks as if the loss of (53-53) is a change that began before the current Old generation, but (21-21) seems to originate in the Middle-class Old generation. (53-53) is a stereotype because the traditional variant has been disappearing for a long time. (21-21) is only a marker because the disappearance of the old variant has not been going on for very long. The change in (21-21) is a result of the rising of the starting point of T1, and T1 changes maybe together with the decrease of (53-53). It has proceeded astoundingly rapidly probably because it is a change caused by the rising in the starting point of T1 and the change in T1 may have proceeded for a long time.

In the next section, I consider some of the possible social and linguistic motivations for this change. Different from (53-21) and (53-53), the change in (21-21) is not the increase or decrease of a local variant, but an old local variant is replaced with a new local variant. I propose that the emergence of the new local variant is triggered by a combination of language-internal and social factors.

7.6.2 Linguistic motivations for the emergence of the new variant

The linguistic motivations for the emergence of the new variant include: 1. metrical structure, tonal complexity and tonal stability; 2. the rising of the starting point in T1.

7.6.2.1 Metrical structure, tonal complexity and tonal stability

One main linguistic motivation for the replacement of T3(LH) with T2(HH) in the output seems to be a combined effect of metrical structure, tonal complexity and tonal stability. Previous phonological studies (Wright, 1983; Duanmu, 1995; Chan, 1985; Wang & Lin, 2017; among others) suggest that stressed syllables tend to keep their original tones while unstressed syllables have a

tendency to lose their original tones. This tendency exists in Tianjin too (Wang & Lin, 2017). Tianjin disyllabic sequences form right-headed feet (J. Wang, 2002; Wee, 2015), which suggests that the second syllable is stronger or more stressed than the first syllable. The fact that it is always the first syllable that changes its underlying tone in Tianjin---unless the second syllable is a neutral tone---supports the claim that unstressed syllables tend to lose their underlying tones in Tianjin (Wang & Lin, 2017).

Zhang (2002) studies the correlation between duration and contour-bearing ability and indicates that the greater the number of pitch targets of a tone, the longer duration it requires. Yip (2002: 27-30) expresses a similar relation between tonal complexity and duration or stress. T3 (214) is more complex than T1 (31/41), T2 (45) and T4 (53) since T3 is a dip-rising tone, which has three pitch targets. As a result, it requires longer duration than T1, T2 and T4, which are simple falling or rising tones and require only two pitch targets. In Tianjin, duration is the most important acoustic correlate of stress (Lu & J. Wang, 2012). It seems that the old variant LH, as a complex tone with three pitch targets, requires longer duration to realize and therefore it should be on stressed syllables.

The view that T3 requires longer duration is supported by the behavior of T3 in Tianjin disyllabic sequences. The first tones in T3+T1, T3+T2, T3+T3 and T3+T4 are all reported to be changed (Wee, 2004; Zhang & Liu, 2011; Li & Chen, 2016; X. Wang, 2016), indicating T3 changes in all non-final positions unless the second syllable is a neutral tone. This further indicates that T3 cannot keep its underlying tone unless it is on the prosodic head. This is consistent with the general correlation between duration, stress and contour-bearing ability in Zhang (2002) that contour tones occur more freely in stressed syllables than unstressed syllables, with no language display the opposite pattern.

In summary, T3 is a complex tone, and therefore it requires longer duration to realize. Based on the iambic foot pattern of Tianjin disyllables (J. Wang, 2002; Wee, 2015), the first syllable is not the prosodic head, and therefore shorter than the second. So complex tone cannot keep its underlying tone when it is on the first syllable. Even the underlying T3 tends to lose its original tone when it is on the first syllable, it is not hard to understand that the derived T3 can only be 13/14, but not 214 when it is not on the prosodic head.

Then why is T3 changed to T2 instead of keeping its 13/14 variant? The reason seems to be related to the rising of the starting point of T1, which will be presented in the next section.

7.6.2.2 The rising in T1 and the emergence of the new variant

As shown in §7.4.2, T1 is a stereotyped tone, which is avoided by the whole Tianjin speech community. The avoidance is reflected in the disappearance of (53-53) sandhi rule, and also reflected in the rising of the starting point of T1, reported in §7.4.2. The tonal value of T1 has changed from 21 to 41 among young speakers, but it is 31 among old speakers.

The change from the old variant LH to the new variant HH is a result of the rising of the starting point of T1. When T1 was LL (21), the starting point of the first syllable is low. Adjacent identical tones are forbidden because of Obligatory Contour Principle (Leben, 1973; Goldsmith, 1976; McCarthy, 1986; Yip, 1988; hereafter OCP). The traditional disyllabic tone sandhi patterns in Tianjin are triggered by OCP (Chen, 2000; J. Wang, 2002; Yip, 2002), either to avoid two adjacent identical tonal units or tonal components. For example, in T4 (HL) + T4 (HL) combination, HL and HL are two adjacent identical tonal units, which violates OCP, so the first HL changes to HH to avoid adjacent identical tonal units. In T4 (HL) + T1(LL) combination, L.LL are adjacent identical tonal components.

To avoid or remedy adjacent identical tonal units, the first tone in T1+T1 undergoes tone sandhi (Yip, 2002), as shown in (2) below. Full syllable in Tianjin is considered to be bimoriac (J. Wang, 2002).

(2) The derivation of the old variant LH

$T1(LL)+T1(LL) \rightarrow T3(LH) + T1(LL)$ (Old variant)

As shown in (2) above, the second tonal feature of the first syllable changes to H feature to avoid adjacent low tonal features. The output is LH because the first tonal feature of the first syllable is low. When T1 was 41 among young speakers, the first tonal feature of the first syllable is no longer low, as shown in (3).

(3) The derivation of the new variant HH

$T1(HL)+T1(LL) \rightarrow T2(HH)+T1(LL)$ (New variant)

As shown in (3) above, the second tonal feature of the first syllable changes to H feature to avoid adjacent low tonal features. The output becomes HH as the first tonal feature of the first syllable is high.

The age difference in (21-21) output further supports the hypothesis that the rising of the starting point of T1 might be related to the appearance of the new output. As for the current acoustic study reported in §7.4.2, T1 is 31 among old speakers, but it is 41 among young speakers. When T1 is 31, as in (2), the first tonal feature is still a low feature, which explains why old speakers tend to use the old variant LH. When T1 is 41, as in (3), the first tonal feature is not low, so young speakers tend to use the new variant HH.

The rising of the starting point of T1 seems to be linked with the appearance of the new output, but the rising of the starting point maybe triggered by sociolinguistic factors. Since the low

feature of T1 is the stereotyped feature in Tianjin dialect, the rising of its starting point makes T1 less salient than it originally was.

The linguistic motivations for the emergence of HH variant has been existing, but why did HH variant emerge in the 1990s instead of other times? The emergence of the new variant in the 1990s is caused by social factors which will be discussed in the next section.

7.6.3 Attitude to Tianjin dialect in the emergence of the new variant

I speculate that the emergence of the new variant in the 1990s is related to migration and attitudes to Tianjin dialect. The emergence time of HH coincides with the time of mass in-migration to Tianjin in the 1990s. Although attitude to migrants is not a significant predictor of (21-21) in neither word list nor interview style, the presence of a large number of migrants may have made Tianjin speakers be more conscious of T1, the salient feature of Tianjin dialect. T1 has become a stereotype and avoided by each social class and age group in Tianjin speech community. One outcome of the avoidance is the raising of the starting point of T1. The emergence time of the new variant supports the proposal that the in-migration to the city started in 1990s triggered or accelerated the change in T1, which in turn causes the emergence of the new variant.

