



2018

Tone Sandhi Phenomena In Taiwan Southern Min

Mao-Hsu Chen

University of Pennsylvania, maohsuchen@gmail.com

Follow this and additional works at: <https://repository.upenn.edu/edissertations>



Part of the [Linguistics Commons](#)

Recommended Citation

Chen, Mao-Hsu, "Tone Sandhi Phenomena In Taiwan Southern Min" (2018). *Publicly Accessible Penn Dissertations*. 3049.
<https://repository.upenn.edu/edissertations/3049>

This paper is posted at Scholarly Commons. <https://repository.upenn.edu/edissertations/3049>
For more information, please contact repository@pobox.upenn.edu.

Tone Sandhi Phenomena In Taiwan Southern Min

Abstract

This dissertation investigates various aspects of the tone sandhi phenomena in Taiwan Southern Min (TSM). Previous studies have reported complete tonal neutralization between the two sandhi 33 variants derived respectively from citation 55 and 24 variants, leading to the claim that tone sandhi in this language is categorical. The fact that tone sandhi in TSM is assumed to possess a mixture of properties of lexical and postlexical rules gives rise to the debate over the status of this phonological rule.

The findings of the dissertation shows incomplete neutralization between the two sandhi 33 variants with an indication of an ongoing sound change towards a near- or complete tonal merger, possibly led by female speakers. In addition, citation form is proposed to be more underlyingly represented on account of the fact that subjects, especially old speakers, have stronger association with citation variants than with sandhi variants in the priming experiment. The spontaneous corpus study suggests that the Tone Circle is merely a phonological idealization in light of the systematic subphonemic difference in f_0 between citation X and sandhi X that are supposed to correspond even with some control of conceivable confounding factors. By comparing direct- and indirect-reference models, I argue that tone sandhi in TSM should be analyzed as a head-left Concatenation rule within a DM-based theoretical framework.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Linguistics

First Advisor

Mark Liberman

Keywords

corpus, Taiwan Southern Min, tonal neutralization, tone sandhi

Subject Categories

Linguistics

TONE SANDHI PHENOMENA IN TAIWAN SOUTHERN MIN

Mao-Hsu Chen

A DISSERTATION

in

Linguistics

Presented to the Faculties of the University of Pennsylvania in Partial
Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2018

Supervisor of Dissertation

Mark Liberman
Christopher H. Browne Distinguished Professor in Linguistics

Graduate Group Chairperson

Eugene Buckley
Associate Professor of Linguistics

Dissertation Committee

David Embick, Professor of Linguistics

Jianjing Kuang, Assistant Professor of Linguistics

STONE SANDHI PHENOMENA IN TAIWAN SOUTHERN MIN

COPYRIGHT

2018

Mao-Hsu Chen

Acknowledgements

This dissertation would not have been possible without the people who have supported and helped me in many different ways along the journey. First, I would like to express my sincere gratitude to my advisor, Mark Liberman, who has contributed his immense knowledge to the work right from the beginning – helping me to develop research questions, challenging me on arguments and analyses, and meanwhile providing illuminating comments and reminding me to maintain a broader perspective and see the big picture. It was his patient guidance and continuous support that have led me to the right direction. I am also grateful to the members of my committee, David Embick and Jianjing Kuang, for numerous constructive discussions and their invaluable feedback on projects relevant to this thesis. My sincere thanks also go to other teachers and members at Penn Linguistics community who played important roles in my graduate education, including Gene Buckley, Robin Clark, Amy Forsyth, Don Ringe, Florian Schwarz, Meredith Tamminga, Jiahong Yuan, and Georgia Zelou. I would like to offer my special thanks to all my Taiwan Southern Min language consultants and subjects for their participation in the study.

Throughout the years my fellow graduate students have offered generous intellectual and social support as well as warm friendship. I am deeply grateful to start an academic experience at Penn with a fun, smart, and loving cohort: Sunghye Cho,

Sabriya Fisher, Aaron Freeman, and Anton Karl Ingason. The camaraderie cultivated among the phonetics labmates, Aletheia Cui, Wei Lai, Yong-cheol Lee, Nari Rhee, Jingjing Tan, Jia Tian, Ting Wang, and Hong Zhang, has continuously encouraged and motivated me during this process. My appreciation also goes to colleagues and friends in the department, including Amy Goodwin Davies, Duna Gylfadottir, Einar Freyr Sigurðsson, Betsy Sneller, and Robert Wilder. I've also benefited a great deal from my non-linguist friends and mentors, who have greatly enriched my grad school life in Philadelphia. I would like to acknowledge my Taiwanese fellows, Yi-Lin Chiang, Chia-Yu Chung, Yen Pei Huang, Yih-Chii Hwang, Chih-Yen Huang, Sandy Jan, Ji-Ying Lee, Hua-Min Shen, Eunice Tsao, and Yi-Ju Tseng, and my church friends and families, Yoonjung Byun, Ang Chen, Pastor Enrique Leal, Qianhui Lin, Joe Park, Lani Borgman Shade, and Ruoxin Wang.

Lastly, I wish to thank my family. My parents, Gin-Sheng Chen and Cheu-Mei Wang, my brother, Huang-Hsu Chen, and his lovely family, and my sister, Lai-Sin Chen, have always been the greatest source of love, comfort, advice, and encouragement all throughout my life. A deeply heartfelt thanks go to my wonderful husband, Han-Tien Chang, for his endless love and tremendous support; without his encouragement, this dissertation would not have materialized. Finally, I am profoundly indebted to the Holy One for His steadfast love and His perfect plan for me.

ABSTRACT

TONE SANDHI PHENOMENA IN TAIWAN SOUTHERN MIN

Mao-Hsu Chen

Mark Liberman

This dissertation investigates various aspects of the tone sandhi phenomena in Taiwan Southern Min (TSM). Previous studies have reported complete tonal neutralization between the two sandhi 33 variants derived respectively from citation 55 and 24 variants, leading to the claim that tone sandhi in this language is categorical. The fact that tone sandhi in TSM is assumed to possess a mixture of properties of lexical and postlexical rules gives rise to the debate over the status of this phonological rule.

The findings of the dissertation shows incomplete neutralization between the two sandhi 33 variants with an indication of an ongoing sound change towards a near- or complete tonal merger, possibly led by female speakers. In addition, citation form is proposed to be more underlyingly represented on account of the fact that subjects, especially old speakers, have stronger association with citation variants than with sandhi variants in the priming experiment. The spontaneous corpus study suggests that the Tone Circle is merely a phonological idealization in light of the systematic subphonemic difference in f_0 between citation X and sandhi X that are supposed to correspond even with some control of conceivable confounding factors. By comparing direct- and indirect-reference models, I argue that tone sandhi in TSM should be analyzed as a head-left Concatenation rule within a DM-based theoretical framework.

Contents

Acknowledgements	iii
Abstract	v
Contents	vi
List of Tables	x
List of Figures	xii
1 Introduction	1
1.1 Overview of Taiwan Southern Min	1
1.2 Tone sandhi	3
1.3 Placement of sandhi domain boundaries	7
1.3.1 Regular tone sandhi behavior	7
1.3.2 Exceptions to general tone sandhi	15
1.3.3 Summary of the regular tone sandhi pattern	21
1.4 Characteristics of tone sandhi in TSM	22
1.5 Research questions	24

1.6	Outline of the dissertation	26
2	Literature Review	28
2.1	Previous studies on neutralization	29
2.2	Previous studies on priming experiments	33
2.3	Previous studies on the circumscription of TSM sandhi domain	38
2.4	Previous studies on the status of tone sandhi rules in TSM	52
3	Tonal Neutralization	60
3.1	Production experiment	60
3.1.1	Subjects and materials	60
3.1.2	Procedure	62
3.1.3	Data analysis	62
3.1.4	Results	63
3.2	Perception experiment	68
3.2.1	Subjects and stimuli	68
3.2.2	Procedure	70
3.2.3	Data analysis	71
3.2.4	Results	72
3.3	Discussion	75
4	Priming effects	77
4.1	Participants	78

4.2	Stimuli	78
4.3	Procedure	81
4.4	Data analysis and results	82
4.5	Discussion	87
5	Corpus Analysis	90
5.1	The Spontaneous Speech Corpus of Taiwan Southern Min	91
5.1.1	Speakers and materials	91
5.1.2	Methods	93
5.2	Results	94
5.2.1	Fitting Smoothing Spline ANOVA models	95
5.2.2	Fitting orthogonal polynomial	106
5.2.3	Dealing with confounding	109
5.3	Discussion	116
6	Tone Sandhi as Phrasal Phonological Rules	119
6.1	Overview of Distributed Morphology	120
6.2	Allomorphy in DM	123
6.3	Tone sandhi as Phrase-left Concatenation	127
6.4	Tone sandhi as Head-left Concatenation	138
6.5	Discussion	145
7	Discussion	147

7.1	Incomplete tonal neutralization	148
7.2	Realization of Tone Circle	150
7.3	Interaction between components of grammar	154
8	Conclusion	157
	Appendix A Orthography for Taiwan Southern Min transcription	161
	Appendix B Stimuli of tonal neutralization production experiment	164
	Bibliography	167

List of Tables

1.1	The tonal inventory of Taiwan Southern Min.	5
1.2	Example of the contrastive tones in Taiwan Southern Min.	5
2.1	Characteristics of lexical and postlexical rules.	53
2.2	The logic of double dissociation.	58
3.1	Subject list of the production experiment.	61
3.2	Statistical result of the interaction between citation tone and gender for young speakers.	65
3.3	Statistical result of the effect of citation tone on mean f0 values for young speakers.	65
3.4	Statistical result of the interaction between citation tone and gender for old speakers.	66
3.5	Lexical frequency of target tokens in the perception experiment of tonal neutralization.	69
3.6	Correct response ratio and confidence level of each token for each speaker.	73
3.7	Overall correct response rates and A prime statistics for each subject.	74

4.1	The stimuli design of the priming experiment.	79
4.2	Results of a linear mixed models analysis for old subjects.	86
5.1	Overview of the TSM Corpus.	93
5.2	Distribution of the citation and sandhi variants of each tone category.	95
5.3	Frequencies of citation and sandhi variants in each subject's interview.	96
5.4	kde.test for orthogonal polynomial fitting f0 curves of different variants of the same tone.	107
5.5	kde.test for orthogonal polynomial fitting f0 curves of the same variant produced by different age groups.	108
5.6	kde.test for orthogonal polynomial fitting f0 curves between variants that are supposed to have the same surface contour.	108
A.1	The list of consonants in Taiwan Southern Min.	163

List of Figures

3.1	F0 contours of target “poe” produced by one male and one female speakers in each age group.	64
3.2	Smoothing spline estimates and 95% Bayesian confidence intervals for comparison of two sandhi 33 f0 contours in each age group.	67
3.3	Interaction effects with Bayesian confidence intervals for two sandhi 33 f0 contours for both age groups.	68
3.4	An example trial of the perception experiment.	71
4.1	The procedure of the priming experiment.	82
4.2	Mean reaction time of all subjects.	83
4.3	Mean reaction time of different generation groups.	84
5.1	Mean f0 curves of citation and sandhi variants of each unchecked tone for the two age groups.	97
5.2	F0 curves of citation and sandhi variants of each tone category modeled by SS-ANOVA.	98

5.3	F0 curves of citation and sandhi forms of each tone produced by both age groups modeled separately using SS-ANOVA.	99
5.4	SS-ANOVA modeling f0 curves of citation and sandhi forms that are supposed to correspond to each other produced by both age groups. .	101
5.5	SS-ANOVA modeling non-prepausal f0 curves of citation and sandhi forms that are supposed to correspond to each other produced by young and old speakers.	103
5.6	SS-ANOVA modeling prepausal f0 curves of citation and sandhi forms that are supposed to correspond to each other produced by young and old speakers.	105
5.7	Visualization of the 3-D orthogonal polynomial data.	107
5.8	Durational distributions of citation and sandhi forms grouped by tone category and age groups.	110
5.9	Durational distributions of citation and sandhi forms grouped by tone category and age groups in prepausal and non-prepausal environments.	111
5.10	F0 curves of citation and sandhi forms of each tone grouped by morpheme frequencies modeled separately using SS-ANOVA for both age groups.	113
5.11	SS-ANOVA modeling f0 curves of citation and sandhi forms that are supposed to correspond to each other grouped by morpheme frequencies for both age groups.	115

6.1	Distributed morphology architecture.	120
7.1	F0 curves of citation and sandhi forms that are taken to be identical grouped by pausing environments modeled separately using SS-ANOVA.	152

Chapter 1

Introduction

This dissertation is an investigation into the tone sandhi phenomena in Taiwan Southern Min. Before I delve into the issues presented by this topic, a brief introduction of the language and the tonal alternation is presented in Sections 1.1 and 1.2, followed by a wide range of Taiwan Southern Min data in Section 1.3 in order to show the general tone sandhi behavior. Section 1.4 describes the characteristics of tone sandhi rules in this language. Relevant research questions are listed in Section 1.5, and Section 1.6 gives the outline of this dissertation.

1.1 Overview of Taiwan Southern Min

Southern Min, or other aliases including Hokkien, Hoklo, and Taiwanese, refers to a variant of the south Min, or Min Nan, Chinese dialects spoken in southeastern China (especially the province of Fujian), Taiwan and other Southeast Asian countries such as Malaysia and Singapore. The Southern Min language can trace its roots

back to at least the Tang Dynasty (618-907 AD) and is said to preserve much more characteristics of Old Chinese than Mandarin, and even Cantonese. On account of long-term language separation and contact with other languages, the Southern Min dialect exhibits all kinds of dialectal variations among speakers from different regions. In this dissertation, the focus is on the regional dialect spoken in Taiwan, here dubbed as Taiwan Southern Min (TSM) for expository convenience. It is the native language of approximately 70% of the population in Taiwan (Government Information Office, 2012), and is more commonly referred to as Taiwanese. Because of the policy that promoted Mandarin Chinese as the official language, Taiwan Southern Min has been marginalized in the realm of education, media, and administration until the rise of indigeneness in recent years.

In Taiwan Southern Min, a typical syllable structure can be schematically represented as [initial [onglide [nucleus + coda]]]. Any consonant except the glottal stop [ʔ] can occur as initials, while only nasals, glides and the stops [p, t, k, ʔ] can occupy the coda position. Syllables ending in stops are called *checked* syllables, as opposed to all others ending in a sonorant referred to as *free* or *unchecked* syllables. Nasal consonants can be syllabic [m̩ n̩], and nasality on vowels is also contrastive. While there is no standard writing system for TSM, an orthography based on *Peh-oe-ji* (POJ) Romanization system¹ with a few modifications is used here for transcribing TSM. A complete list of the orthography can be found in Appendix A.

¹*Peh-oe-ji* (POJ), an orthography developed by Western missionaries in the 19th century, uses a modified Latin alphabet along with some diacritics to transcribe variants of Southern Min Chinese, spoken TSM in particular, and it has remained as one of the most popular romanization systems to write TSM, especially among online users.

1.2 Tone sandhi

Tone sandhi, a phenomenon abundant in Chinese dialects, refers to the tonal alternation between variants of a word or morpheme that in most cases only differ in tones, and this tonal change is conditioned by adjacent tones or by where the word or morpheme is located prosodically or morphosyntactically in an utterance.

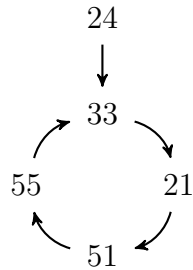
In Taiwan Southern Min, every lexical word has two tonal variants: a *citation* form, also known as *junction* or *base* form, and a *sandhi* form. The process in which lexical tones in citation forms undergo morphophonemic changes to the corresponding sandhi forms is referred to as *tone sandhi* (TS), the application of which depends solely on the position of the word within a phrasal sandhi domain, often termed *tonal group* or *tone group* (TG). The citation tone of a syllable, as the name suggests, surfaces when it is pronounced in isolation or occurs at the right edge of a sandhi domain, namely precedes a phrase boundary, whereas its individually assigned sandhi tone is realized whenever the syllable is not phrase-final, or TG-final. There are in total seven contrastive tones in TSM, including five *free* or *unchecked* tones for free syllables and two *checked* tones for checked syllables. The tone sandhi systems of the two groups are illustrated in (1)²³ and (2) respectively. The former is often referred to as Taiwanese Tone Circle, as suggested by its quasi-circular chain shift. Here checked tones are

²The phonetic values of the tones vary among studies and subdialects. See Tsay (1994) for a chart showing the variation in phonetic values earlier researchers have given to the five long tones in TSM, indicating that while there is disagreement on the absolute values of the tones, the relative differences are clear.

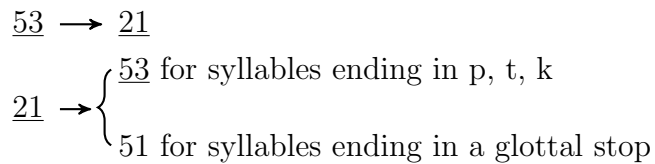
³The 5-point scale of Chao (1930) is used to indicate the pitch height with 5 being the highest and 1 the lowest.

marked with underlines, and the direction of the arrow shows the selection of the surface sandhi form. For instance, the sandhi form of a citation tone of high level 55 is a mid level tone 33, and the sandhi form of a citation tone of mid level 33 is a low falling tone 21, and so on. Table 1.1⁴ lists the corresponding pairs of citation and sandhi tones in manifestation of the eight-tone system in traditional Chinese linguistics.

- (1) Tone sandhi for free syllables (Tone Circle)



- (2) Tone sandhi for checked syllables



- (3) moa -iu 'sesame oil'
Citation form: 24 24
Sandhi form: **33** 24

The application of tone sandhi is exemplified in (3), where the two free syllables

⁴The tone category here refers to the traditional Chinese categories of four tones (*ping*, *shang*, *qu*, and *ru* 'even, rising, leaving/falling, and entering') and two registers (the upper and lower, *yin* and *yang*, registers). In the case of TSM, *yangshang* (lower rising) tone was merged with *yangqu* (lower falling) tone.

Table 1.1: The tonal inventory of Taiwan Southern Min.

Tone	Category	Description	Citation tone	Sandhi tone
Ia	<i>yinping</i>	high level	55 (HH)	33
Ib	<i>yangping</i>	low rising	24 (LM)	33
IIa	<i>yinshang</i>	high falling	51 (HL)	55
IIb	<i>yangshang</i>	-	-	-
IIIa	<i>yinqu</i>	low falling	21 (ML)	51
IIIb	<i>yangqu</i>	mid level	33 (MM)	21
IVa	<i>yinru</i>	mid checked	<u>21</u>	<u>53</u>
IVb	<i>yangru</i>	high checked	<u>53</u>	<u>21</u>

Table 1.2: Example of the contrastive tones in Taiwan Southern Min.

Tone category	Citation form	Sandhi form
Ia /si/ ‘poem’	gim33- si55 ‘to recite poems’	si33 -jin24 ‘poet’
Ib /si/ ‘time’	sui33- si24 ‘anytime’	si33 -kan55 ‘time’
IIa /si/ ‘death’	thai33- si51 ‘to kill’	si55 -lo33 ‘dead end’
IIb	-	-
IIIa /si/ ‘four’	te21- si21 ‘fourth’	si51 -hong55 ‘four sides’
IIIb /si/ ‘temple’	hut <u>21</u> - si33 ‘Buddhist temple’	si21 -bio33 ‘temple’
IVa /sik/ ‘color’	aN21- sik21 ‘red’	sik53 -chhai51 ‘color’
IVb /si/ ‘ripe’	bin21- sik53 ‘look familiar’	sik21 -sai33 ‘to know’

moa ‘sesame’ and *iu* ‘oil’ have the same citation tone 24, but only the first syllable, occurring at a non-TG-final position, undergoes the sandhi rule and changes into mid level 33, as marked in bold. The phonetic output of the second syllable remains its original citation tone since the syllable occurs at utterance-final, also the right edge of a sandhi domain. Table 1.2 illustrates this with all seven lexical tones occurring at the first and second syllable of a disyllabic word respectively, in which the first syllable is realized with the sandhi tone while the second syllable is pronounced with the citation tone. Assuming that citation form is underlying, regardless of the phonetic details, the tone sandhi pattern illustrated in (1) and (2) can be schematized as the

Tone Sandhi Rule (TSR) in (21), which has the effect of converting all but the final citation tone within a tone group into their corresponding sandhi tones.

(4) *Tone Sandhi Rule* (TSR) (Chen 1987: 113)

$T \rightarrow T' / _ T$ within a sandhi domain, where T is citation tone and T' is sandhi tone.

There is essentially no limit on the length of a sandhi domain which can be as small as a mono- or disyllabic word, as in (3), a phrase, or even a sentence, as in (5) adopted from Chen (1987:118(8)), where sandhi tones are marked in bold. A sandhi domain, as a consequence, may consist of an arbitrarily long string of sandhi tones followed by a single citation tone, as illustrated with the regular expression in (6), where the asterisk indicates zero or more instances of the preceding element and the # sign demarcates the right edge of a sandhi domain, following the notational conventions adopted in Chen (1987). This boundary notation is used in all relevant examples in this dissertation.

(5)

yi	kiong	-kiong	kio	gua	koh	khoaN	poaN	tiam	-cheng	ku	chheh
Citation:	55	24	24	21	51	21	21	21	51	55	51 21
Sandhi:	33	33	33	51	55	51	51	51	55	33	55 21
	he by-forece		ask me more		read	half	hour	long book			

‘He insisted that I read for another half an hour.’

(6) # (T')* T #

1.3 Placement of sandhi domain boundaries

1.3.1 Regular tone sandhi behavior

While the size of a sandhi domain can theoretically be infinite, the placement of sandhi domain boundaries is certainly not arbitrary. A variety of tone sandhi examples are given in this section in order to reveal some pattern of sandhi domain distribution. Since the main concern lies not in the phonetic correspondence between citation and sandhi tones but in what determines a tone sandhi domain in TSM, tonal specifications are omitted here.

The first observation can be made from the contrast shown in (7), where the = sign denotes the absence of # such that syllables on either side are linked together as members of the same sandhi domain, and thus obligatory sandhi is expected⁵. (7a) and (7b) have exactly the same string but differ in their meanings, sandhi domain groupings, and syntactic structures. In (7a), *khah kin* ‘faster’ is the adverbial complement of the VP headed by the verb *sia* ‘write’, but in (7b), it is an adjective predicate of the sentential subject *i sia* ‘he writes’. This distinction suggests that a sandhi domain boundary is observed not only at the end of the sentence⁶ but also at the right edge of a non-pronominal subject.

- (7) a. i [sia = khah kin]_{VP} (Chen, 1987, p.119(19))
he write more fast

⁵The marking of ‘=’ is optional, so there is no difference between A = B # C and A B # C.

⁶Here, as in most examples in this dissertation, utterance-final # is omitted.

‘He writes faster.’

- b. [i sia]_s # khah kin
he write more fast
‘It would be faster for him to write.’

Another contrast is manifested in the minimal pair of sentences in (8), where the only difference between the two utterances resides in the surface form of the second syllable in *moa-a* ‘sesame seed(s)’, since the NPs *moa-a* in (8a) and (8b) play distinct roles in modifying the bun with respect to the adjective *toa* ‘big’. The NP *moa-a* in (8a) serves as the subject of a relative clause modifying *sio-piaN* ‘bun’, whereas it functions as an adjunct of the adjective *toa* ‘big’, suggesting that tone sandhi applies between an adjunct and the following constituent to which the adjunct attaches; accordingly, the right edge of the adjunct is not marked with a sandhi domain boundary. In the same way, no sandhi domain boundary is found after an adjunct modifying the ensuing verb, as illustrated in (9a).

- (8) a. *moa-a* # *toa* # *e sio-piaN* (Chen, 1987, p.124(29))
sesame seed big E bun
‘bunds with big sesame seeds’
- b. *moa-a* = *toa* # *e sio-piaN*
sesame seed big E bun
‘bunds as big as sesame seeds’ (‘tiny buns’)
- (9) a. *i loan-chu* = *kong*
he mindlessly talk
‘He is talking mindlessly.’
- b. *gua ka-chai* # *che chit pan ki*
I fortunately take this Cl flight

‘Fortunately, I am taking this flight.’

However, not all adjuncts behave the same. The contrast in (9) shows that tone sandhi in TSM makes a distinction between sentential adverbs (S-adverbs) and VP adverbs (VP-adverbs). While the sentential adverb *ka-chai* ‘fortunately’ in (9b) and each of the S-adverbs in (10) constitute a sandhi domain on its own, the VP-adverbs *loan-chu* in (9a) and *chiah* in (11)⁷ phrase together with the following verb. The two types of adverbs can be distinguished by their complementary distribution within a sentence as illustrated in (12). First, while only S-adverbs as in (12d) can be preposed to pre-subject, sentence-initial position, VP-adverbs cannot as shown in (12a). Second, suppose that the negative particle *bo* is the head of a NegP located between IP/TP and VP, *bo* can only precede a VP-adverb as in (12c) but not an S-adverb as in (12f), i.e., VP-adverbs cannot scope over negation as verified in (12b) whereas S-adverbs can as in (12e). Note that locatives and temporal modifiers such as *keh-tng-kang* ‘the next day’ and *au-pai* ‘next time’ in (10) pattern with sentential adverbs and constitute their own sandhi domains.

- (10) ki-sit # au-pai # kho-leng # i e khi Bi-kok #
actually next time maybe he will go USA
‘Actually, next time maybe he will go to USA.’

⁷Note that 0 indicates the occurrence of a neutral tone, where this tone-bearing syllable is de-emphasized by losing its original tone and has a surface form different from its base and sandhi variants. Neutral tones generally occur at phrase-finally and are pronounced in a lighter and softer way with either a low falling tone 21 or a pitch depending on the previous tone. The preceding syllable is emphasized insofar as its corresponding base form is surfaced.

- (11) keh-tng-kang # chiah e tng#-lai0
 the second day just will return
 ‘(She) will not return until the next day.’
- (12) a. *[loan-chu]_{AdvP} i kong
 mindlessly he talk
 ‘He talked mindlessly’
- b. *i [loan-chu]_{AdvP} bo kong
 he mindlessly not talk
 ‘He did not talk mindlessly.’
- c. i bo [loan-chu]_{AdvP} = kong
 he not mindlessly talk
 ‘He didn’t talk mindlessly.’
- d. [ka-chai]_{AdvP} # gua che chit pan ki
 fortunately I take this Cl flight
 ‘Fortunately I took this flight.’
- e. gua [ka-chai]_{AdvP} # bo che chit pan ki
 I fortunately not take this Cl flight
 ‘Fortunately I did not take this flight.’
- f. *gua bo [ka-chai]_{AdvP} che chit pan ki
 I not fortunately take this Cl flight
 ‘Fortunately I did not take this flight.’

In a double-object construction like (13), tone sandhi applies between verb and its following object, either the direct object or the indirect object, but is blocked between the two objects such that a sandhi domain boundary is located between *chit e lau peng-iu* ‘an old friend’ and *in bO* ‘his wife’ in (13a). When one of the objects is moved forward and precedes the verb, that fronted object phrases separately from the following verb, as shown in (13b) and (13c).

- (13) a. i kai-siao chit e lau peng-iu # hO in bO
 he introduce one Cl old friend to his wife
 ‘He introduced an old friend to his wife.’
- b. i ka tang-oh # kai-siao chit e lu-peng-iu
 he to schoolmate introduce one Cl girlfriend
 ‘He introduced a girlfriend to his schoolmate.’
- c. i chiong hit pun chheh # sang hou tang-oh
 he obj-marker that Cl book give to schoolmate
 ‘He gave that book to his schoolmate.’

Except under focus or in contrastive use⁸, pronouns, as shown in (13) and (7), generally do not form their own sandhi domain but phrase together with whatever following them, either as subjects such as *i* ‘she’ in (14), or as objects, e.g. *in* ‘them’ and *lang* ‘someone’ in (14). On the other hand, non-pronominal counterparts always form their own sandhi domain, such as *hit pun chheh* ‘that book’ in (13c), the CP-clause subject in (7b), and ‘sesame seeds’ as the subject of the CP in (8a).

- (14) i kau-tai in nng e lang # m-thang sui-pian # ka lang khui
 she urge them two E people cannot arbitrarily for someone open
 mng #
 door
 ‘She urges them not to arbitrarily open the door for strangers.’

For modification structures across varieties of Chinese, all modifiers, including adjectives, nouns or relative clauses, generally precede what is being modified. In TSM, the modifier(s) and the modified object(s) are often separated by a modificational

⁸When pronouns are in contrastive focus, they could be realized either in their citation form or in their sandhi form but with greater vocal effort such as intensity or longer duration.

marker (or its other aliases such as subordinator, complementizer, or relative clause marker) *e*, parallel to the modificational marker *de* in Mandarin. The presence of particle *e*, although sometimes being optional as shown in (15), is much more common in natural speech, and there is always a tone sandhi boundary before marker *e* and following the preceding modifier, either a noun as in (16), an adjective as in (15b) and (17), or a relative clause as in (8a). For structures involving coordinate constructions such as (18), not only the conjunct before marker *e* but also all the other previous conjuncts are marked with a sandhi domain boundary respectively.