Previous sociolinguistic studies have shown that attitude influences language change and variation (Labov, 1972; Eckert, 2000; Kristiansen & Jørgensen, 2005; Giles & Ogay, 2007; Johnsen, 2015). Classical sociolinguistic studies such as Labov's (1972) explanation of reversed sound change on Martha's Vineyard, and Eckert's (2000) description of the sound changes represented among students at Belten High attribute sound changes to attitudes to local identity. More recent sociolinguistic studies also show the influences of attitudes on language change and variation. Johnsen (2015) studies the language used in South-East Norway, and finds that people there do not adopt new linguistic features from the upper Oslo speech community, which is often

referred to as ‘educated daily speech’, ‘Standard East Norwegian’ (Vanvik, 1972:120-121). Instead, they adopt features from the lower Oslo dialect, which is a less prestigious dialect than the upper Oslo dialect. The reason Johnsen (2015) gives is the negative views on the upper Oslo speech community prevent people in South-East Norway from adopting their language.

The mixed regression analyses of (21-21) in word list style indicate that attitude to Tianjin dialect is a significant factor affecting the application of (21-21). People with negative attitude to Tianjin dialect have higher application frequency of the new variant HH than people with positive attitude to migrants. The distribution of HH variant by age group, attitude to Tianjin dialect and style is given in Table 7.3.

Table 7.5 The distribution of HH variant by age group, attitude to Tianjin and style

Style Age group	Wordlist				Interview			
	Negative attitude		Positive attitude		Negative attitude		Positive attitude	
	App / N	%	App / N	%	App / N	%	App / N	%
Young	285/288	99	448/469	95.5	327/387	84.5	699/812	86.1
Middle	136/180	75.6	162/270	60	162/313	51.8	258/427	60.4
Old	43/90	47.8	49/324	15.1	46/139	33.1	97/419	23.2
Total	437/558	83.2	679/1063	62	535/839	63.8	1054/1658	63.6

Speakers with negative attitudes to Tianjin dialect are more likely to use the new prestigious HH variant in wordlist style (83.2%) than in interview style (63.8%). Speakers having negative attitudes to Tianjin dialect do not like the more local old variant, so they select the more standard-like new variant in their more careful wordlist style.

In word list style, people with negative attitude to Tianjin dialect are a lot more likely to use the prestigious HH variant (83.2%) than those with positive attitude to Tianjin dialect (62%).

But there is almost no difference in interview style, with HH being used at a rate of 63% regardless of their attitude to the dialect. People are more conscious of their speech in the more careful word list style than in interview style, so the significant difference in word list between people with negative attitude and positive attitude suggests that they are at least somewhat conscious of the change in (21-21) sandhi. The situation is the same with LH variant: attitudes differentiate LH use in word list style but not in interview style. Because interview style is a less careful style, speakers have less control of their speech in interview style. As a result, their speech in interview style cannot reflect their language attitudes as wordlist style does,

7.7 Summary

This chapter mainly studies the social meaning of different variants by examining stylistic differences of the three variables. The steep decrease of the application of the local variant of (53-53) from interview style to word list style indicates that speakers consciously avoid the local variant in their more careful style. The big style difference in Middle-aged and Old groups suggest that the change is not complete in these age groups, but the style difference in the Young group is minimal, suggesting the change in this age group is almost complete. The big style difference in (53-53) by social class and gender indicates the high salient of the variable.

Different from (53-53), speakers do not avoid the local variant of (53-21) in their more careful style, and the style difference in (53-21) is very small, suggesting the change is not salient and below the level of social awareness. But the change show consistent stylistic difference, with local variant appear more frequently in more careful wordlist style, suggesting it is a ‘marker’ instead of an ‘indicator’. People with negative attitude to migrants consistently have higher application frequency of the local variant than people with positive attitude to migrants, suggesting

people with negative attitude lead the revival of the local feature to keep their local identity and mark them from migrants.

The style difference in (21-21) is bigger than (53-21), but smaller than (53-53), suggesting the social salient of (21-21) is higher than (53-21), but lower than (53-53). The change is triggered by the rising of the starting point in T1, which is the stereotyped feature in Tianjin. This point is confirmed by age difference. The acoustic analysis and impressionistic coding show that the tonal values of T1 are 41 in the Young group, but the tonal values of T1 are 31 in the Old group, which explains why the output of the Young group is HH, but the output of the Old group is LH.

Chapter 8 Conclusions and future directions

This dissertation has shown that there are significant age differences in the application of the (53-53), (53-21) and (21-21) sandhi variables. Coupled with less systematic evidence from prior phonetic and dialectology studies, we can infer from the apparent time distribution that all three sandhi variables are undergoing diachronic change. The change directions are opposite between (53-53) and (53-21). (53-53) is becoming obsolete as young and middle-aged speakers almost do not have it, while its application frequency is still pretty high among older speakers. (53-21) is on the rise as young and middle-aged speakers have it almost without exception, while its application among old speakers remains low. The change in (21-21) is the replacement of the old variant LH with the new variant HH.

The change in (53-53) sandhi rule seems to be a change towards the prestigious non-applicant variant HL led by young speakers. Young middle class speakers are less likely to use the application non-prestigious variant LL than any other subgroups by age group and social class. This is in line with prior research on (53-53): Gao & Lu (2003) reported that the change in (53-53) seemed to be led by young middle class speakers. We also saw that women are more likely to not apply the (53-53) rule than men, particularly women in the middle class. This is consistent with sociolinguistic studies of changes from above, such as Labov (1966)'s study of (r) in New York City.

Stylistic variation also suggests that the change in (53-53) is a change from above. The application variant is consistently avoided in the more formal word list style by age group, social class and gender, indicating the variant has been avoided by the whole Tianjin speech community for some time. The big difference between the two styles is indicative of the high salience of the variable in the speech community. And indeed, (53-53) is definitely a stigmatized feature, since

my participants had metalinguistic knowledge of its application variant LL as the more salient feature of Tianjin dialect.

Language attitude has influences on the decrease of (53-53). Since it is a stereotype of old Tianjin identity related to the bad connotations of uneducated and rude, it is especially avoided by people with a negative attitude to Tianjin dialect. The results showing that people with positive attitude to Tianjin are more likely to use the local variant in both styles support the proposal that negative attitude to the dialect accelerated the decrease of (53-53).

Contrary to (53-53), the application local variant of (53-21) is increasing over time. The increase of (53-21) seems to be related to in-migration to the city over the past thirty years. The results that people with negative attitude to migrants are more likely to use the local variant echo this proposal: they do not like migrants, so they are more eager to mark differently from them. Local identity awakened by large-scale in-migration to the city is proposed to be the cause of the recycling of the unmarked local feature (53-21).

Stylistic variation suggests that the change in (53-21) is a change from below as the less standard variant appears more frequently in the more careful word list style. This shows that speakers do not avoid the local variant even when they are paying careful attention to their own speech. The difference in (53-21) application rates across the two styles is consistent in every age group, but the degree of difference between styles is much smaller than that in (53-53), indicating the social salience of (53-21) is not as great as for (53-53).