- (15) a. pin-toaN = hak-seng
 lazy student
 ‘lazy student(s)’
- b. pin-toaN # e hak-seng
 lazy E student
 ‘lazy student(s)’
- (16) hak-hau e # lau-su
 school E teacher
 ‘School teacher(s)’
- (17) toa-han # e cha-bO gin-a # kio-cho a-khim #
 elder E female child call as A-Khim
 ‘The elder daughter is called A-Khim.’
- (18) chhong-beng # khun-lat # e = gni-a
 smart hard-working E child
 ‘a smart and hard-working child’

While there is generally no distinction between countable and uncountable nouns,

and singular and plural forms of nouns in Chinese languages, nouns can be counted with numerals along with specific classifiers associated with the nouns, such that an extended DP/NP can be expressed with the construction of [(demonstrative) + numeral + classifier + noun]. (19) shows the interactions between tone sandhi application and movement of such a determiner/noun phrase. In (19a), the DP *chit chhut liok-iaN-phiN* ‘a video movie’ as the direct object of the preceding verb is followed by a sandhi domain boundary, and tone sandhi applies among the elements within it. When the NP *liok-iaN-phiN* ‘video movie’ is preposed to sentence-initial topic position as in (19b), it itself forms a single sandhi domain, whereas the rest of the DP remains in situ and phrases together with the following infinitival clause. (19c) demonstrates a case in which the entire DP is preposed and the NP is further fronted to sentence-initial, and both the NP and the “DP-leftover” *chit chhut* are followed by a sandhi domain boundary.

- (19) a. chO chit chhut = liok-iaN-phiN # lai khoaN
 rent one Cl video movie to watch
 ‘rent a video movie to watch’
- b. liok-iaN-phiN # chO chit chhut = lai khoaN
 video movie rent one Cl to watch (same as (a))
- c. liok-iaN-phiN # chit chhut # chO loa-choe chiN
 video movie one Cl rent how much money
 ‘How much does a (cassette of) video movie rent for?’

For biclausal structures, the tone sandhi pattern is less clear and has not been discussed in previous studies. The two examples given below are original and are drawn

from the corpus study. The contrast between (19a) and (19b), where the matrix verb *chO* ‘rent’ takes a purposive infinitival clause complement⁹, reveals that two clauses together can form a single sandhi domain, which is also true for sentences with a matrix-clause verb that introduces indirect speech, thought, or belief taking a finite clausal complement as in (20). However, there exist cases involving non-application of tone sandhi across clausal boundaries as in (21), where two clauses form individual sandhi domains.

- (20) a. Chhun-Kiau # siong-sin = $_{CP}$ [A-Bin # boe lai]
 Chhun-Kiau believe A-Bin will not come
 ‘Chhun-Kiau believed that A-Bin would not come.’
- b. gua jin-ui = $_{CP}$ [A-Bin # kong # e khah u to-li]
 I think A-Bin say E more have sense
 ‘I think what A-Bin said makes more sense.’
- (21) a. $_{CP}$ [sia chit phiN bun-chiong # chin-cheng,] # gua su-iau khah
 write this Cl essay before I need more
 che gian-kiu
 many research
 ‘Before writing this essay, I need to do more research.’
- b. gua chin thiam, # $_{CP}$ [sO-i # lan lai tng#-khi0]
 I very tired so we let go back
 ‘I’m very tired, so let’s go home.’
- c. $_{CP}$ [i na-si = boe = hiau tai-oan-oe] # li to ka i hoan-ek
 he if Neg know Taiwanese you then to him translate
 ‘If he doesn’t know Taiwanese, you translate for him.’

⁹The *lai*-purposive is characterized by the presence of the element *lai* with the literal meaning ‘come’ but is usually translated into ‘in order to/for’. Lin and Liao (2008) further noted that *lai* can also be replaced by *khi* ‘go’ without substantial change in meaning except that it conveys a strong sense of directionality.

- d. *CP*[i sui-lian # ni-ki # ia toa] # (put-ko #) sin-the # ia
 he although age quite big however body quite
 ho
 good
 ‘Although he is quite old, he is pretty healthy.’

1.3.2 Exceptions to general tone sandhi

Apart from the morphosyntactic conditions mentioned above, there are certainly other factors at play in determining tonal alternations in TSM. First of all, phrases like idiomatic expressions as in (22) and verses in (23) distinctly exhibit some rhythmic effects on tone sandhi application. While each of the monosyllabic words phrases individually in (22a), the four-syllable idiom in (22b) is divided into two (or three) sandhi domains.

- (22) a. hong # chhoe # hou # ak
 wind blow rain pour
 ‘stormy weather (lit. wind blows and rain pours)’
- b. kok (#) thai # bin = an
 country prosper people peace
 ‘The country prospers and the people live in peace.’

As both of these idioms share the same configuration [NP VP NP VP], which can essentially break into two clauses, one would expect the two idioms be pronounced in the same way in terms of tone sandhi application. Based on the abovementioned tone sandhi pattern, each of the NP subjects and VP predicates is supposed to form its own sandhi domain and therefore each character should be bounded by #, just

as exemplified in (22a). (22b), on the other hand, shows a common alternative to parse such quadrisyllabic expressions into two iambic feet, in which the first syllable is de-emphasized and surfaces with its sandhi variant whereas the second syllable is stressed and remains in its citation form.

- (23) a. kO lin # put kian # kim-si guat # (Chen, 2000, p.472(88b))
 ancient people not see today moon
 ‘people of yore do not see today’s moon’
- b. han tek # ho # ha thO #
 sweat drip seedlings of crops under soil
 ‘(his) sweat drips on the ground beneath the seedling’

As typical in classical Chinese poetry, verses composed of pentasyllabic and heptasyllabic lines as in (23) are generally parsed into rhythmic structures of $[(\sigma\sigma)(\sigma\sigma\sigma)]$ and $[(\sigma\sigma)(\sigma\sigma)(\sigma\sigma\sigma)]$ in each line respectively. The heptasyllabic poetic line in (23a) conforms to such a metrical template irrespective of its syntactic structure, which would suggest a reading made up of only two sandhi domains: *kO lin # put kian kim-si guat #*; namely, there should not be a sandhi domain boundary between the verb *kian* ‘see’ and the direct object *kim-si guat* ‘today’s moon’. (23b) shows another case of tone sandhi application being compromised with rhythmic effect. In ordinary speech, the subject NP *han* ‘sweat’ ought to form a separate sandhi domain, and the verb *tek* ‘drips’ should join its argument *ho* ‘crop seedlings’ in one sandhi domain.

Second, TSM has other kinds of tonal alternation phenomena. Neutral tone sandhi, for example, has the effect of placing some syllables in the neutral tone, which

is generally lighter and softer¹⁰ in a way to de-stress the syllable and allow emphasis to fall on the preceding syllable that remains in its citation tone. For instance, the entire sentence in (24) forms a single sandhi domain, in which the last monosyllabic word *goa* ‘me’ undergoes neutral tone sandhi (annotated with 0) and its preceding syllable *sin* in bold is realized in its citation tone. For situations emphasizing that it is the speaker and not someone else that the addressee must believe, there is a contrastive focus on *goa* such that it remains in its citation form while regular tone sandhi applies to all other antecedent syllables.

- (24) li ai siong-**sin** goa0 (or ‘li ai siong-sin **goa.**’)
 you must believe me
 ‘You must believe me.’

In cases where consecutive syllables are affected and realized in the neutral tone, the syllable before the first neutral tone syllable stays in its citation form. Such instances are commonly found in complex verb phrases, such as VPs with direction complement as in (25), VPs with resultative complement, and VPs with degree complement. As this sandhi rule applies to one or more syllables near the right edge of a sandhi domain, neutral tones are often found at the end of a clause or sentence, but can also occur in the middle of a clause.

¹⁰Ou and Hsiao (1997) report that the affected syllables first undergo tone loss and then are realized as either 1) ‘constant low neutral tone’, or 2) ‘neutral tone derived from preceding tone spreading’. The tone value of the first type of neutral tone is always a low tone that is lower than tone 21, while the tone value of the second type depends on the offset of the base tone of the preceding syllable, referred to as a phenomenon of tone spreading.

- (25) **thiau** khi0-lai0
jump upward
'jump up'

Furthermore, there exist some lexical specific tone changing rules in TSM, e.g. the one applying to the syllable preceding the diminutive suffix *-a*, which shows a tonal value of high falling 51. While the noun suffix has the meaning of 'small thing', it has now become a noun marker and does not necessarily imply smallness. Previous researches propose that the realization of the syllable before *-a* depends on the result of the application of regular tone sandhi. If the sandhi tone of that syllable is a high falling 51 tone, it will further change to a high level 55 tone. If the outcome is a non-high level 33 tone or a low-falling 21 tone, the syllable will surface with a rising 24 tone, and there is no further tonal alternation when the sandhi form is a high level 55 tone (Tung et al., 1967, Zhang, 1988). In other words, the syllable directly preceding the diminutive *-a* can only surface in either a high level 55 tone or a rising 24 tone, as exemplified in (26).

- (26) Realization of the syllable before suffix *-a*
- a. high level 55 tone
 - (i) kau-a 'dog' (*kau*: citation 51 → sandhi 55)
 - (ii) iN-a 'swallow (bird)' (*iN*: citation 21 → sandhi 51)
 - b. rising 24 tone
 - (i) ti-a 'pig' (*ti*: citation 55 → sandhi 33)

- (ii) he-a ‘shrimp’ (*he*: citation 24 → sandhi 33)
- (iii) bo-a ‘hat’ (*bo*: citation 33 → sandhi 21)

In addition, TSM has a special construction involving reduplication with monosyllabic adjectives that shows irregularity in tonal alternation. A monosyllabic adjective morpheme can repeat itself twice, resulting in triplication, to intensify the meaning of the adjective. If the morpheme has a falling citation tone, either tone 51 or 21, the first two syllables surface the same with the regular sandhi tone whereas the last syllable stays in its citation form. However, if the citation tone of the morpheme is 55, 33, or 24, the first syllable of the triplication shifts to an extraordinary rising tone, similar to tone 24, rather than the corresponding sandhi tone, while the second and third syllables conform to regular tone sandhi, as exemplified in (27).

- (27) First syllable of the adjective triplication (citation → sandhi)
- a. regular tone sandhi
 - (i) nN55-nN55-nN51 ‘extremely soft’ (*nN*: 51 → 55)
 - (ii) phong51-phong51-phong21 ‘remarkably swollen’ (*phong*: 21 → 51)
 - b. rising tone, similar to tone 24
 - (i) sin**24**-sin33-sin55 ‘very new’ (*sin*: 55 → 33)
 - (ii) tam**24**-tam33-tam24 ‘soaking wet’ (*tam*: 24 → 33)
 - (iii) kau**24**-kau21-kau33 ‘very thick’ (*kau*: 33 → 21)

Another lexical exception to general tone sandhi relates to titles. Among terms of address, only *sian-siN* ‘Mr.’ is subject to neutral tone sandhi when preceded by a surname, in which case both syllables of *sian-siN* are realized with the neutral tone, while the monosyllabic surname remains in its citation tone, which is marked in bold in (28). Other titles such as *sio-chia* ‘Ms., madam’, *lau-su* ‘teacher’, *i-su* ‘doctor’, etc. comply with regular tone sandhi rules, and so does *sian-siN* when used as a title with the meaning of ‘sir’, ‘doctor’, or ‘teacher’.

- (28) **tan** sian0-siN0
 Chen Mr.
 ‘Mr. Chen’

A final case of lexical idiosyncrasies is reported by (Tsay and Myers, 1996) as shown in (29), where the verb meaning ‘give’ normally conforms to the general tone sandhi pattern, as shown in (29a); however, it could be surfaced with an unexpected sandhi tone when preceding certain pronouns, as listed in (29b). The unusual tone surfaces obligatorily when the verb is followed by a third person pronoun, be it singular or plural (in fact, both have the same citation tone 55), while it appears optionally before pronouns of first and second person. Thus far, no other verbs are found to behave this way.

- (29) a. normal tone sandhi
- | | | |
|----------------------|--------------------|-------------------|
| <i>citation form</i> | <i>sandhi form</i> | |
| hO 33 ‘give’ | hO 21 kau51 | ‘give to the dog’ |

- b. idiosyncratic tone sandhi
- | | | | |
|----------------|---|-----------------------------|--------------------------|
| free variation | { | hO 21 goa51 | ‘give to me’ |
| | | hO 33 goa 33 | |
| free variation | { | hO 21 li51 | ‘give to you’ |
| | | hO 33 li 33 | |
| | | *hO 21 i(n)55 | ‘give to him/her (them)’ |
| | | hO 33 i(n) 33 | |

1.3.3 Summary of the regular tone sandhi pattern

In spite of several lexical exceptions, the general tone sandhi behavior in TSM can be recapitulated as follows:

TS applies... (i.e., no sandhi domain boundary is inserted)

1. between a noun, an adjective, or a verb and its preceding modifier adjunct such as the adjective ‘lazy’ in (15a), the VP-adverb *chiah* ‘just’ in (11), and *moa-a* ‘sesame seeds’ as an adjunct of ‘big’ in (8b).
2. between a verb and the first object to its right, either D.O. or I.O. as in (13).
3. between a pronoun and the first word to its right, e.g. *i* ‘(s)he’ in (13).
4. between a negative particle and its following verb/adjective/adverb as in (18c).
5. between a matrix-clause verb and an infinitival complement in (19b).
6. between a matrix-clause verb that introduces indirect speech/thought/belief and its clausal complement in (20).
7. among the elements within a DP of the form [(demonstrative) + numeral +

classifier + NP] as in (19a).

TS is blocked... (i.e., a sandhi domain boundary exists)

1. between a non-pronominal subject and what follows to its right as in (17), (7b), and ‘sesame seeds’ as the subject of the CP in (8a).
2. between a preverbal object and the following verb as in (13a), (13c) and (19b).
3. between the modificational marker E and its preceding modifier, which can be a noun, an adjective as in (15b) and (17), or a clause as in (8a).
4. between two objects in a double-object structure as in (13b).
5. between a sentential adverb and the first word to its right as in (10).
6. across clausal adjunct boundaries as in (21).
7. between XPs in coordinate XP constructions such as the APs in (18).

1.4 Characteristics of tone sandhi in TSM

In Taiwan Southern Min, tone sandhi is a general change affecting any non-domain-final tone, irrespective of lexical category or contiguous tonal environment, and the domains of application are defined with reference to syntactic structure (Chen, 1987, Lin, 1994). In terms of the phonological nature, it has long been assumed impressionistically that the tonal alternation between citation and sandhi forms is categorical, meaning that citation X and sandhi X variants have the same phonetic properties, such as f₀, where X can be any contrastive tones excluding tone 24, which only sur-

faces at domain-final positions. Comparing the columns of citation and sandhi tones of the same tone category in Table 1.1, one can observe that there seems to be no clear connection between the corresponding citation and sandhi tones, i.e., no phonetic or phonological conditioning can completely explain the direction of the change (either having citation or sandhi variant as the underlying form), but phonetic arbitrariness and phonological opacity (Kiparsky, 1973). Since sandhi tones form a proper subset of the set of citation tones and no new tones are created in the sound change, the tone sandhi process itself is structure-preserving. And as illustrated above in Section 1.3, there exist cases of lexical idiosyncrasies that do not conform to the general tone sandhi. In addition, previous studies using wug tests such as Hsieh (1970) and Wang (1995), among others, show that tone sandhi rules in TSM are semi-productive according to the obtained poor correct response rates.