Changes from below may not show consistent stylistic variation in ‘indicator’ stage or the more prestigious variant appear more frequently in the less formal style in ‘marker’ stage. Here the more overtly prestigious variant appears less frequently in the more formal style consistently, suggesting it is a ‘marker’.

As introduced in subsection 3.5, Tianjin dialect has a function that SC does not have: Tianjin dialect is potentially a resource for marking residents of the six central districts from those who lived in the suburbs in the past, and for marking Tianjin locals from migrants when the presence of migrants is evident. As the lingua franca of China, SC does not have this function. If this function of Tianjin dialect is taken into consideration, the stylistic variation of (53-21) may indicate that the local variant has some covert prestige.

The fact that (53-21) is not so salient as (53-53) is perhaps why it is selected to mark Tianjin identity. (53-21) has never been stigmatized; it appears to be below public awareness. As such it is available for ‘recycling’ (Dubois & Horvath, 2000) as a positive marker (Labov, 1972) of ‘new’ Tianjin identity. By applying the (53-21) rule, natives of Tianjin may be linguistically contrasting themselves with the many migrants who have moved to the city in the last three decades. The rise of (53-21) coincides well with the history of migration to Tianjin.

The change in (21-21) is reflected in the emergence of a new local variant HH. The new local variant is preferred by young speakers, while the old local variant still exists among old speakers. The change towards the new and more SC-like variant seems to be led by young middle class speakers with a negative attitude to Tianjin. The change in (21-21) is proposed to be caused by the avoidance of the most salient feature of Tianjin T1, and the avoidance is not only reflected in the decrease of (53-53), but also the rising of the starting point of T1 (21-21) as a lexical tone, which in turn leads to the change from LH to HH. The acoustic analysis results, that the value of T1 is 41 among young speakers, but 31 among old speakers, seem to support this view.

Different from (53-53) and (53-21), the stylistic variation of (21-21) exhibits a crossover age distribution. The new variant appears more frequently in the more formal style in young and middle-aged groups, and also in the middle class of the old groups, suggesting it is a change from

above. The different stylistic variation in the old group is caused by working class. It seems the change towards the new variant has not be spread to old working class speakers.

Dialect will change in the conflicts between forces in favor of standardization and in favor of the ‘vernacular maintenance’. Standardization which “works by promoting invariance or uniformity in language structure” (Milroy J. , 2001: 531) conflicts with forces in favor of keeping language variances. Like other studies on changes of Tianjin dialect, the decrease in (53-53) and the emergence of the new variant in (21-21) reflect the process of standardization in which dialectvariances are reduced. Different from other studies on Tianjin dialect, this is the first one that actually shows the effect of local identity and relate it to the strengthen of an unmarked local variable (53-21).

The decrease of prominent stereotyped dialect features has been observed in the languages with large numbers of speakers (Hinskens, 1998). The unmarked dialect features have been observed to strengthened or revived to keep local identity (Labov,1972; Dubois & Horvath, 2000). These two conflicting forces might lead to a stable compromise dialect (Hinskens, 1998). The changes in Tianjin tone sandhi echo these earlier generalizations made from Western speech communities. While the stereotyped features like (53-53) and ‘special consonants’ are disappearing, an unmarked local feature like (53-21) is increasing to keep local identity. The change in Tianjin dialect also indicates that Tianjin dialect is changing towards a compromise form. Some Tianjin speakers have noticed the compromise change in the dialect. For example, # 72 (an old male speaker) said that young Tianjin speakers only keep some tones or tunes of Tianjin dialect, but the ‘special consonants’ and old Tianjin vocabulary have disappeared. The Tianjin tones or tunes mentioned by him might be the Tianjin tone sandhi kept among young and middle-aged speakers. It is a form of compromise to keep unmarked local features while giving up stereotyped

local features. This case study of Tianjin Chinese adds to the expanding number of non-Western case studies of language change (Stanford & Preston 2009) that support earlier generalizations made from Western communities.

Abrupt changes in language sometimes co-occur with huge changes in society, as shown in Elliott (2000). Over the past decades, Chinese society has undergone big changes and these changes are almost certain to be reflected in language. But sociolinguistic study of these changes has been very limited. The changes exhibited in Tianjin tone sandhi, can contribute to the understanding of the mechanism of language change in China. This study fills an important gap by examining the distribution of (53-53), (53-21) and (21-21) by age, sex, style and social class, and by contextualizing the results against Tianjin's sociodemographic history.

But there are some questions that need further study.

First, since the present dissertation studies the social meaning of the three variables from production data alone, it needs data from perception to further check the social meanings and social motivations of the changes. "Since meaning is constructed in the interaction between speaker and hearer, evidence of social meaning ultimately must be sought by bringing together data on production and perception" (Eckert & Labov, 2017: 471). Perception tests on Chinese are very limited, and most of them are matched guise tests (Lambert et al., 1960). I agree with Eckert and Labov (2017) that the study of perception must obtain situated meaning, since the meanings of the variables are not fixed, but must be reflected in the speakers' situated use of the variation (Eckert, 2008). Recent perception studies (Campbell-Kibler, 2008; Podesva et al. 2015) have shown that hearers perceive meanings in the context of utterances, but similar perception studies have not been conducted on Chinese, let alone Tianjin Chinese. It will be important and rewarding to study the social meaning of the three variables from perception and test whether the perceptual

placements and evaluation correspond to production results, and how the context or situation in which the utterance is made affect the social meaning. This type of exploration will not only to enrich the studies on Chinese and Chinese dialects, but also beneficial to the broad sociolinguistic study.

Second, the middle-aged group includes speakers from 35 to 65. And the speakers in this group show diversity in the application of (53-53). For example, some speakers exhibit a great deal of stylistic variation, but not other speakers. A study aiming at this age group with narrowed age division is needed to explore the change happened to this group.

APPENDICES

APPENDIX A. Interview questions

The interview modules were conducted in different orders (Labov 1984). The questions might vary depending on different participants and the interview situations.

The Young group

[Basic demographic information]

1. How old are you?
2. Where do you live?
3. Have you always lived in this district? If no, where else have you lived? When did you live there?
4. Where are your parents from?
5. Had you ever been away from Tianjin for a long time before you grew up?
6. What do your parents do?
7. How much education did your parents get a chance to finish?
8. Do you speak Standard Mandarin in addition to Tianjin?
 - a. Do you speak other dialects?
 - b. Do you speak any other languages? [e.g. English]
9. Where did you go to school?

[Peers]

1. Did you have many friends when you were in high school? How do you keep in touch with them?
2. Do you have many friends in the university?
3. How did you assign space in your dorm? Do you like the place of your bunk?
4. Are your room-mates like you? Do you get on with your room-mates?

[Fight]

1. Have you ever argued with a room-mate? What was it about?
2. Have you fought with students from other places?
3. Have you ever quarreled with a migrant? Did you speak Tianjin when you quarreled with them?

[Dating]

1. Do you have a girl friend or boyfriend? Where is she or he from?
2. How did you get to know each other?
3. Have you told your parents that you have a girl friend or boyfriend?

[Family]

1. Do your parents like your girl friend or boy friend? Are your parents strict with you?
2. Are they satisfied that you are admitted to this university?
3. Do they want you to stay in Tianjin or go to other places after graduation?
4. Did you ever get blamed by your parents for something you didn't do?

[School]

1. Are you satisfied at being admitted to this university?
2. Do you like the university?
3. Do you like the campus? The transportation is inconvenient since the new campus is at suburb.
4. What is your major? Could you tell me something about it?

[Work]

1. Do you want to go to graduate school or work after graduation?
2. Do you have plans to go to other cities like Beijing after graduation?

3. Do you like living in Tianjin? Will you plan to go back Tianjin later after you work or study in other cities like Beijing?
4. Do you think Tianjin has developed as fast as other cities?
5. Do you think the city can provide good opportunities for your work or study?

The Middle-aged group

[Basic demographic information]

1. How old are you?
2. Where do you live?
3. Have you been living in this district or you moved to the district after you grew up?
4. How many years of education did you get a chance to finish?
5. Where are your parents from?
6. Do you speak Standard Mandarin in addition to Tianjin? Do you speak other dialects?
 - a. Do you speak other dialects?
 - b. Do you speak any other languages? [e.g. English]
7. Had you ever been away from Tianjin for a long time before you grew up?
8. What do you do?

[Work]

1. Is your job demanding?
2. Are you satisfied with your salary?
3. Do you want to change a job? Do you think the work is related to your education?
4. Do you think you have equal chances to receive good education?
5. Is it hard to do the housework after a day's working?

[Fights]

1. Do you have conflicts with others in work?
2. How do you treat the conflicts?
3. Have you ever fought with your husband or wife?

[Dating]

1. How did you know your husband or wife?
2. Are you and your husband or wife from families of equal social ranks?
3. Where did you get married? Did you have a lot of guests? Were you nervous on the big day?
Wedding can be stressful. Was there a moment when you thought, “this will never come together!”
Did anything funny happen?

[Family]

1. How old is your kid?
2. Do you have a boy or a daughter?
3. His/her school or work
4. How do you share the education of your kid and housework with your husband or wife?
5. Are you strict with your kid?

[Migrants]

1. Do you like the migrants coming to Tianjin?
2. Do you think the city is becoming better or worse with them?
3. Are there many migrants among your colleagues? Do they hang around with themselves or with Tianjin natives?

The Old group

[Basic demographic information]

1. How old are you?
2. Where do you live?
3. Have you been living in this district or you moved to the district after you grew up?
4. How many years of education did you get a chance to finish?
5. Where are your parents from?
6. Do you speak Standard Mandarin in addition to Tianjin? Do you speak other dialects?
7. Do you speak any other languages? [e.g. English]
8. Had you ever been away from Tianjin for a long time before you grew up?

[Games]

1. Have you been in touch with your former classmates or colleagues?
2. How do you relax in your spare time?

[Family]

1. How old is your kid?
2. Do you have a boy or a daughter? Do you have grandchildren?
3. His/her school or work or marriage
4. How do you kill time with your children or/and grandchildren?

[Dating]

1. How did you know your husband or wife?
2. Are you and your husband or wife from families of equal social ranks?

3. Where did you get married? Did you have a lot of guests? Were you nervous on the big day? Wedding can be stressful. Was there a moment when you thought, “This will never come together!” Did anything funny happen?

[Work]

1. Is your job demanding?
2. Are you satisfied with your salary?
3. Do you want to change a job? Do you think the work is related to your education?
4. Do you think you have equal chances to receive good education?
5. Is it hard to do the housework after a day’s working?

[Fights]

1. Do you have conflicts with others?
2. How do you treat the conflicts?

[Migrants]

1. Do you like the migrants coming to Tianjin?
2. Do you think the city is becoming better or worse with them?

Interviews of every age group end with the same questions related to language attitudes, as shown below.

[Language]

1. What is your first language? Standard Mandarin or Tianjin?
2. Is it pleasant to hear Tianjin dialect?
3. Do you prefer to speak Tianjin dialect at home?
4. Have you ever tried to change the language you speak?
5. Do/Will you allow your children/grand children to speak Tianjin?

6. Do you think people speaking Tianjin dialect are less-educated or/and from a lower social class?
7. What do you think are the features of Tianjin? What do you think is the most salient feature of Tianjin?

Briefing

This interview is about Tianjin people's attitudes towards Tianjin city and Tianjin Chinese. Tianjin has changed a lot because of the influences of Standard Mandarin and the migrants. The project is a sociolinguistic study on the changes and variations of Tianjin. All the data collected will be reported without referring to any specific informant. Your privacy will be carefully protected.

Thank you very much for your cooperation!

APPENDIX B. Word list

pinyin	Gloss	Parts of speech
da ⁵³ hui ⁵³	conference	Noun
lu ⁵³ dou ⁵³	green bean	Noun
da ⁵³ jiu ⁵³	uncle	Noun
jian ⁵³ zhu ⁵³	building	Noun
shou ⁵³ qi ⁵³	be bullied	Adjective
hui ⁵³ bao ⁵³	report	Verb
kuo ⁵³ da ⁵³	amplify	Verb
zhong ⁵³ shu ⁵³	plant trees	Verb
kan ⁵³ bing ⁵³	go to see a doctor	Verb
zhao ⁵³ xiang ⁵³	take a picture	Verb
dui ⁵³ xiang ⁵³	boy friend or girl friend	Noun
jiao ⁵³ shou ⁵³	professor	Noun
jian ⁵³ yi ⁵³	suggest	Noun
si ⁵³ lu ⁵³	route 4	Noun
jin ⁵³ li ⁵³	try one's best	Verb
lang ⁵³ fei ⁵³	waste	Verb
jin ⁵³ bu ⁵³	progress	Verb
bian ⁵³ hua ⁵³	change	Verb
dao ⁵³ qian ⁵³	apologize	Verb
shang ⁵³ ke ⁵³	have class	Verb
da ⁵³ bing ²¹	soldier	Noun
da ⁵³ gu ²¹	aunt	Noun
jiao ⁵³ shi ²¹	teacher	Noun
mian ⁵³ bao ²¹	bread	Noun
kuai ⁵³ can ²¹	fast food	Noun
qi ⁵³ che ²¹	vehicles	Noun
shi ⁵³ bing ²¹	soldier	Noun
hua ⁵³ zhuang ²¹	put on makeup	Verb
jie ⁵³ guang ²¹	excuse me	verb
lu ⁵³ yin ²¹	record	Verb
shi ⁵³ wei ²¹	protest	Verb
hua ⁵³ zhuang ²¹	put on makeup	Verb
jie ⁵³ guang ²¹	excuse me	verb
lu ⁵³ yin ²¹	record	Verb
shi ⁵³ wei ²¹	protest	Verb
bao ⁵³ xiao ²¹	reimburse	Verb
nai ⁵³ xin ²¹	patient	adjective
xi ⁵³ xin ²¹	careful	adjective
ju ⁵³ shuo ²¹	it is said that...	adverb
bi ⁵³ xu ²¹	must	Adverb
sha ²¹ fa ²¹	sofa	Noun
feng ²¹ yi ²¹	wind coat	Noun