These attributes altogether is a mixture of properties of lexical and postlexical rules, which gives rise to the debate over the status of tone sandhi rules in TSM. The characteristics of lexical rules that TSM tone sandhi possesses include phonological categoricity, phonetic arbitrariness, structure preservation, semi-productivity, obligatory application, with lexical exceptions, and insensitivity to rate. At the same time, TSM tone sandhi applies to all categories and has the power to look outside the word to the phrasal context to determine whether a syllable is in sandhi or non-sandhi position — properties of postlexical rules.

1.5 Research questions

In a few words, tone sandhi phenomenon in TSM involves different parts of grammar: the tonal change itself is phonological in nature while the conditioning environment is sensitive to syntactic information (the syntax-phonology interface), and the potential case of tonal neutralization is related to the phonetics-phonology interface; thus, it provides us with a great avenue to study the interfaces among different components of the grammar. In this dissertation, two broad research questions are addressed:

- (1) Does the Tone Circle really exist? Are citation and sandhi variants of the same tone realized according to the assigned tonal values? Whether citation and sandhi variants that are taken to be identical on the surface realized the same in production? Can they be perceived as the same by native speakers? Does the realization of tone sandhi show any generational difference?
- (2) How is such a complex tone sandhi system represented in the grammar? In what ways do different components of the grammar interact during the tone sandhi process?

The answers for the questions above are expected to grasp an understanding of the bigger picture listed below.

- (3) What is the phonological representation of the tones in Taiwan Southern Min? Which form, citation or sandhi, is more underlying? Should the gradient phonetic properties be encoded in the phonological grammar?

- (4) How do native speakers acquire the circular tone sandhi in Taiwan Southern Min?

While answering the two groups of questions in (1) and (2), I also aim to answer the following questions to shed some light on less clearly understood aspects of the tone sandhi phenomena in TSM:

- (5) Whether the tonal neutralization case in TSM shows complete or incomplete neutralization in production? Can native speakers perceptually distinguish the two sandhi variants involved in tonal neutralization? Is there a generational difference in the realization of the tonal neutralization case?
- (6) How citation and sandhi variants are represented in the listeners' mental lexicon? Whether priming effects can be induced by primes that correspond to the citation or sandhi variant of the targets? Are listeners more sensitive to the surface representations or underlying representations in spoken word recognition?
- (7) Should tone sandhi in TSM be analyzed as a lexical or postlexical rule, or as a case of allomorphy? How to account for the fact that tone sandhi in TSM have access to nonlocal information about the syntactic structure?

The first two groups of questions are answered throughout Chapters 3, 4, 5, and 6. The fifth research question is answered in Chapter 3, and the sixth question's answer is discussed in Chapter 4. The last research question is answered in Chapter 6.

1.6 Outline of the dissertation

The remainder of this dissertation is organized as follows. Chapter 2 reviews previous studies on tonal neutralization involving tone sandhi phenomena, phonological priming with tones, and the circumscription of sandhi domains in Xiamen/TSM. In particular, Section 2.3 describes in detail existing algorithms showing that sandhi domains in TSM is syntactically determined. Chapter 3 presents an experimental study on whether the two sandhi 33 variants, one derived from citation tone 55 and the other from citation tone 24, are completely neutralized by native speakers in production and perception. In Chapter 4, I investigate whether priming effects can be induced by primes, which are the corresponding citation or sandhi tones of the targets' first syllable, and discuss how the priming result can shed light on the representation of citation and sandhi variants in the mental lexicon. In Chapter 5, I analyze a Taiwan Southern Min spontaneous speech corpus to examine the realization of the Tone Circle and the tonal neutralization case in spontaneous speech. Chapter 6 focuses on analyzing the circumscription of TSM sandhi domains in an articulated derivational model of the syntax-phonology interface developed in Pak (2008), which is based on the DM theoretical framework, and elaborates on how two such approaches — Pak's suggestion and my proposal — can account for the tone sandhi data in TSM, as well as how phonological rules, e.g. tone sandhi in this language, can have access to nonlocal information about syntactic structure. I argue that sandhi domains in

TSM are determined by a Head-left Concatenation rule and that tone sandhi is a phonologically context-free rule, rather than a case of allomorphy. In Chapter 7, I review the experimental results and the proposed theoretical analysis of tone sandhi rules in TSM, and present the overall discussion in relation to the research questions. Lastly, Chapter 8 concludes the dissertation.

Chapter 2

Literature Review

In this chapter, I present previous studies on topics relevant to the themes discussed in this dissertation: tonal neutralization, phonological priming, the circumscription of sandhi domains in TSM, and the status of TSM tone sandhi rules. In Section 2.1, I briefly describe previous studies on tonal neutralization cases involving tone sandhi phenomena, including the Mandarin T3 sandhi and the two tone sandhi rules in TSM, namely $T55 \rightarrow T33$ and $T24 \rightarrow T33$. Section 2.2 presents previous findings on phonological priming with tones. In Section 2.3, I provide details about two existing algorithms for the circumscription of sandhi domains in Xiamen/TSM, which are arguably the most successful analyses in the literature. Section 2.4 summarizes the debate on the status of tone sandhi rules in TSM, i.e., to regard them as lexical or postlexical rules.

2.1 Previous studies on neutralization

Phonological neutralizations have been proposed to be incomplete, posing a threat to formal phonological theory as to what phonological and phonetic information account for the mental representation of a lexical item. A strong claim comes from Port and Crawford (1989) studying on the acoustic contrast of German voiced and voiceless stops at syllable-final position. For example, a devoiced obstruent in *Bund* ‘group, association’ is said to become indistinguishable from its underlying voiceless congener, say the final stop consonant in *bunt* ‘colorful’. Their discriminant analysis and identification task, however, shows significant differences in both production and perception, supporting their idea that the final devoicing in German must be represented not only with a categorical phonological rule that signals the discrete change in the voicing feature, but also with a phonetic implementation rule that justifies continuously variable values of phonetic features. Other effects of incomplete neutralization have been found in the final devoicing in Catalan (Charles-Luce and Dinnsen, 1987), Polish (Slowiaczek and Dinnsen, 1985), Russian (Olga Dmitrieva and Sereno, 2010), Turkish (Rudin, 1980), and Dutch (Warner et al., 2004).

While most of the early neutralization studies have been focused exclusively on the encoding of segmental information, the research interest also extends into autosegmental information, allotones or tone sandhi phenomena, in particular. Peng (2000) investigates the Mandarin third tone sandhi in which tone 3 (a dipping tone

denoted with tone value 214¹) becomes tone 2 (a rising tone represented as 35) when followed by another tone 3. The production experiment finds marginally significant effect of tone type (underlying tone 2 vs. the sandhi tone) on the mean f0 values; nevertheless, native speakers in the identification experiment failed to perceive the small pitch difference found in the production experiment, suggesting that the sandhi tone is not completely neutralized with the underlying tone 2 acoustically but is perceptually neutralized to some degree such that the sandhi tone is indistinguishable from tone 2.

With the convoluted tone sandhi system, TSM serves as a great place for studying tonal neutralization, which is part of the longstanding puzzle in TSM concerning the realization of the Tone Circle. The traditional impressionistic assumption is that the tone alternation between citation and sandhi forms is categorical, often referred to as the categoricity of tone sandhi in TSM, e.g. citation X and sandhi X have the same phonetic properties such as F0, where X can be any contrastive tones excluding tone 24, as it only surfaces at domain-final positions. And as noted, sandhi tones form a proper subset of the set of citation tones, a phenomenon often attributed to structure preservation.

To test the categoricity hypothesis, Tsay et al. (1999) compare citation and sandhi forms that have been described as having the same surface tones and examine the apparent neutralization of the two distinct citation tones, 55 and 24, into a single

¹The numerical values in parentheses represent pitch height on a five-point scale introduced in Chao (1948), where 5 indicates the highest pitch and 1 the lowest.

sandhi tone 33. The analyzed data includes recordings of three male and four female subjects in their late 20's to early 30's. The designed material for the production task consists of five pairs of sentences with both citation and sandhi forms realized the same on surface as 55, 33, and 21 tones in the target syllable, as exemplified in (1) where both the citation *te* 'emperor' and the sandhi *te* 'earth' surface tone 21, and another three pairs covering the neutralization case as in (2), where both citation *si55* 'poetry' and citation *si24* 'times' have been described to surface as tone 33 at sandhi position.

- (1) a. hit e hong **-te#** kiu chin che (Tsay et al., 1999, p.2408(2-3))
 that CL loyal -emperor ball very many
 'That emperor has many balls.'
- b. hit liap ang **te-** kiu# chin tat chiN
 that CL red earth -ball very worth miney
 'That red globe is very valuable.'
- (2) a. gun tao ting tiong-kok **si55** -po (Tsay et al., 1999, (4-5))
 my home subscribe China poetry -report
 'At home, we subscribe to China Poetry Journal.'
- b. gun tao ting tiong-kok **si24** -po
 my home subscribe China times -report
 'At home, we subscribe to China Times.'

As shown in the examples, the target syllable always occurs approximately in the middle of the sentence. F0's at the beginning, middle, and end points of the citation and sandhi forms of the target syllable along with the syllable duration are compared using ANOVA analyses. Regarding the neutralization case, no significant difference

is found between the F0's of citation and sandhi forms at the three points of the target syllable, and hence they report the neutralization being complete. As to the F0 comparison between citation and sandhi variants that are said to be realized the same, significant difference is absent at the beginning and end points of the syllable, but present at the mid point, where the citation variant has lower F0 than the sandhi variant, which according to them is ascribed to what they termed as the Intonation Hypothesis, where the citation form might be lowered in F0 due to phrase-final declination. They concludes that the alternation between citation and sandhi tones is categorical, as indicated by the neutralization of the f0's of juncture and context forms, confirming the categoricity hypothesis.

Myers and Tsay (2008) employ minimal pairs of sentences to examine whether neutralization occurred in two conditions: (a) across-positionally (juncture vs. context), in which the target words occupy different positions in a sandhi domain, one in the juncture position (i.e. the position followed by a sandhi domain boundary, and the word is realized with its citation tone 33) while the other in the context position (namely, a position that does not precede a sandhi domain, and the word is realized as the sandhi tone 33 derived from citation tone 55), and (b) within-position (context vs. context), where the pairs of sentences have the same circumscription of sandhi domains, whereas the target words have different underlying tones (55 or 24) but the same surface tone 33. Discourse contexts (listener-absent or -present) are also taken into consideration. No significant effect of discourse context on duration and f0 is reported, and the difference in overall f0 between sandhi tones 33 derived from

citation tone 55 and citation tone 24 respectively is not only insignificant but also extremely small, a mere 2.3 Hz. The duration of sandhi 33 derived from citation 24 is significantly longer, 8 ms on average, than sandhi 33 derived from citation 55. No overall difference in slope between these two sandhi tones in context position is found, contrary to the claim that the processes giving rise to incomplete neutralization involve temporal adjustment of gestures and their relations (Zsiga, 1993, Port, 1996, Port and Leary, 2005).

2.2 Previous studies on priming experiments

Upon hearing some speech, listeners are able to process the acoustic-phonetic input, identify the intended words and understand the linguistic content in just an instant within a conversation in spite of the immense acoustic and phonetic variability in the signal (e.g. speaking rate, pitch, voice quality, and prosodic prominence). In general, linguists believe that spoken word recognition involves categorization of some acoustically different stimuli into instances of the same word, and different accounts have been proposed as to how words are mentally represented to allow for this complex categorization. The traditional account assumes a single highly abstract phonological representation for each word with non-distinctive phonological features unspecified which can accommodate any pronunciation variants in speech. Any speaker-specific information is regarded as a source of noise and has no effect on the word recognition process. Studies of recognition memory, Goldinger (1996, 1998) among others,

have shown that listeners implicitly retain nonlinguistic information such as characteristics of the talker's voice in memory, and thus bring forth the episodic account which considers lexical representations as ensembles of memory traces (or episodes) of word instances. These episodes could be detailed or abstract with varying degrees of strength in memory, which reflects the listener's frequency of exposure to each particular variant form.

In principle, it is possible to distinguish two functionally different processes involved in the perception of spoken words: *word recognition* and *lexical access*, the former of which is often defined as the pattern recognition process that allows a listener to identify a spoken stimulus as a word, where the latter generally refers to the process that mediates access to abstract knowledge (e.g., syntactic, semantic, pragmatic information) about a lexical entry (Pisoni and Luce, 1987). Studies on auditory word recognition indicate that the processing of a target word is influenced by the prior presentation of a phonologically related prime. A large number of experiments have used a priming procedure to examine the effects of phonological similarity on word recognition. For instance, in a lexical decision task, Jakimik et al. (1985) found facilitation to process auditory presented monosyllabic target words (either real words or nonwords) only when the polysyllabic primes are related phonologically and orthographically (e.g., facilitation is found for *message-mess*, but not for *definite-deaf*). While early studies primarily focus on phonological overlap in segmental information, a number of recent experiments have devoted to autosegmental information such as accents and tones.

Previous studies have used auditory-auditory priming lexical decision task to investigate the representation of context-conditioned allophonic tonal variants in Chinese languages. Zhou and Marslen-Wilson (1997) design two auditory-auditory priming lexical decision tasks to examine Mandarin disyllabic compounds that undergo T3 sandhi, where T3 is realized with a rising pitch contour denoted as T3V, similar to T2, when immediately followed by another T3. In the first experiment, targets are disyllabic T3T3 compounds that undergo T3 sandhi, while primes are disyllabic compounds of one of the three following forms: (1) T2TX, the first syllable of which matches the surface variant of the first syllable of the target, (2) T3TX, the first syllable of which has an underlying representation match with the first syllable of the T3T3 target, or (3) a control prime, where its first syllable has no overlap in either the surface or the underlying representation of the target's first syllable. The first syllable of the target and the first syllable of the non-control primes share segmental information, and here X refers to any of the four contrastive tones in Mandarin except for T3. Their results show that while T3TX primes have the shortest reaction time, T2TX primes have the longest and the reaction time of control primes is in-between. The former suggests that T3V (the first syllable of T3T3 target) = T3 (the first syllable of T3TX prime) and the latter implies that T3V \neq T2 (the first syllable of T2TX prime). Given their assumption that T3T3 undergoes categorical tone sandhi and changes into T2T3, the results are contradictory in that T2 = T3 but T2 \neq T2 / T3. In the second priming task, targets are T2TX that do not undergo T3 sandhi whereas primes are one of the following four types: (1) T2TX, the

first syllable of which matches both the surface and the underlying representation of the target’s first syllable, (2) T3T3, the first syllable of which undergoes T3 sandhi and turns into T3V similar to T2, (3) T3TX that does not undergo T3 sandhi, or (4) a control prime with its first syllable bearing a tone other than T2 and T3. The results show that control primes have the shortest reaction time while T3TX primes have the longest reaction time. In between the two extremes are the reaction times of T2TX and T3T3 primes, which are comparable to each other, suggesting that $T2 = T3V$. The significant difference in the reaction times of T2TX/T3T3 and T3TX primes indicates that $T2/T3V \neq T3$. While the results from the second task seem to support their categorical assumption that a T3 undergoing tone sandhi behaves just like a T2, there exist some potential confounding factors such as the contribution of the prime’s second syllable in the priming effect, the lexical integrity of Mandarin compound words, and some possible frequency effects.

Chien et al. (2016a), on the other hand, conduct an auditory-auditory priming experiment with a lexical decision task. While targets are disyllabic T3T3 that undergo T3 sandhi, primes are monosyllabic that takes one of the three forms: (1) a T2 prime that has a surface-tone overlap with the first syllable of the target, (2) a T3 prime which matches the underlying tone of the target’s first syllable, or (3) a control prime serving as the baseline condition has no overlap in either the surface or the underlying tone with the target’s first syllable. Among them T3 primes show significantly shorter reaction times than the other two prime types, the reaction times of which do not differ significantly. They thus conclude that surface matching

does not result in faster sandhi word recognition. The word frequency effect is also reported where high-frequency targets elicit significantly faster reaction times than low-frequency ones.

Chien et al. (2016b) to my knowledge is the first to study the tone sandhi in TSM using auditory-auditory priming. In the experiment, a lexical decision task is adopted to compare two sandhi rules in TSM, namely (1) T51 \rightarrow T55 and (2) T24 \rightarrow T33. These two rules are chosen for comparison since they differ dramatically in terms of their productivity in novel words — 80% application rate for the rule T51 \rightarrow T55 but only 40% for the latter rule T24 \rightarrow T33 as reported by Zhang et al. (2011). Take the first sandhi rule /T51TX/ \rightarrow [T55TX] as an example. Disyllabic /T51TX/ targets that undergo TS are preceded by one of the three monosyllabic primes: (1) a surface-tone prime, which has a surface-tone match with the first syllable of the target, T55 in this case, (2) an underlying-tone prime, which has an underlying-tone overlap with the target’s first syllable, here T51, or (3) an unrelated control prime, say T21. For the sandhi rule T51 \rightarrow T55, they report that surface-tone primes elicit a stronger facilitation effect than underlying-tone primes with sizeable frequency effects. For the other sandhi rule T24 \rightarrow T33, on the other hand, it is the underlying-tone primes that elicit a stronger facilitation effect than surface-tone primes. They conclude that the contradictory results may be due to the influence of sandhi productivity on participants’ sensitivity to the surface and underlying representations of tone sandhi words. Compared to the experimental design in Zhou and Marslen-Wilson (1997), the priming direction in this study is always one-way, i.e., the prime is always the

citation variant and the target is always the sandhi variant. It is therefore interesting to see if there is any priming effect when the direction is reversed, namely, whether a sandhi variant can prime its corresponding citation variant. The priming experiment presented in Chapter 4 tries to accommodate such prime-target pairs using a lexical exception case that does not conform to the general tone sandhi rules.

2.3 Previous studies on the circumscription of TSM sandhi domain

As described in Sections 1.2 and 1.3, tone sandhi in TSM is a phrase-level tonal alternation between corresponding citation and sandhi forms. As a matter of course, many efforts have been made to define the application domain of tone sandhi. Earlier attempts focus on tone sandhi in Xiamen — a Southern Min dialect to which TSM is most closely related, including Duanmu (1995)'s stress-based account involving destressing and rebracketing, Chiu (1931)'s syntax-based assumption of a one-to-one correspondence between sandhi domains and syntactic constituents, and Cheng (1970)'s suggestion that a tone sandhi domain boundary tends to co-occur with strong syntactic junctures such as that between subject and predicate. While contributing most of the major tone sandhi data in TSM, Cheng (1968, 1973) conducts the earliest comprehensive description of the sandhi phenomena under discussion and draws attention to the distinct tone sandhi patterns among different word classes, and claims that a sandhi domain is bounded at the right end of NP, VP, S (and S'), and sentence

AdvP. In that case, phonological rules such as tone sandhi in TSM would have to be sensitive to some distinction among categories, which, as pointed out in Chen (1987), do not even form a natural class. Later Chen (1987) and Lin (1994) respectively show that tone sandhi in Xiamen is a phrase-level phenomenon conditioned by syntactic structure. In the remaining part of the section, I will first describe in detail the two algorithms, which are arguably the most successful analyses in the literature that I am aware of, and then evaluate how the two approaches account for the tone sandhi behavior in TSM illustrated in Section 1.3.

As the first to analyze the syntactic conditions under which tone sandhi operates in Xiamen, Chen (1987) refers to the domain of tone sandhi application as a *Tone Group* (TG) and demonstrates that TGs cannot be defined by reference to phonological structures, but rather are syntactically determined phrase-level constituents. The proposed account called TGF is listed in (3).

The maximal domain condition in the first half of the TGF rule covers the range of categories to which Cheng refers (e.g. NP, VP, S, S', and AdvP), indicating that TGF is not sensitive to categorial distinction, as opposed to Cheng's view, but rather makes reference to categorial hierarchy. The second half of the rule concerns the distinct differences in tone sandhi behavior between arguments and adjuncts. As pointed out in Jackendoff (1977), among the two types of complements, functional arguments and modifiers or adjuncts, only the former are strictly subcategorized with respect to the head. The evidence Chen offers relates to the optionality of # in (4). In one reading, *Bi-kok* 'U.S.' serves as the subject of *chhut* 'make' within a relative clause,

and is therefore bounded by a sandhi domain boundary. In another reading, *Bi-kok* is interpreted as an adjunct of *chhut*, parallel to expressions like *Boston trained*, and hence tone sandhi applies between *Bi-kok* and *chhut*.

(3) *Tone Group Formation* (TGF) (Chen, 1987, 1992)

- a. Mark the right edge of every XP with # (a sandhi domain boundary), except where XP is an adjunct c-commanding its lexical head.²
- b. XP is an adjunct of Y, if XP
 - (i) appears in [...XP...]_{YP} and
 - (ii) is not a strictly subcategorized argument of Y.
- c. A c-commands B if the first branching node dominating A dominates B.

(4) Bi-kok (#) chhut # e khi-chhia
 U.S. make E car
 ‘U.S.-made cars’

While the first property in (3b) includes all possible complements, the second property restricts adjuncts only to those that are optional in regard to the meaning of the head, such as the ones referring to time, place, benefactive, instrumental and other oblique relations. Take (5), repeated from (8), for example. The NP *moa-a* ‘sesame seed’ functions as the argument of *toa* ‘big’ in (5a) but serves as the degree adjunct of the adjective *toa* in (5b); TGF therefore predicts a sandhi domain boundary following the NP only in the former case, as borne out by the facts.

²The word ‘lexical’ in (3a) is absent in the original version of TGF, but is added in his 1992 manuscript.

- (5) a. moa-a # toa # e sio-piaN
 sesame seed big E bun
 ‘bunds with big sesame seeds’
- b. moa-a = toa # e sio-piaN
 sesame seed big E bun
 ‘bunds as big as sesame seeds’ (‘tiny buns’)

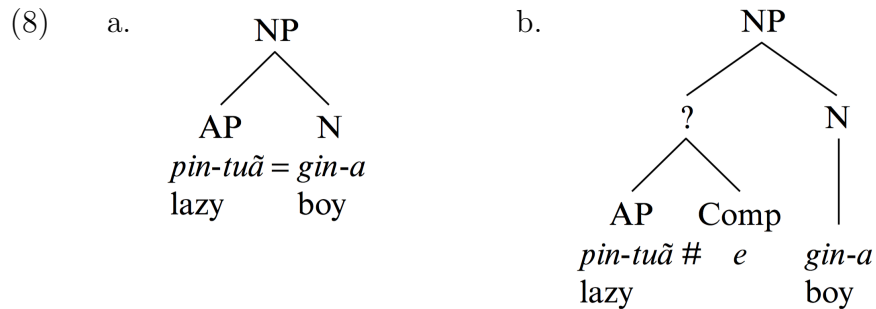
According to TGF, adjuncts are further distinguished into two types, and only the one that does not c-command a lexical head is marked with a sandhi domain boundary. This distinction corresponds with the asymmetry between sentential adverbs and VP adverbs with respect to tone sandhi behavior. Consider example (9), repeated below as (6). The VP adverb *loan-tsu* ‘mindlessly’ in (6a) as an adjunct lies within the VP and c-commands its lexical head V, so tone sandhi applies between the VP adverb and the following verb. The sentential adverb *ka-chai* ‘fortunately’ in (6b), on the other hand, is an adjunct outside the domain of V and the head it c-commands is Infl, a functional head; thus, its right margin corresponds to a sandhi domain boundary.

- (6) a. i loan-chu = kong
 he mindlessly talk
 ‘He is talking mindlessly.’
- b. gua ka-chai # che chit pan ki
 I fortunately take this Cl flight
 ‘Fortunately, I am taking this flight.’

Lastly, the c-command condition is motivated to account for cases involving the modificational marker *e* (also characterized as a complementizer, hence it is marked as *Comp* in (7b)) such as the contrast in (7), similar to (15), where a modifier AP does

not block tone sandhi when it joins directly to a nominal head, but induces a sandhi domain boundary when the particle *e* intervenes between the AP and the nominal head. The corresponding syntactic representations proposed by Chen are shown in (8). While the AP *pin-toaN* ‘lazy’ functions as the modifier of the nominal *gin-a* ‘boy’ in both cases, it is only in the first case that the head c-commanded by the AP is lexical rather than functional, and therefore TGF predicts that a sandhi boundary is absent between the AP and the nominal head in (7a) but follows the AP and precedes the grammatical particle *e* in (7b), as attested by the data.

- (7) a. *pin-toaN* = *gin-a*
 lazy boy
 ‘lazy boy(s)’
- b. *pin-toaN* # *e* *gin-a*
 lazy Comp boy
 ‘lazy boy(s)’



In the mean time, the c-command condition also explains the tone sandhi behavior of coordinate XPs construction such as (18), reproduced below as (9), where both conjuncts are marked with a sandhi domain boundary. As both APs do not c-command

the nominal head *gin-a* which they modify, TGF's prediction that each of these APs is bounded by a sandhi domain boundary is borne out.

- (9) chhong-beng # khun-lat # e = gni-a
 smart hard-working E child
 'a smart and hard-working child'

- (10) Xiamen Chinese Phrasing Parameter (Lin 1994: 248)

$]_{max}, X^{max}$ not lexically governed

- a. Government

A governs B iff A m-commands B and every barrier for B dominates A.

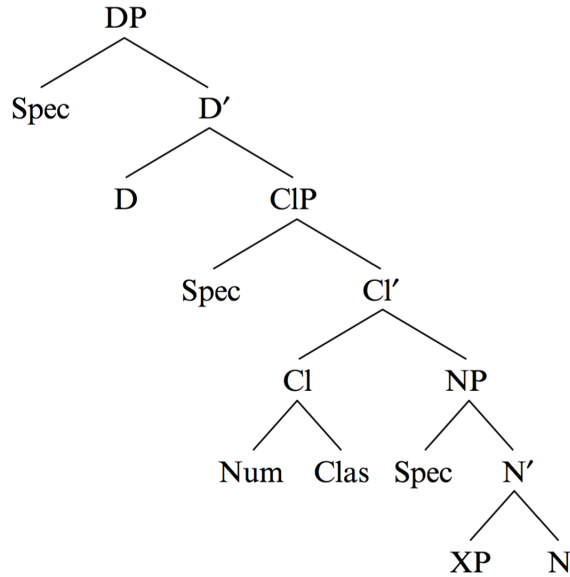
- b. A m-commands B iff A does not dominate B and every maximal projection M that dominates A dominates B.

- c. Barriers are defined roughly as follows: no maximal projection may appear between B and the maximal projection of A that dominates B.

Lin (1994), on the other hand, adopts the way how Hale and Selkirk (1987) define tone groups in Papago to determine tone sandhi domains in Xiamen in terms of lexical government, as listed in (10). Following Chomsky (1986), Lin assumes that categorial projections include IP and CP, of which the heads are Infl and Comp respectively. Based on the fact that noun phrases can be expanded as demonstrative + numeral + classifier + noun, he further embraces Tang (1990)'s proposal that the internal structure of noun phrases in Chinese should be fleshed out as (11), where Cl, Cl', and

ClP correspond to Q, Q', and QP respectively³, and Cl, or Q, is a functional category. ClP or QP is further ignored whenever its projection is irrelevant, in which case the NP complement is directly governed by the head D.

(11) Internal structure of DP in Chinese languages



While Chen's TGF refers to both structural (X^{max}) and functional (argument vs. adjunct) conditions, Lin's government-theoretic approach is based on a purely configurational property defined in terms of syntactic trees. First consider the contrast between sentential and VP adverbs shown in (6) with regard to their tone sandhi behavior. As VP adverbs are generally assumed to be adjoined to V', the VP adverb *loan-chu* 'mindlessly' in (6a) is by definition lexically governed by the head V, and thus is not marked with a sandhi domain boundary. Sentential adverbs, on the contrary, are generally assumed to be adjoined to I'; hence *ka-chai* 'fortunately' in

³Cl and Q, and correspondingly Cl' and Q', ClP and QP are used interchangeably in the literature.

(6b) lies outside of VP and is non-lexically governed by Infl, and its right margin is predicted to correspond to a sandhi domain boundary, a correct result.

Next, turn to prenominal modifiers such as the APs in (7). Given the assumption that prenominal APs are adjoined to N', it follows naturally that they are lexically governed by the head N, and thus no sandhi domain boundary is expected to occur between the AP *pin-toaN* 'lazy' and the following head noun *gin-a* 'boy' in (7a). For prenominal modifiers followed by the modificational marker *e* such as the AP in (7b), Lin postulates that the particle *e* as a functional category projects its own maximal projection called FP, substituting for the question mark in (8). Clearly, the AP is non-lexically governed by the functional head *e*, ergo a sandhi domain boundary is expected to follow the AP, as borne out by the facts. Similarly, as both APs in the coordinate construction in (9) are not lexically governed by any lexical category, both conjuncts are expected to be marked by a sandhi domain boundary, as corroborated by the attested facts.