jiao ²¹ tong ²¹	transportation	Noun
can ²¹ zhuo ²¹	dining table	Noun
fei ²¹ ji ²¹	airplane	Noun
zhong ²¹ yang ²¹	center	Noun
shu ²¹ bao ²¹	school bag	Noun
gao ²¹ sheng ²¹	be promoted to a higher position	Verb
can ²¹ jia ²¹	participate	Verb
fa ²¹ sang ²¹	make arrangement for a funeral	Verb
jia ²¹ ban ²¹	work overtime	Verb
jiao ²¹ hua ²¹	water flower	Verb
xi ²¹ sheng ²¹	sacrifice	Verb
ban ²¹ jia ²¹	move	Verb
gong ²¹ ji ²¹	attack	Verb
kai ²¹ che ²¹	drive	Verb
shang ²¹ xin ²¹	sad	Adjective
guan ²¹ xin ²¹	careful	Adjective
sheng ²¹ yin ²¹	voice	Noun
xin ²¹ xian ²¹	fresh	Adjective
an ²¹ pai ⁴⁵	arrangement	Noun
fen ²¹ hong ⁴⁵	receive a dividend	Verb
jian ²¹ zhi ⁴⁵	take a part time job	Verb
kai ²¹ men ⁴⁵	open the door	Verb
gong ²¹ chang ²¹³	factory	Noun
ke ²¹ zhang ²¹³	director	Noun
jiao ²¹ shui ²¹³	water	Verb
an ²¹ wen ²¹³	stable	Adjective
gao ²¹ jian ⁵³	good idea	Noun
kuan ²¹ dai ⁵³	broadband	Noun
biao ²¹ shi ⁵³	mark	Verb
kai ²¹ dian ⁵³	run a store	Verb
nan ⁴⁵ fang ⁵³	south	Noun
liang ⁴⁵ xin ⁵³	conscience	Noun
liao ⁴⁵ tian ⁵³	chat	Verb
ti ⁴⁵ gao ²¹	improve	Verb
yin ⁴⁵ hang ⁴⁵	bank	Noun
he ⁴⁵ ping ⁴⁵	peace	Noun
tuan ⁴⁵ yuan ⁴⁵	united	Verb
reng ⁴⁵ ran ⁴⁵	still	Adverb
xie ⁴⁵ di ⁴⁵	sole	Noun
mian ⁴⁵ ao ²¹³	cotton-padded jacket	Noun
xu ⁴⁵ n si ²¹³	try to kill oneself	Verb
tiao ⁴⁵ zheng ²¹³	adjust	Verb
ling ⁴⁵ jian ⁵³	parts	Noun
chi ⁴⁵ dao ⁵³	late	Verb
mi ⁴⁵ xin ⁵³	superstitious	Adjective

qi ⁴⁵ guai ⁵³	strange	Adjective
bao ²¹³ wen ²¹	keep warm	Verb
chan ²¹³ sheng ²¹	produce	Verb
xu ²¹³ duo ²¹	much	Adjective
pu ²¹³ tong ²¹	ordinary	Adjective
fen ²¹³ hong ²¹	pink	Noun
jian ²¹³ cha ⁴⁵	check	Verb
jian ²¹³ zhi ⁴⁵	simply	Adverb
dao ²¹³ mei ⁴⁵	bad luck	Adjective
ba ²¹³ jiao ²¹³	corner	Noun
jiang ²¹³ li ²¹³	reasoning	Verb
ba ²¹³ dui ²¹³	pay one's debts	Verb
yong ²¹³ gan ²¹³	brave	Adjective
gao ²¹³ jian ⁵³	paper	Noun
kuan ²¹³ dai ⁵³	entertain	Verb
biao ²¹³ shi ⁵³	express	Verb
you ²¹³ yong ⁵³	useful	Adjective
ban ⁵³ nian ⁴⁵	half a year	Noun
hou ⁵³ men ⁴⁵	back gate	Noun
kan ⁵³ qi ⁴⁵	keep abreast with	Verb
xing ⁵³ ji ⁴⁵	hot-tempered	Adjective
shi ⁵³ zhang ²¹³	mayor	Noun
zhong ⁵³ shu ²¹³	heatstroke	Verb
mai ⁵³ shui ²¹³	sell water	Verb
dao ⁵³ di ²¹³	in the end	Adverb
du ²¹	governor	Noun
ke ²¹	knock	Verb
chi ²¹	eat	Verb
du ⁴⁵	poison	Noun
ke ⁴⁵	shell	Noun
chi ⁴⁵	late	Adjective
du ²¹³	bet	Verb
ke ²¹³	thirsty	Adjective
chi ²¹³	ruler	Noun
du ⁵³	cross	Verb
ke ⁵³	lesson	Noun
chi ⁵³	red	Adjective

APPENDIX C. Language attitude survey results

Participant	A	B	C	D	E	Total
P1	-1	1	-1	0	0	-1
P2	-1	-1	-1	0	0	-3
P3	1	-1	1	0	0	1
P4	1	1	1	0	0	3
P5	1	1	-1	0	0	1
P6	1	1	1	0	0	3
P7	1	1	1	0	0	3
P8	1	1	1	0	0	3
P9	1	1	-1	0	0	1
P10	1	1	1	0	0	3
P11	1	1	-1	0	0	1
P12	-1	1	-1	0	0	-1
P13	1	1	-1	0	0	1
P14	1	1	1	0	0	3
P15	1	1	1	0	0	3
P16	1	1	1	0	0	3
P17	1	1	-1	0	0	2
P18	1	1	-1	0	0	2
P19	1	1	1	0	0	3
P20	1	-1	-1	0	0	-1
P21	1	1	1	0	0	3
P22	-1	-1	1	0	0	-1
P23	1	1	1	0	0	3
P24	1	1	1	0	0	3
P25	1	1	-1	0	0	1
P26	1	1	-1	0	0	1
P27	-1	1	-1	0	0	-1
P28	-1	-1	-1	0	0	-3
P29	1	1	-1	0	0	1
P30	-1	1	-1	0	0	-1
P31	1	1	1	0	0	3
P32	-1	1	-1	0	0	-1
P33	-1	1	-1	0	0	-1
P34	1	-1	-1	0	0	-1
P35	1	1	1	0	0	3
P36	-1	1	-1	0	0	-1
P37	1	-1	1	0	0	1
P38	1	1	1	0	0	3
P39	-1	-1	-1	-1	0	-4
P40	1	1	-1	1	0	3
P41	1	1	1	1	0	4
P42	1	1	-1	1	0	3

P43	1	1	1	1	0	4
P44	1	1	1	1	0	4
P45	-1	-1	-1	-1	-1	-5
P46	1	-1	1	1	0	2
P47	1	1	1	1	1	5
P48	1	1	1	1	1	5
P49	1	1	1	1	1	5
P50	-1	1	1	1	-1	1
P51	-1	1	1	1	0	2
P52	-1	1	1	1	-1	1
P53	-1	1	-1	1	-1	-1
P54	-1	1	1	1	-1	1
P55	1	1	1	1	0	4
P56	-1	-1	-1	-1	-1	-4
P57	1	-1	-1	-1	0	-2
P58	-1	-1	-1	1	-1	-3
P59	-1	-1	-1	-1	0	-4
P60	1	1	-1	1	0	2
P61	1	1	1	0	0	3
P62	1	1	-1	1	0	2
P63	-1	-1	-1	-1	0	-4
P64	-1	-1	-1	-1	0	-4
P65	1	1	1	1	0	4
P66	1	1	-1	1	-1	1
P67	1	1	1	1	0	4
P68	1	1	1	1	1	5
P69	1	-1	-1	-1	-1	-3
P70	1	1	1	1	1	5
P71	1	-1	1	1	-1	1
P72	1	1	1	1	1	5
P73	-1	1	-1	1	-1	-1
P74	1	1	1	1	0	4
P75	-1	1	-1	-1	-1	-3
P76	1	1	1	1	-1	3
P77	-1	-1	-1	-1	-1	-5
P78	-1	1	-1	-1	0	-2
P79	1	1	1	1	1	5
P80	1	1	-1	1	-1	1
P81	1	1	1	1	0	4
P82	1	1	-1	1	0	3
P83	-1	1	1	1	0	3
P84	-1	1	1	1	0	3
P85	-1	1	-1	0	0	-1
P86	1	1	1	0	0	3
P87	1	1	-1	0	0	1