Now reconsider the asymmetry between arguments and adjuncts in terms of their tone sandhi behavior shown in (5) in light of the lexical-government approach. As suggested by Chen, *moa-a* 'sesame seed' in both examples are interpreted as noun phrases, which should be projected to NP and DP according to the nominal structure in (11) Lin assumes. That said, *moa-a* in both cases would be bounded by # since it is non-lexically governed by the head D, which yields a correct result for (5a) but a wrong one for (5b). The first solution Lin offers is to regard the modifier *moa-a* in (5b) as an NP without projection to DP, which is motivated by Stowell's (1989)

claim that a predicational noun phrase might not project to DP in a discussion of the difference between referential and predicational noun phrases. As *moa-a* in (5b) is used as a metaphor and does not refer to actual sesame seeds, it is more like a predicational expression, and by analyzing it as an NP instead of DP⁴, *moa-a* is lexically governed by the adjective head *toa* ‘big’; thus no sandhi domain boundary is expected to follow it, as desired. Another alternative is to view *moa-a toa* in (5b) as an N + A compound, parallel to expressions like *thian-chin* ‘as innocent as heaven (lit. innocent)’. In this case, the compound should be treated as a lexical unit, which accounts for the obligatory tone sandhi of *moa-a*.

With the rationales behind TGF and the government-theoretic approach in mind, let’s review the examples discussed in Section 1.3 and see how both algorithms account for them. First consider the contrast in (7), reproduced below as (12). In (12a), *khah kin* ‘faster’ is a VP adverb c-commanding its lexical head *sia* ‘write’; thus, TGF predicts that there is no sandhi domain boundary between V and the adjunct. In contrast, *khah kin* in (12b) is interpreted as a predicate taking a sentential subject *i sia* ‘he writes’, and the subject as an argument is expected to be marked with a sandhi domain boundary, a correct result. The lexical-government approach accounts for the lack of # after the verb *sia* in (12a) by the fact that *sia* and its following

⁴A piece of evidence Lin uses to support that *moa-a* in (5b) does not project to DP relates to his claim that it cannot be preceded by demonstratives such as *hit* ‘that’, which is later compromised by example (84) in Chen (2000), shown below as (i).

- (i) [[hit [liap [moa-a]_{NP}]_{CIP}]_{DP} = toa]_{AP} # e sio-piaN
 that Cl sesame seed big Comp bun
 ‘a bun as big as that sesame seed’

adjunct together constitute the maximal projection. The sentential subject *i sia* in (12b) is non-lexically governed by Infl, and therefore is predicted to be followed by a sandhi domain boundary, as desired.

- (12) a. i [sia = khah kin]_{VP}
 he write more fast
 ‘He writes faster.’
- b. [i sia]_S # khah kin
 he write more fast
 ‘It would be faster for him to write.’

Consider the double-object constructions in (13), reproduced below as (13). In (13a) the verb and its following complements together constitute the maximal projection, and thus according to the structural condition in both TGF and the lexical-government approach, tone sandhi applies between the verb and the following object. As both the direct object and the indirect object in (13a) are arguments, and so do the fronted ones in (13b) and (13c), each of them is bounded by a sandhi domain boundary, in accordance with TGF. In comparison, objects as complements of V are lexically governed by V, thereby according to the government-theoretic approach their right margin should not correspond to a sandhi domain boundary, an ostensibly wrong prediction. In fact, the sandhi domain boundaries come from an inner complement of the argument. As noun phrases are assumed to be DPs with the structure shown in (11), the inner NP complement is non-lexically governed by the functional head D, and hence a sandhi domain boundary is expected at the right margin of the

NP, accounting for the attested facts.

- (13) a. i kai-siao chit e lau peng-iu # hO in bO
he introduce one Cl old friend to his wife
'He introduced an old friend to his wife.'
- b. i ka tang-oh # kai-siao chit e lu-peng-iu
he to schoolmate introduce one Cl girlfriend
'He introduced a girlfriend to his schoolmate.'
- c. i chiong hit pun chheh # sang hou tang-oh
he obj-marker that Cl book give to schoolmate
'He gave that book to his schoolmate.'

For the same reason stated above, in (19a), repeated as (14a), a sandhi domain boundary is found after the object *chit chhut liok-iaN-phiN* 'a video movie' but not between the verb *chO* 'rent' and the object, successfully accounted for by both accounts. (14b), however, presents a rather puzzling phenomenon, in which the original sandhi domain boundary after the object is ignored when the head noun *liok-iaN-phiN* 'video movie' is topicalized. According to TGF, the object containing the headless QP *chit chhut* as an argument should be marked with a sandhi domain boundary, just like the headless QP in subject position in (14c); Lin, by contrast, assumes $[Q[e]_{NP}]_{QP}$ as the relevant structure of the object, and since NP is non-lexically governed by the functional head Q, a # is predicted to follow the empty NP, contrary to fact.

- (14) a. chO chit chhut = liok-iaN-phiN # lai khoaN
rent one Cl video movie to watch
'rent a video movie to watch'

- b. liok-iaN-phiN # chO chit chhut = lai khoaN
 video movie rent one Cl to watch (same as (a))
- c. liok-iaN-phiN # chit chhut # chO loa-choe chiN
 video movie one Cl rent how much money
 ‘How much does a (cassette of) video movie rent for?’

Assuming that the prenominal QP in (14b) has the structure [QP *e*]_{NP}, where *e* is an empty category co-indexed with the topic, Chen proposes to reanalyze the QP as an adjunct *c*-commanding the lexical head *V*, such that no sandhi domain boundary is expected to follow it. He argues that the reanalysis is motivated by the fact that frequency adjuncts modifying verbs such as *chit pai* ‘one time’ in (15) have to be reanalyzed as a prenominal adjectival adjunct due to violation of Huang’s (1982) Phrase Structure Constraint, which specifies that Chinese phrase structure is left-branching, i.e. head-final, except on the lowest level of projection.⁵ Chen claims that the suggested reanalysis for (14b) is simply a mirror image of the restructuring instantiated by (15); that is, the adnominal QP in (14b) is reinterpreted as an adverbial, while the adverbial complement QP in (15) is reanalyzed as an adnominal complement.

- (15) khoaN chit pai = tian-yaN
 watch one time movie
 ‘watch a movie once’

⁵As an evidence to support the reanalysis of *chit pai* in (15), Chen brings up (i), where the frequency adverbial adjunct and its following head seem to move as a unit.

- (i) A-bing # liam chit pai tian-yaN # long bO khoaN
 A-bing even one time movie also not watch
 ‘A-bing hadn’t even watched the movie once.’

- (16) a. lin-gO tui chhiu-a ting chit liap chit liap lak loh0-lai0
 apple from tree top one Cl one Cl fall down
 ‘Apples fall on the ground one by one from the (apple) tree.’
- b. *lin-gO tui chhiu-a ting chit liap lin-gO chit liap lin-gO lak loh0-lai0
 apple from tree top one Cl apple one Cl apple fall down
 (same as (a))

The first solution Lin offers is to simply assume that a sandhi domain boundary at the right end of an empty category is invisible to phrase phonology. The second alternative is to treat expressions like *chit chhut* not as fully articulated noun phrases but as pure QPs without an NP complement, such as the one, i.e. *chit liap*, in (16a)⁶, where the numeral + classifier sequence interpreted as a manner adverbial cannot be followed by an NP complement, as exemplified in (16b). In this case, the QP in (13b) as the complement of the verb is lexically governed by the verb, and hence no sandhi domain boundary follows it, as predicted by the lexical-government approach.

- (17) a. Chhun-Kiau # siong-sin =_{CP}[A-Bin # boe lai]
 Chhun-Kiau believe A-Bin will not come
 ‘Chhun-Kiau believed that A-Bin would not come.’
- b. gua jin-ui =_{CP}[A-Bin # kong # e khah u to-li]
 I think A-Bin say E more have sense
 ‘I think what A-Bin said makes more sense.’
- (18) a. _{CP}[sia chit phiN bun-chiong # chin-cheng,] # gua su-iau khah
 write this Cl essay before I need more
 che gian-kiu
 many research
 ‘Before writing this essay, I need to do more research.’

⁶The example is borrowed from Lin’s example (35) on p.255. The circumscription of sandhi domains is not provided in the original example.

- b. gua chin thiam, #_{CP}[sO-i # lan lai tng#-khi0]
 I very tired so we let go back
 ‘I’m very tired, so let’s go home.’
- c. _{CP}[i na-si = boe = hiau tai-oan-oe] # li to ka i hoan-ek
 he if Neg know Taiwanese you then to him translate
 ‘If he doesn’t know Taiwanese, you translate for him.’
- d. _{CP}[i sui-lian # ni-ki # ia toa] # (put-ko #) sin-the # ia
 he although age quite big however body quite
 ho
 good
 ‘Although he is quite old, he is pretty healthy.’

Lastly, consider the structures in (20) and (21), reproduced as (17) and (18). In the former case, the finite clausal complements are arguments of the verb, thereby TGF successfully predicts there is no sandhi domain boundary between the VP and its following complement CP. The government-theoretic approach accounts for the lack of # after the verb by the fact that the clausal complements are lexically-governed by the head V. For biclausal structures in (18), where two clauses form separate sandhi domains, the first CP as an adjunct does not c-command a lexical head, so TGF predicts that its right margin corresponds to a sandhi domain boundary, a correct result. The government-theoretic approach, on the other hand, successfully locates a sandhi domain boundary between the two clauses in view of the fact that the first CP is non-lexically governed by the head Comp.

To summarize, while both Chen’s TGF and Lin’s government-theoretic approach account for a broad range of data as demonstrated above, the two differ in their proposed workarounds for problematic cases, such as the one presented by (14b), as

well as in the properties involved in their algorithms — while Chen’s TGF combines both structural (X^{max} and c-commanding) and functional (adjuncthood) conditions, Lin’s approach relies solely on configurational properties defined in terms of syntactic trees. Significantly, both analyses rested upon a framework based on some version of the prosodic hierarchy theory regard sandhi domains in TSM as prosodic constituents, but do not specify how exactly tone sandhi rules can access to these morphosyntactic information that is required for the circumscription of sandhi domains in TSM.

2.4 Previous studies on the status of tone sandhi rules in TSM

In generative phonology, multiple stages of phonological rule application are postulated: while lexical phonological rules operate in the lexicon and are confined to morphemes and single words, postlexical rules, also referred to as phrasal rules, apply across word-boundaries and have access to phrasal or syntactic structure. The difference in their domains of application suggests that lexical rules must apply before postlexical rules and that distinct properties are attributed to these two types of phonological rules, as shown in Table 2.1 (Kiparsky, 1985, et seq.). For instance, the velar softening of /k/ to [s] before [i] or [ai] in English is a lexical rule that applies across morphemes but within word boundaries, such that *electric* [ilɛkt.rɪk] and *critic* [kɪtɪk] become *electricity* [ilɛkt.rɪsəti] and *criticize* [kɪtɪsəɪz] respectively. Nasal assimilation in English, on the other hand, as a postlexical rule is feature changing

and optional. It assimilates nasals to the place of articulation of the following stops, and hence we get alternating forms such as [hʌŋgə] ~ [hʌŋgə] for *hunger* and [ɪm pɪnsəpəl] ~ [ɪm pɪnsəpəl] for *in principle*.

Table 2.1: Characteristics of lexical and postlexical rules.

Lexical rules:	Postlexical rules:
a. word-bounded	not word-bounded
b. access to word-internal structure assigned at the same level only	access to phrase structure only
cyclic	non-cyclic
c. apply in derived environments	apply across the board
d. structure-preserving	not (necessarily) structure-preserving
e. apply to lexical categories only	apply to all categories
f. may have exceptions	automatic
g. semi-productive	fully productive
h. categorical output	may have gradient output
i. obligatory	optional and may be sensitive to rate, register and pause

While the rule diagnostics in Table 2.1 as a guideline correctly accounts for a number of phonological data, its reliability has been challenged by many attested rule applications that share characteristics of both lexical and postlexical rules. For example, Sproat (1993) shows that /l/-darkening in English is sensitive to word-internal structure but is also gradient and non-structure-reserving. As a postlexical rule, English rhythem rule decreases its application as the number of intervening syllables increases. However, it applies cyclically and yields outputs that may be lexicalized (e.g. *ábstract* < *ábstract art*), and is blocked by lexical exceptions such as *exact* and *superb* (Kaisse, 1990).

As far as tone sandhi in TSM is concerned, the same conflicting result is obtained. As described in Section 1.4, it is widely believed among Chinese linguists that tone sandhi in TSM is structure-preserving, semi-productive, categorical, and obligatory; also, it makes lexical exceptions and is insensitive to rate — all properties of lexical rules. At the same time, it applies to all categories and has the power to look outside the word to the phrasal context to determine whether a syllable is in sandhi or non-sandhi position. The debate over the status of tone sandhi rules in TSM has thus attracted great attention, especially from those interested in the syntax–phonology interface.

The question that arises in light of such phenomenon is *how* exactly the phonological rules access to morphosyntactic structures. Two general views of the syntax–phonology interface have been proposed and differ in the way phonological rules operate on the syntax. The direct–reference approach, as the name suggests, claims that phonological operations apply directly to the morphosyntactic structure as it exists at the given stage in PF. On the other hand, the indirect–reference approach such as the Prosodic Hierarchy Theory (PHT) argues that phonological rules do not apply directly to syntactic structures, but instead refer to some kind of derived intermediate prosodic structure which contains a hierarchically organized set of prosodic constituents (Selkirk, 1986, et seq.), e.g., Phonological Phrase, Intonational Phrase, etc. Distinct levels of constituents in the Prosodic Hierarchy vary in size and in the way they are derived. Each level is argued to be generated from syntactic structure by a unique parameterized algorithm, e.g. Hale and Selkirk’s (1987) Papago Phrasing

Parameter which aligns the right ends of non-lexically governed maximal projections in syntactic representation with the right ends of tonal phrases in Papago⁷.

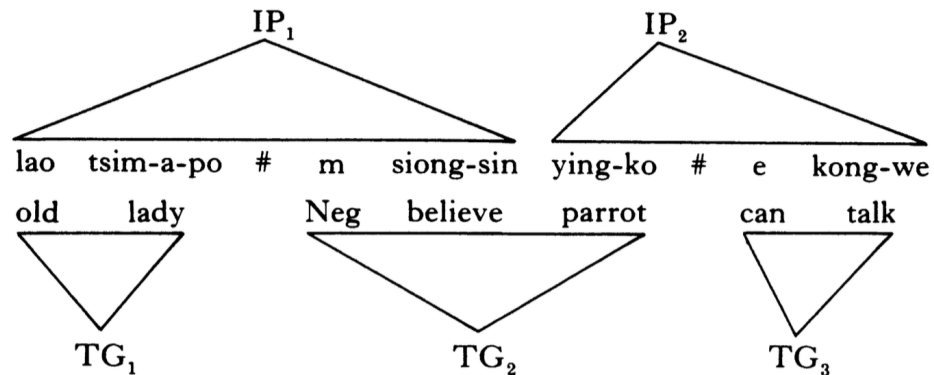
While Chen (1987) adopts PHT and regards tone group (TG) in Xiamen as a prosodic unit corresponding closely to the ‘phonological phrase’, intermediate between the clitic group and the intonational phrase (IP), he reports an interesting case of domain mismatches in (20) which violates PHT’s core tenets formalized as the Strict Layer Hypothesis in (19) (Selkirk, 1986, Hayes, 1989, among others).

(19) *Strict Layer Hypothesis* (SLH)

The categories of the Prosodic Hierarchy may be ranked in a sequence C_1, C_2, \dots, C_n , such that

- a. all segmental material is directly dominated by the category C_n , and
- b. for all categories $C_i, i \neq n, C_i$ directly dominates all and only constituents of the category C_{i+1} .

(20) ‘The old lady doesn’t believe that parrots can talk.’ (Chen, 1987:143)



⁷This exact phrasing parameter is later borrowed by Lin (1994) to account for the sandhi domain phrasing in Xiamen.

According to the hypothesis listed in (19), a lower-level prosodic unit is exhaustively contained within a higher level prosodic unit. Assuming that a tone sandhi domain (i.e., TG) is a phonological phrase ranked somewhere between intonational phrase (IP) and smaller units such as phonological word in the Prosodic Hierarchy, one would expect TGs to be fully contained within IPs, in accordance with SLH. However, (20) shows an intersecting relation between IPs and TGs. More specifically, while IP₁ contains TG₁ *lao tsim-a-po* ‘old lady’ and part of TG₂, namely, *m siong-sin* ‘not believe’, IP₂ consists of the remainder of TG₂ *ying-ko* ‘parrot’ and the entire TG₃. To put it another way, TG₂ spans across intonational phrase boundaries between IP₁ and IP₂, which is marked by the prolongation and/or pausing after the verb *siong-sin*. Note that hesitation pauses such as the one between the two IPs in (20) do not affect tone sandhi application. Which form is realized on the pre-pause syllable fully depends on the syntactic structure.

Hayes (1990) advances the Precompilation Theory in order to account for all phenomena of phrasal phonology that PHT fails to explain. In this model, rules that are sensitive to morphosyntactic information such as French liaison⁸ are treated as allomorphic rules termed precompiled rules, which are a subset of lexical rules stored in the lexicon. The rules then generate multiple diacritically-marked allomorphs lexically such that they are expected to have the other characteristic properties of lexical rules, e.g. structure-preservation and insensitivity to rate. Later at the interface of

⁸In colloquial style, liaison is frequent or obligatory between prenominal adjectives/quantifiers and nouns, but much less frequent between nouns and postnominal adjectives and impossible between an adverb and a following verb/adjective (Selkirk, 1986, Pak, 2008).

syntax and phrasal phonology, the derived allomorphs are inserted in the relevant syntactic contexts. To put it simply, two separate mechanisms are postulated in this theory: while the generation of allomorphs occurs lexically, the selection of allomorphs for insertion into syntax occurs postlexically.

In the case of tone sandhi in Xiamen, Hayes argues that tone sandhi does not refer to a prosodic domain and therefore the overlap of TGs with IPs (e.g. the case in (20)) does not contradict the Strict Layer Hypothesis. Based on Chen's original TGF, he claims that tone sandhi as a precompiled rule inserts 'citation' allomorphs at the right edges of non-adjunct maximal projections and 'sandhi' allomorphs elsewhere. Note that such proposals require the standard locality conditions on allomorphy to be relaxed to a great extent in that "...lexical insertion frames can not only 'see' linearly adjacent words but also have access to nonlocal information about syntactic structure such as what kind of XP contains the allomorph, and whether that XP is incorporated in the rest of the structure as an argument or an adjunct, thus admitting a host of unattested alternations like the following: (a) Insert allomorph X at the right edge of a clausal adjunct, if the following word is a vowel-initial adjective. (b) Insert allomorph Y at the left edge of a parenthetical XP, if the word preceding the parenthetical is a bimoraic noun" (Pak, 2008, p. 246).

Tsay and Myers (1996) try to defend the Precompilation Theory by showing that both allomorph generation and allomorph selection are necessary and independent via the demonstration of double dissociation in Table 2.2. Considering that tone sandhi in TSM shows sensitivity to syntactic structure (Chen, 1987, Lin, 1994) while possess-

ing other properties of lexical rules such as lexical idiosyncrasies, semi-productivity (Hsieh, 1970, Wang, 1995, among others), and phonological categoricity (Peng, 1997, Tsay et al., 1999, Myers and Tsay, 2008), they argue that tone sandhi in TSM can be accounted for by Hayes’s theory but that it involves only allomorph selection and not allomorph generation; namely, lexical insertion would choose between allomorphs that are essentially listed in the lexicon. Since tone sandhi affects every morpheme listed in the lexicon, the surface tone of a morpheme is determined solely based on the morpheme’s abstract tone-class diacritic.

Table 2.2: The logic of double dissociation.

Allomorph Generation	Allomorph Selection	Example
no	no	standard postlexical phonology
yes	no	standard lexical phonology
yes	yes	standard lexicalized phonology (e.g. Hausa final vowel shortening)
no	yes	English <i>a/an</i> allomorphy (Hayes, 1990), tone sandhi in TSM

Following this line of thought, based on the well-established fact that lexical frequency affects the retrieval of lexical items from memory (Jescheniak and Levelt, 1994, Meyers and Guy, 1997, among others), Chen et al. (2010) design a production experiment to test the lexicon-based hypothesis that tone sandhi in TSM is a process of allomorph selection between citation and sandhi tonal variants. Meanwhile, age effects (young vs. old), allomorph frequency, and positions of the morpheme are examined. In the experiment, 12 native speakers in each age group were asked to

produce the target morpheme Y in a carrier sentence [gua⁵⁵ be⁵⁵ khi⁵¹ XY+tiam²¹ # be⁵⁵ XY] ‘I am going to the XY shop to buy XY.’ when presented with a disyllabic word XY, and while both X and Y are actually occurring morphemes, the combination XY is either a real or a pseudo word. On account of their major findings, including (1) older speakers had significantly higher accuracy rate than younger speakers, (2) tone sandhi productivity was significantly influenced by allomorph frequency, and (3) subjects only seem to choose between citation and sandhi forms and rarely make non-sandhi errors, they conclude that both allomorphs are stored in the mental lexicon while there may be some asymmetry in the storage and/or retrieval of the two forms. However, note that this line of reasoning is only valid with the assumption that citation and sandhi forms are tonal allomorphs stored in the lexicon, and that from their results one can only deduce that frequency effect plays a great role in tone sandhi production and that speakers are aware of the two tonal alternations of a single morpheme.

Chapter 3

Tonal Neutralization

To examine the tonal neutralization case in TSM, a production experiment was conducted to see whether the two mid level sandhi tones 33, one derived from the high level citation tone 55 and the other from the low rising citation tone 24, are completely neutralized in the same context position. An identification experiment was carried out to test if native speakers of TSM can perceptually distinguish the two sandhi tones 33 derived from different citation variants.

3.1 Production experiment

3.1.1 Subjects and materials

13 native speakers of TSM, 6 males and 7 females, were recruited partially in Taiwan and partially at the University of Pennsylvania for this study. They were divided into two age groups, seven speakers, 4 males and 3 females, in the young speaker group (with an average of age 28.43) and six speakers, 2 males and 4 females, in the old

speaker group (with an average of age 58.83). All speakers were able to speak and read Mandarin.

Table 3.1: Subjects of the production experiment. The number in parentheses indicates subject’s age at the time of the experiment.

age group	old	young
male	CJS (66)	HCJ (25)
	ZJQ (60)	LCH (27)
		LJF (35)
		TSY (30)
female	CXY (57)	CSH (27)
	YMA (52)	TYJ (26)
	WCM (60)	WYW (29)
	WMA (58)	

The study included 10 target minimal pairs of monosyllabic words, all of which were embedded in minimal pairs of sentences, as exemplified in (1) and (2), where the target words are in bold and are marked with their citation tones, 24 and 55, but they are both realized as 33 on the surface, i. e., the sandhi tone. The complete list of stimuli can be found in Appendix B. Each of the stimuli sentences had a length of four to ten syllables with an average length of 6.23 syllables. All the pairs of target words were at sentence-medial position. A randomized reading list including 300 sentences ((10 minimal pairs \times 2 words + 10 fillers) \times 10 repetitions) was generated for each subject.

- (1) beh kio i **poe24** kau tang-si
 want ask him accompany until when
 ‘How long do you still want him to accompany you?’

- (2) beh kio i **poe55** kau tang-si
want ask him fly until when
'How long do you still want him to travel by flight?'

3.1.2 Procedure

During the recording, one experimenter was either on-site or remotely monitoring the experiment to provide detailed instructions and ensure that the participants understood and performed the task as expected. The randomized stimuli were presented one by one with Chinese characters on a computer screen to elicit subjects' production. Subjects were asked to first read the number of the order for each token followed by the stimuli sentence as naturally as possible, and they were allowed to rest at any time. The recording lasted for about 45 minutes with 300 tokens recorded from each speaker, 200 of which were the intended targets used in further analyses.

3.1.3 Data analysis

The nucleus of each target syllable was hand-labeled in Praat (Boersma and Weenink, 2015). The f0 values at every tenth and twelfth in time of the syllable nucleus were extracted with VoiceSauce (Shue et al., 2010) and converted into semitones with a reference of 100 Hz. Each f0 contour was further divided into four regions, and the mean f0 of each quartile as well as the overall mean f0 were calculated. Each f0 contour was therefore represented by five mean values for regression analysis. The data of the two age groups were analyzed separately in order to see the generational

difference.

As some tokens were pronounced differently from what were expected, they were excluded from the analysis. In total, 2,389 f0 contours, 1,258 tokens from the young speaker group and 1,131 tokens from the old speaker group, were analyzed with the repeated-measures linear mixed effects model and the Likelihood Ratio Test in R (R Core Team, 2014) using the *lme4* package (Bates et al.). Citation tone and gender were the fixed effects while subject and token were the random effects. A random slope model was adopted to account for inter-subject and inter-token variations. Interactions between citation tone and gender were also included. The resulting full and reduced models can be summarized as in (3).

- (3) a. Full: $\text{mean} \sim \text{citation tone} * \text{gender} + (1 + \text{citation tone} \mid \text{subject}) + (1 + \text{citation tone} \mid \text{token})$
- b. Reduced1: $\text{mean} \sim \text{citation tone} + \text{gender} + (1 + \text{citation tone} \mid \text{subject}) + (1 + \text{citation tone} \mid \text{token})$
- c. Reduced2: $\text{mean} \sim \text{gender} + (1 + \text{citation tone} \mid \text{subject}) + (1 + \text{citation tone} \mid \text{token})$

3.1.4 Results

Figure 3.1 illustrated the difference in the realizations of the sandhi tones 33 between different age groups. The solid lines indicate the f0 contours for tokens with tone 55

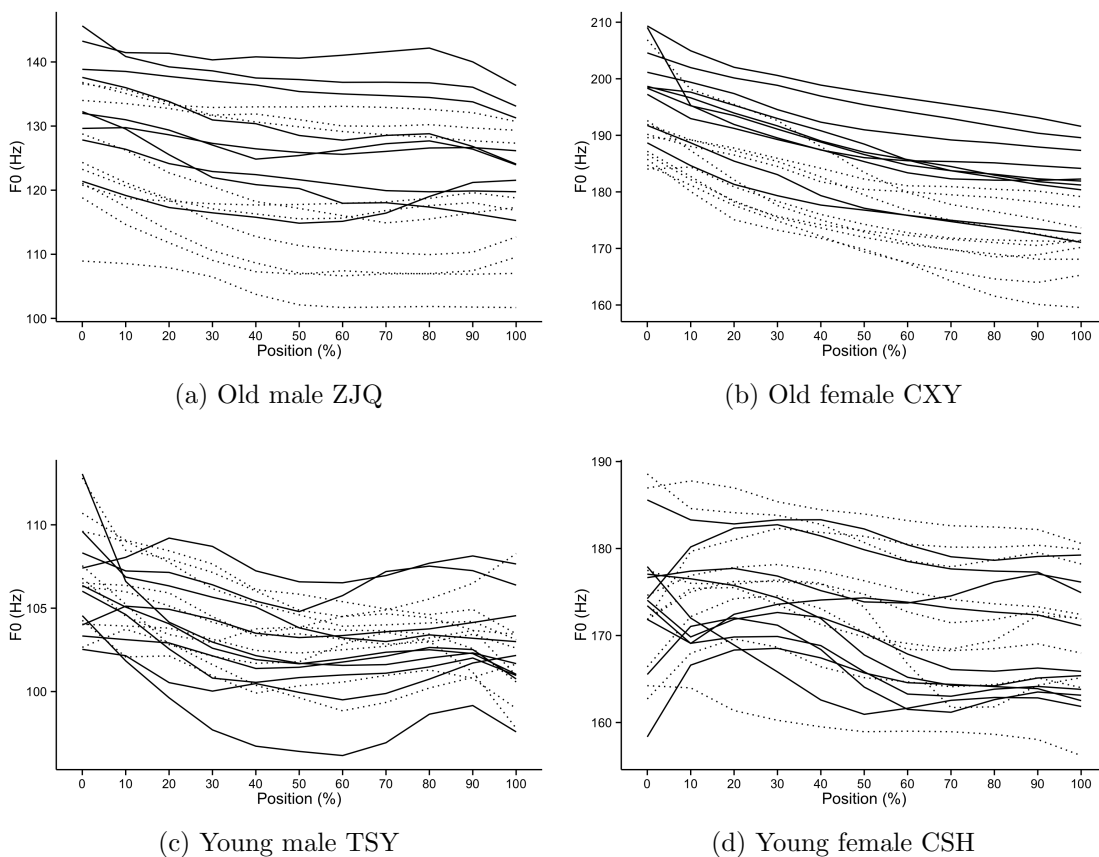


Figure 3.1: The f_0 contours of the target token “poe” produced by one male and one female speakers in each age group. The solid lines indicate the f_0 contours for tokens with tone 55 as the citation tone whereas the dotted lines refer to the f_0 contours of tokens with citation 24 tone.

as the citation tone whereas the dotted lines refer to the f_0 contours of tokens with citation tone 24. In Figures 3.1(a) and 3.1(b), most of the solid lines were above the dotted lines, suggesting that the old speakers tended to pronounce the target words of citation tone 55 with a higher pitch than those of citation tone 24. As compared with the lower two graphs, one could observe that the sandhi tones 33 derived from citation tones 55 and 24 produced by the old speakers were more distinct from each other than those produced by the young speakers, where more overlaps were found.