P88	-1	-1	-1	0	0	-3
P89	-1	1	-1	-1	-1	-3
P90	1	1	-1	-1	-1	-1

BIBLIOGRAPHY

BIBLIOGRAPHY

- Auer, P. (2005). Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations. In N. Delbecque, J. Auwera, & D. Geeraerts (Ed.), *Perspectives on variation: Sociolinguistic, historical, comparative* (pp. 7-42). New York: Mouton de Gruyter.
- Baranowski, M. (2008). The fronting of the back upgliding vowels in Charleston, South Carolina. *Language Variation and Change*, 20(3), 527-551.
- Barrett, R. (2008). Linguistic differentiation and Mayan language revitalization in Guatemala. *Journal of Sociolinguistics*, 12(3), 275-305.
- Bates, D., Macchler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* <doi:10.18637/jss.v067.i01>, 67(1), 1-48.
- Bell, A. (1984). Language style as audience design. *Language in Society*, 13(2), 145-204.
- Boersma, P., & Weenink, D. (2014, January). *Praat: Doing phonetics by computer*. Retrieved from <http://www.praat.org/>
- Campbell-Kibler, K. (2008). I'll be the judge of that: Diversity in social perceptions of (ING). *Language in Society*, 37, 6-37.
- Carmichael, K. (2017). Displacement and local linguistic practices: R-lessness in post-Katrina Greater New Orleans. *Journal of Sociolinguistics*, 21(5), 696-719.
- Carvalho, A. M. (2004). I speak like the guys on TV: Palatalization and the urbanization of Uruguayan Portuguese. *Language Variation and Change*, 16(2), 127-151.
- Chan, M. K.-M. (1985). *Fuzhou phonology: A non-linear analysis of tone and stress*. Seattle, WA: University of Washington dissertation.
- Chao, Y. (1968). *A grammar of spoken Chinese*. Berkeley: University of Berkeley Press.
- Chen, M. (1986). The paradox of Tianjin tone sandhi. *Proceedings of the Chicago Linguistic Society Meeting*, (pp. 98-114).
- Chen, M. (2000). *Tone sandhi: Patterns across Chinese dialect*. Cambridge: Cambridge University Press.
- Chen, P. (1999). *Modern Chinese: History and sociolinguistics*. New York: Cambridge University Press.
- Cofer, T. M. (1972). *Linguistic variability in a philadelphia speech community*. Dissertation available from ProQuest. AA17313386.

- Cotter, W. M., & Uri, H. (2015). Social integration and dialect divergence in coastal Palestine. *Journal of Sociolinguistics*, 19(4), 460-483.
- Coupland, N. (1980). Style-shifting in a Cardiff working-setting. *Language in Society*, 9(1), 1-12.
- Coupland, N. (1985). 'Hark hark the lark': Social motivations for phonological style-shifting. *Language and Communication*, 5, 153-171.
- Duanmu, S. (1995). Metrical and tonal phonology of compounds in two Chinese dialects. *Language*, 71(2), 225-259.
- Dubois, S., & Barbara, H. (2000). When the music changes, you change too: Gender and language change in Cajun English. *Language Variation and Change*, 11(3), 287-313.
- Eckert, P. (2000). *Linguistic variation as social practice: The linguistic construction of identity in Belten High*. Oxford: Blackwell.
- Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12(4), 453-476.
- Eckert, P., & Labov, W. (2017). Phonetics, phonology and social meaning. *Journal of Sociolinguistics*, 21(4), 467-496.
- Eckert, P., & McConnell-Ginet, S. (2003). *Language and gender*. Cambridge: Cambridge University Press.
- Eckert, P., & Rickford, J. R. (Eds.). (2001). *Style and sociolinguistic variation*. Cambridge: Cambridge University Press.
- Elliott, N. C. (2000). Rhoticity in the accents of American film actors: A sociolinguistic study. *Standard Speech, Voice and Speech Review*, 1(1), 103-130.
- Feng, H. (2007). *Dialect leveling and dialect maintenance: contact-induced changes in Tianjin speech in the context of globalization (Doctoral dissertation)*. Tianjin: Nankai University.
- Feng, H. (2015). *Dialect leveling and dialect maintenance: Contact-induced changes in Tianjin speech in the context of urbanization*. Tianjin: Tianjin University Press.
- Ferguson, C. (1959). Diglossia. Word 15: 325-340. In D. Hymes (Ed.), *Reprinted in Language in culture and society (1964)* (pp. 429-437). New York: Harper & Row.
- Ferrer, R. C., & Sankoff, D. (2003). Identity as the primary determinant of language choice in Valencia. *Journal of Sociolinguistics*, 7(1), 50-64.
- Fischer, J. (1958). Social influences on the choice of a linguistic variant. *Word*, 14(1), 47-56.
- Fought, C. (1999). A majority sound change in a minority community: /u/-fronting in Chicano English. *Journal of Sociolinguistics*, 3(1), 5-23.

- Fu, C. (1991). *Shiji zhijiao de chengshi jianshe: Tianjin shinian jianshe jishi* (City construction at the turn of century: Documentation of Tianjin city construction in ten years). Beijing: Kexue Chubanshe (Science publisher).
- Gallois, C., Ogay, T., & Giles, H. (2005). Communication accommodation theory: A look back and a look ahead. In W. B. Gudykunst, *Theorizing about intercultural communication* (pp. 121-148). Thousand Oaks, California: Sage.
- Gao, J., & Lu, J. (2003). The changing FF rule in Tianjin dialect. *Studies on Modern Phonetics and Phonology* (pp. 252-259). Tianjin: Tianjin Social Science Press.
- Giles, H., & Ogay, T. (2007). Communication accommodation theory. In B. B. Whaley, W. Samter, & (eds.), *Explaining communication: Contemporary theories and exemplars (Lea's communication Series)* (pp. 293-310). Mahwah, New Jersey: Lawrence Erlbaum Association.
- Goldsmith, J. (1976). *Autosegmental phonology*. Cambridge, MA: MIT PhD dissertation.
- Gu, X., & Liu, Y. (2003). Tianjin fangyin bianhua de shehui yuyanxue fenxi (A sociolinguistic analysis of the change in Tianjin dialect). *Studies on Modern Phonetics and Phonology* (pp. 260-269). Tianjin: Tianjin Social Science Press.
- Han, G. (1993). *Tianjin dialect*. Beijing: Beijing Yanshan Press.
- Hinskens, F. (1998). Dialect levelling: A two-dimensional process. *Folia Linguistica Historica*, 32(1/2), 35-51.
- Holmes, J. (1995). *Women, men and politeness*. London: Longman.
- Jiang, M. (2007). Cong Shanghaihua de bianhua kan Putonghua dui fangyan de yingxiang (The influences on Putonghua on dialects: Changes in Shanghai dialect. *Jichang Xueyuan Xuebao* (Journal of Jichang College)(4), 94-96.
- Johnsen, S. S. (2015). Dialect change in South-east Norway and the role of attitude in diffusion. *Journal of Sociolinguistics*, 612-642.
- Kiesling, S. (1998). Men's identities and sociolinguistic variation: The case of fraternity men. *Journal of sociolinguistics*(2), 69-99.
- Kwan, M. (2001). *The salt merchants of Tianjin state-making and civil society in late imperial China*. Honolulu: University of Hawai'i Press.
- Labov, W. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, W. (1984). Field methods of the project on linguistic change and variation. In J. Baugh, & J. Sherzer, *Language in use* (pp. 28-53). Englewood Cliffs: Prentice Hall .

- Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change*(2), 205-254.
- Labov, W. (1994). *Principles of linguistic change: Internal factors*. Oxford, UK: Blackwell.
- Labov, W. (2001a). *Principles of linguistic change: Social factors*. Oxford: Blackwell.
- Labov, W. (2001b). The anatomy of style-shifting. In P. Eckert, & J. Rickford, *Style and sociolinguistic variation* (pp. 85-108). Cambridge: Cambridge University Press.
- Labov, W. (2006). *The social stratification of English in New York*. Cambridge: Cambridge University Press.
- Labov, W., Cohen, P., & Robin, C. (n.d.). *A preliminary study of the structure of English used by Negro and Puerto Rican speakers in New York City*. Final report, Cooperative Research Project 3091. [ERIC ED 03 019].
- Lambert, W., Hodgson, R., Gardner, R., & Fillenbaum, S. (1960). Evaluative reactions to spoken language. *Journal of Abnormal and Social Psychology*, 60, 44-51.
- Leben, W. (1973). *Suprasegmental phonology*. Cambridge, MA: MIT PhD Dissertation.
- Lennig, M. (1978). *Acoustic measurement of linguistic change: The modern paris vowel system*. Dissertation available from ProQuest. AA17824743.
- Li, S., & Han, G. (1991). Luelun Tianjin fangyandao (On Tianjin dialect island). *Tianjin Shifan Daxue Xuebao-Shehui Kexue Ban (Journal of Tianjin Normal University-Social Science)*(2), 71-76.
- Li, C. (2013). Ruhe dingyi Zhongguo zhongchan jieji: Huafen Zhongguo zhongchan jieji de sange biao zhun (Defining Chinese middle class: The three standards classifying Chinese middle class). *Xue Hai (Academia Bimestris)*(3), 62-71.
- Li, Q., & Chen, Y. (2016). An acoustic study of contextual tonal variation in Tianjin Mandarin. *Journal of Phonetics*, 54(1), 123-150.
- Li, S. (1956). Shuo Tianjinhua de ren zenyang xuexi Putonghua (How to learn Standard Mandarin for people who speak Tianjin dialect). *ZHONGGUO YUWEN (Studies of the Chinese Language)*, IV(1), 24-27.
- Li, X., & Liu, S. (1985). Tianjinhua de liandu biandiao (Tone sandhi in Tianjin dialect). *Zhongguo Yuwen (Studies of the Chinese Language)*(1), 76-80.
- Lin, H.-S. (2008). Variable directional application in Tianjin tone sandhi. *Journal of East Asian Linguistics*, 17(3), 181-226.
- Lu, J. (1997). Tianjin fangyan zhong de yizhong xin de liandu biandiao (A new tone sandhi rule in Tianjin dialect). *Tianjin Shifan Daxue Xuebao-Shehui Kexue Ban (Journal of Tianjin Normal University-Social Science)*(4), 67-72.

- Luthin, H. (1987). The story of California (ow): The coming-of-age of English in California. *Variation in language: NWAV-XV at Stanford* (pp. 312-324). Stanford, CA: Stanford University Department of Linguistics.
- Macaulay, R. (1977). *Language, social class, and education-a Glasgow study*. Edinburgh: Edinburgh University Press.
- Max Planck Institute for Psycholinguistics. (2016). Retrieved from The Language Archive[Computer software]: Retrieved from <http://tla.mpi.nl/tools/tla-tools/elan/>
- McCarthy, J. (1986). OCP effects: gemination and antigemination. *Linguistic Inquiry*, 17, 207-263.
- Menard, S. (1995). Applied logistic regression analysis. *Sage University Paper Series on Quantitative Applications in the Social Sciences*, 07-106. Thousand Oaks, CA: Sage.
- Meyerhoff, M. (2011). *Introducing sociolinguistics*. New York: Routledge.
- Milroy, J. (2001). Language ideologies and the consequences of standardization. *Journal of Sociolinguistics*, 5(4), 530-550.
- Milroy, J., & Milroy, L. (1985). Linguistic change, social network and speaker innovation. *Journal of Linguistics*, 21, 339-384.
- Milroy, J., & Milroy, L. (1997). Varieties and variation. In F. Coulmas (Ed.), *The Handbook of Sociolinguistics* (pp. 47-64). Oxford: Blackwell.
- Milroy, L. (1980). *Language and social networks*. Oxford: Basil Blackwell.
- Myers, R. (1990). *Classical and modern regression with application (2nd ed.)*. Boston: Duxbury.
- Podesva, R. (2008). Three sources of stylistic meaning. In K. Shaw, S. Wagner, & E. Yasue (Ed.), *Texas Linguistic Forum Proceedings of the Symposium about Language and Society-Austin (SALSA 15)*, (pp. 134-143).
- Podesva, R., D'Onofrio, A., Hofwegen, J., & Kim, S. K. (2015). Country ideology and the California Vowel Shift. *Language Variation and Change*, 27, 157-186.
- Poplack, S., & Walker, D. (1986). Going through (L) in Canadian French. In D. Sankoff (Ed.), *Diversity and Diachrony* (pp. 95-107). Amsterdam: Benjamins.
- R Core Team. (2018). *R: A language and environment for statistical computing [Computer software]*. Retrieved from Retrieved from <http://www.R-project.org/>
- R Foundation for Statistical Computing, Vienna, Austria. (2018). *R Core Team (2018)*. Retrieved from R: A language and environment for statistical computing: <http://www.R-project.org/>

- Reid, E. (1978). Social and stylistic variation in the speech of children: Some evidence from Edinburgh. In P. Trudgill (Ed.), *Sociolinguistic patterns in British English* (pp. 158-173). London: Edward Arnold.
- Schilling-Estes, N. (2002). Investigating stylistic variation. In J. Chambers, & N. Schilling-Estes, *The Handbook of Language Variation and Change* (pp. 375-401). Oxford: Blackwell.
- Schleef, E., Meyerhoff, M., & Clark, L. (2011). Teenagers acquisition of variation: A comparison of locally-born and migrant teens realisation of English (ing) in Edinburgh and London. *English world-wide*, 32(2), 206-236.
- Sharma, D. (2018). Style dominance: Attention, audience, and the 'real time'. *Language in Society*, 47(1), 1-31.
- Shi, F. (1986). Tianjin fangyan shuangzizu shengdiao fenxi (Analysis of disyllabic tones in Tianjin dialect). *Yuyan Yanjiu (Linguistic Research)*, 1(1), 77-90.
- Shi, F. (1988). Tone and tone sandhi in Tianjin dialect. *Zhongguo Yuwen (Studies of the Chinese Language)*(5), 351-359.
- Shi, F., & Wang, P. (2004). Tianjinhua shengdiao de xin bianhua (New changes in Tianjin tones). In F. Shi, & Z. Shen (Ed.), *The joy of research: a festschrift in honor of Professor William S.-Y. Wang on his seventieth birthday* (pp. 176-188). Tianjin: Nankai University Press.
- Stanford, J., & Preston, D. (2009). Variation in indigenous minority languages. Amsterdam/Philadelphia: John Benjamins.
- Tagliamonte, S. (2006). *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press.
- Tagliamonte, S. A. (2012). *Variationist sociolinguistic: Change, observation, interpretation*. Wiley-Blackwell.
- Tan, F. (1986). Yetan Tianjin fangyan de liandu biandiao (On Tianjin dialect tone sandhi, revisited). *Zhongguo Yuwen (Studies of the Chinese Language)*(6), 447-450.
- Tan, R. (2015). *Tianjin fangyan yu jingu wenhua (Tianjin dialect and Tianjin culture)*. (1, Ed.) Beijing: Zhongguo Guoji Guangbo Chubanshe(Chinese International Broadcasting Publishing House).
- Thomas, E. (1997). A rural/metropolitan split in the speech of Texas Anglos. *Language variation and Change*(9), 309-332.
- Tianjin Municipal Bureau of Statistics & Survey Office of the National Bureau of Statistics in Tianjin. (2018). *Tianjin statistical yearbook*. Beijing: China Statistics Press.
- Trudgill, P. (1972). Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society*, 1(2), 179-195.

- Trudgill, P. (1974). Linguistic change and diffusion: description and explanation in sociolinguistic dialect geography. *Language in Society*(3), 215-246.
- Trudgill, P. (1995). *Sociolinguistics: An introduction to language and society*. Penguin Group.
- Tsung, L. (2014). *Language power and hierarchy: Multilingual education in China*. New York: Bloomsbury publishing.
- Wang, D. (2011). The structural readjustment in Tianjin industry: Present status, characteristics, and strategies. *Qiu Zhi (Seeking Knowledge)*, 3, 33-34.
- Wang, J. (2002). Youxuanlun he Tianjinhua de liandu biandiao ji qingsheng (Optimality Theory and tone sandhi and neutral tone in Tianjin dialect). *Zhongguo Yuwen (Studies of the Chinese Language)*(4), 363-371.
- Wang, L. (2016). *Looking for Tianjin dialect maintainers: Build sound archives (Xunzhao Tianjinhua daiyanren: Wei Tianjin fangyan jiangang haixuan 'hao xiangyin')*. Retrieved 11 2019, http://ex.cssn.cn/zt/zt_zh/fybhztjh/bhtjhzx/201611/t20161125_3290739.shtml
- Wang, X. (2003). Tianjin fangyan sanzizu de liandu biandiao (Tianjin trisyllabic tone sandhi). *Zhongguo Yuwen (Studies of the Chinese Language)*(2), 165-171.
- Wang, X. (2015). *Tianjin trisyllabic tone sandhi: Description and analysis (doctoral qualifying paper)*. East Lansing: Michigan State University .
- Wang, X. (2016). *A sociolinguistic study on Tianjin tone sandhi (Doctoral qualifying paper)*. East Lansing, Michigan: Michigan State University.
- Wang, X. (2017). Migration, local identity and change in Tianjin tone sandhi. *Proceedings of Linguistic Society of America* . 2, pp. 40: 1-15. Linguistic Society of America.
- Wang, X., & Lin, Y.-H. (2017). A unified approach to Tianjin trisyllabic tone sandhi: Metrical conditions and tonal complexity. *Proceedings of the Linguistic Society of America, II*, pp. P19: 1-13.
- Wang, X., & Wagner, S. E. (2019, 10). Local identity and standardization: Evidence from Tianjin Chinese tone sandhi. Eugene, OR.
- Wee, L.-H. (2004). Unraveling the relation between Mandarin tones and musical melody. *Journal of Chinese Linguistics*, 35(1), 128-144.
- Wee, L.-H. (2015). Prominence from complexity: Capturing Tianjin ditonal patterns. *Language and Linguistics*, 16(6), 891-926.
- Wolfram, W. (1969). *Black/White speech differences revisited: A preliminary report*. Washington, D.C.: Center for Applied Linguistics.

- Wolfram, W., & Fasold, R. (1974). Field methods in the study of social dialects. In W. Wolfram, & R. Fasold, *The study of social dialects in American English* (pp. 36-72). Englewood Cliffs, NJ: Prentice-Hall.
- Wood, A., Fletcher, P., & Hughes, A. (1986). *Statistics in language studies*. Cambridge: Cambridge University Press.
- Wright, M. S. (1983). *A metrical approach to tone sandhi in Chinese dialects*. Amherst, MA: University of Massachusetts Dissertation.
- Yip, M. (1988). The obligatory contour principle and phonological rules: A loss of identity. *Linguistic Inquiry*, 19(1), 65-100.
- Yip, M. (2002). *Tone*. Cambridge & New York: Cambridge University Press.
- Yu, M., & Zhang, Y. (2015, 1 12). *Maintain pure Tianjin dialect (Liucun yuanzhiyuanwei 'hao xiangyin')*. Retrieved 11 2019, from Zhongguo dangan zixunwang (Chinese archives information website): http://www.zgdazxw.com.cn/news/2015-01/12/content_81795.htm
- Yuan, J. (2000). *Hanyu fangyan faiyao (Introduction to Chinese dialects)*. Beijing: Yuwen Chubanshe.
- Zhang, J. (2002). *The effects of duration and sonority on contour tone distribution*. New York: Routledge.
- Zhang, J. (2014). Tones, tonal phonology, and tone sandhi. In J. Huang, Y.-H. Li, & A. Simpson (Ed.), *The handbook of Chinese linguistics* (pp. 443-464). John Wiley & Sons. Inc.
- Zhang, J. (2019). Tone mergers in Cantonese: Evidence from Hong kong, Macao and Zhuhai. *Asia-Pacific Language Variation*, 5(1), 28-49.
- Zhang, J., & Liu, J. (2011). Tone sandhi and tonal coarticulation in Tianjin Chinese. *Phonetica*, 68(3), 161-191.
- Zhang, Q. (2005). A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society*, 34(3), 431-466.
- Zhang, S. (1995). Shilun putonghua dui fangyan yuyin de yingxiang (The influences of Putonghua on dialectal sounds). *Yuyan Wenzi Yingyong (Applied Linguistics)*(4), 94-98.
- Zhao, H. (2018). Social meaning in the perception of neutral tone variation in Putonghua. *Asia-Pacific Language Variation*, 4(2), 161-196.
- Zhou, Y. (1997). Tianjin fangyan de qinshu chengwei (Kinship terms in Tianjin dialect). *Guoji Guanxi Xueyuan Xuebao (The Journal of University of International Relations)*(2), 45-52.