Table 3.2: The statistical result of the interaction between citation tone and gender for the young speaker group, with the reduced1 model compared with the full model.

Result	Q1	Q2	Q3	Q4	All
p-value	0.463	0.651	0.769	0.718	0.619
$\chi^2(1)$	0.538	0.205	0.086	0.131	0.247

Tables 3.2 and 3.3 showed the results of the Likelihood Ratio Test for the young speaker group. The former indicates that the interaction term between citation tone and gender for the young speakers is not significant throughout the entire f0 contour, and thus there is no interdependence between the two fixed effects. The latter table suggests that in all cases of the mean f0 of each target syllable among the young speakers, the citation tone shows no significant effect on the surface f0 values.

Table 3.3: The statistical result of the effect of citation tone on mean f0 values for the young speaker group, with the reduced2 model compared with the reduced1 model.

Result	Q1	Q2	Q3	Q4	All
p-value	0.130	0.111	0.068	0.066	0.069
$\chi^2(1)$	2.297	2.545	3.340	3.388	3.310

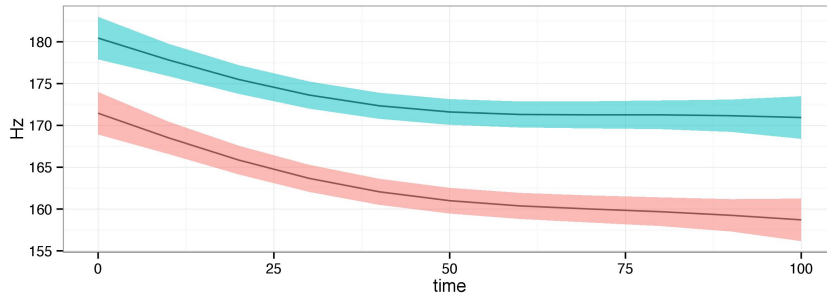
Table 3.4 shows the results of the Likelihood Ratio Test for the old speaker group. As shown in the table, the interaction term between citation tone and gender for the old speaker group is significant, suggesting that citation tone has a significant effect on all the mean f0 values that were somehow modulated through gender, namely, citation tone and gender were significantly inter-dependent on each other. The by-subject and by-token coefficients for the effect of citation tone are different for each subject and token as estimated in our random slopes model. However, the values are

always positive and that many of them are quite similar to each other, revealing that despite individual variation, there is consistency in how citation tone affected all the mean f0 values; that is, for all the old speakers, all five mean f0 values went up when the citation tone of the token was 55, but for some people they went up slightly more so than for others.

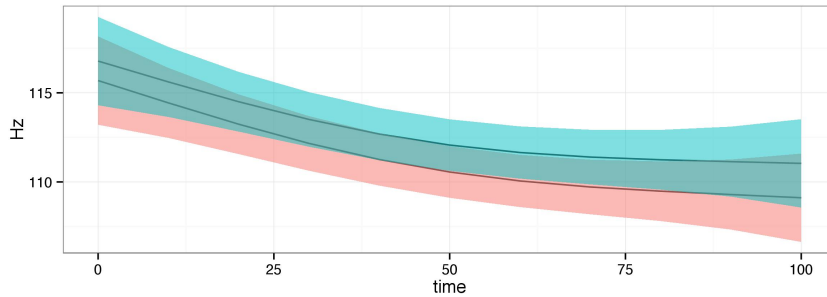
Table 3.4: The statistical result of the interaction between citation tone and gender for the old speaker group, with the reduced1 model compared with the full model.

Result	Q1	Q2	Q3	Q4	All
p-value	0.026*	0.006**	0.014*	0.018*	0.009**
$\chi^2(1)$	4.9603	7.405	6.035	5.574	6.829

In addition to the analysis with linear mixed effects models, each f0 contour was also used to build the smoothing spline analysis of variance (SS ANOVA) models implemented in R R Core Team (2014) with the *gss* package Gu et al. (2014). This method has been used to determine similarities and differences of multiple curve shapes (Gu, 2013, Wang, 2011), such as vowels’ formant contours, circadian rhythms (Wang et al., 2003), and tongue shapes from ultrasound imaging (Davidson, 2006). It has the advantage of taking entire curve shape into consideration and report statistical significance on the interaction term even when the difference lies only in a small portion of the curves. The reported Bayesian confidence intervals are used to determine at which point in the comparison the curves are statistically different. For more details on the methods and applications of SS-ANOVA, please refer to Gu (2013) and Wang (2011).



(a) old speakers

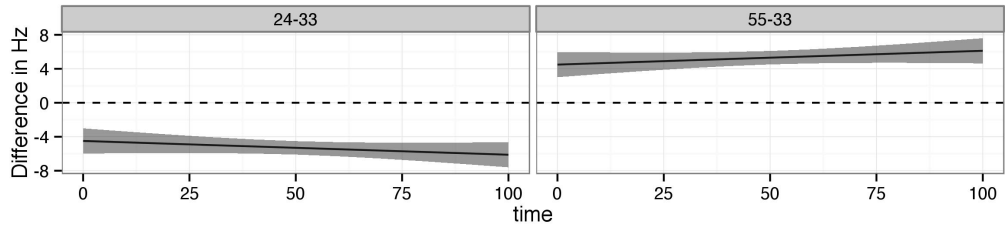


(b) young speakers

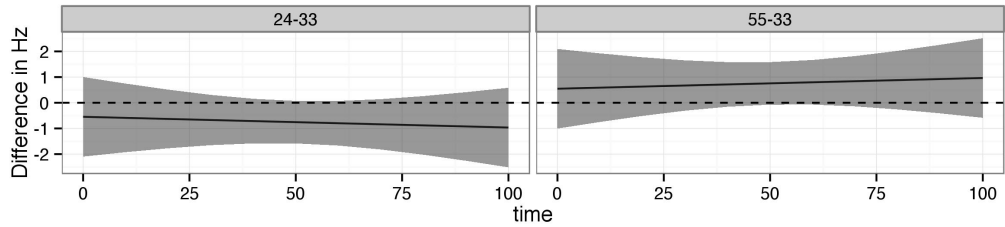
Figure 3.2: Smoothing spline estimates and 95% Bayesian confidence intervals for comparison of the two sandhi 33 f0 contours in each age group, where the one deriving from citation 24 is marked in red, and the one from citation 55 in green.

Figure 3.2 shows the 95% Bayesian confidence intervals around the smoothing splines for the main effects (here the two types of sandhi tones 33) curves. For the young speaker group, an overlap between the confidence intervals is observed throughout the entire curve in Figure 3.2(b), representing there is insignificant difference between the f0 contours of the two target sandhi tones. Figure 3.2(a), on the other hand, shows the opposite pattern, revealing that there is a significant difference between the two types of f0 contours produced by the old speakers. The result from SS ANOVA is perfectly in line with that from linear mixed effects models.

The interaction effects for the main effects curves are illustrated in Figure 3.3.



(a) old speakers



(b) young speakers

Figure 3.3: Interaction effects with Bayesian confidence intervals for the f_0 contours of the two target sandhi tones for both age groups, where the one deriving from citation 24 is on the left and the one from citation 55 on the right.

Again, as shown in Figure 3.3(b), there is no significant difference between the two f_0 contours produced by the young speakers, since the confidence intervals covered the zero on the y axis along the entire interaction curve. Figure 3.3(a) confirms the significant difference between the two f_0 contours produced by the old speakers.

3.2 Perception experiment

3.2.1 Subjects and stimuli

All speakers of the old age group from the production experiment, i.e., two males and four females, were recruited to participate in this preliminary perception study,

Table 3.5: Lexical frequency of the target tokens. Blue shades indicate the preferred token in that specific context out of subjective judgment.

token	<i>citation tone 55</i>		<i>citation tone 24</i>	
	occurrence	frequency	occurrence	frequency
choan	8	0.002%	77	0.019%
khoan	8	0.002%	3	0.0007%
kiaN	6	0.0015%	16	0.004%
koan	5	0.0013%	170	0.042%
kun	22	0.0054%	22	0.0054%
peng	7	0.0017%	6	0.0015%
poe	16	0.0039%	33	0.0082%
seng	341	0.0841%	1	0.0003%
thau	8	0.0019%	3	0.0007%
ti	3	0.0007%	NA	NA

as their production data shows significant f0 differences between the sandhi tones 33 derived from citation 55 and those derived from citation 24.

The identification test was generated with a self-written PHP script for each subject such that the participants could gain access to it online. The script was run on a browser on a computer with all the audio stimuli embedded in the webpage. The stimuli used in this task were extracted from the sentences with target tokens recorded by the subjects themselves during the production experiment and no fillers were included; therefore, there were ideally 200 sentence stimuli for each listener (10 minimal pairs \times 2 words \times 10 repetitions). Mispronounced sentences were further excluded, and the total number of stimuli for each subject varied according to their production result, ranging from 140 to 200 sentences with an average of 179.33 sentence stimuli per subject.

As it is hard to find minimal pairs of words as well as to embed them into the same

carrier sentences, preferences toward a certain context are expected. Where there is no clear preference between the two contexts of the minimal pair of sentences, lexical frequency is likely to be determinant. In view of the fact that no Taiwan Southern Min corpus was large enough to provide lexical frequency information and that most Chinese dialects share a common written form, the *Chinese Spoken Wordlist*¹ (Tseng, 2013) based on Taiwan Mandarin conversations was used for reference of lexical frequency. Table 3.5 shows the frequency information of the target tokens, where blue shaded cells indicate the preferred choice in that specific context out of subjective judgment. Take tokens *choan*, *kun*, and *peng* for example, given the contexts provided by the sentence stimuli, sentences with tokens of citation tone 55 are usually preferred to those of citation tone 24, while the reverse is true for tokens *seng*, *thau*, and *ti*. For cases with no evident context preference, a potential bias lies in the difference in lexical frequencies, which is relatively small for most cases except the one marked in red, where the lexical frequency of *koan24* is much larger than that of *koan55*, as shown in the table.

3.2.2 Procedure

Listeners ran through the trials individually in a quiet room, with the stimuli played to them through headphones. Subjects were able to replay the audio stimuli as many times as they wanted. In each trial, subjects were asked to identify an audio stimulus

¹The word list consisting of 16,683 word types and 405,435 word tokens, equivalent to 607,016 syllables, is based on natural conversations in Taiwan Mandarin collected and processed at Academia Sinica, with a total of 42 hours of speech recording. The recording took place from 2001 to 2003 and the speakers' age ranged from 14 to 63.

by choosing between the corresponding minimal pair of sentences, and to specify their confidence level when making the forced-choice decision on a 5-point Likert scale ranging from “not confident at all” (the leftmost) to “completely confident” (the rightmost). Figure 3.4 shows an example trial. The entire experiment took no more than 45 minutes.



Figure 3.4: An example trial of the perception experiment.

3.2.3 Data analysis

The ratio of the correct responses (e.g., the sentence with the target token of citation tone 55 was opted for when the audio stimulus contained the surface tone 33 derived from the target with citation tone 55) and the confidence level were calculated for each token and for each subject. To control the response bias, the non-parametric analog of d' , the measure A' (Grier, 1971), ranging from 0 to 1, was calculated for each listener using the formula shown in (4).

$$(4) \quad A' = 0.5 + [(x - y)(1 + x - y) / 4 x (1 - y)],$$

where x is the ratio of correct responses and y is the ratio of false alarms. An A' near 1.0 refers to good discriminability, while a value near 0.50 indicates chance performance.

3.2.4 Results

The correct response rate and the confidence level of each token for each subject are shown in Table 3.6. Notwithstanding the limited sample size, the general pattern indicates that the male subjects performed better than the female subjects, in particular the oldest male subject CJS. Note that where CJS has lower correct response rate, so do the other participants (especially for tokens *kiaN*, *koan*, and *thau*), indicating that the correct response rates varied according to the minimal pairs of tokens.

The cells highlighted in green denote that the results pattern to what is expected from the lexical frequency effect or context preferences. For instance, *choan* with citation 55 is preferred by the context, and thus in the data of the male subject CJS, higher correct response rate for citation 55 (=1) is observed, as opposed to that for citation 24 (=0.4). However, in the data of subject ZJQ, the difference in the correct response rates between the two sandhi tones is rather trivial (=1 vs. 0.9); hence the cells are highlighted in red, indicating that the opposite of what is expected is shown.

The confidence level was intended to see whether native speakers experienced any difficulty in distinguishing between the two tones in certain contexts, but the high

Table 3.6: The ratio of the correct responses and the confidence level of each token for each subject. CR stands for the correct response rate and CL the confidence level.

(a) male subject CJS					(b) male subject ZJQ				
token	<i>citation 55</i>		<i>citation 24</i>		token	<i>citation 55</i>		<i>citation 24</i>	
	CR	CL	CR	CL		CR	CL	CR	CL
choan	1	100%	0.4	100%	choan	1	86%	0.9	86%
kiaN	0.8	100%	0.9	100%	khoan	0	70%	0	70%
koan	0.8	92%	1	100%	kiaN	0	78%	0	88.89%
kun	1	100%	0.8	100%	koan	0	86%	0	88%
seng	0.9	100%	0.7	100%	kun	0.9	86%	1	90%
thau	0.4	100%	0.9	100%	peng	0.8	82%	1	84%
ti	1	100%	1	100%	poe	1	73.33%	0.8	82%
					seng	0.7	74%	0.9	74%
					thau	0	72%	0	76%
					ti	1	84%	0.9	86%

(c) female subject CXY					(d) female subject YMA				
token	<i>citation 55</i>		<i>citation 24</i>		token	<i>citation 55</i>		<i>citation 24</i>	
	CR	CL	CR	CL		CR	CL	CR	CL
choan	0.7	90%	0.125	85%	choan	0.1	82%	0.8	84%
khoan	0	82%	0	84%	khoan	0	82%	1	86%
kiaN	0	86%	0	92%	kiaN	0.7	84%	0.3	84%
koan	0	96%	0	92%	koan	0.7	78%	0.5	82%
kun	0.6	84%	0.6	86%	kun	1	86%	0	86%
poe	0.7	82%	0.5	94%	peng	0	86%	1	82%
seng	0	86%	0.9	96%	seng	0	86%	1	86%
thau	0	88%	0	86%	thau	0	84%	0.9	86%
ti	0.8	88%	0.3	86%	ti	0	84%	1	86%

(e) female subject WCM					(f) female subject WMA				
token	<i>citation 55</i>		<i>citation 24</i>		token	<i>citation 55</i>		<i>citation 24</i>	
	CR	CL	CR	CL		CR	CL	CR	CL
choan	0.9	90%	0	96%	choan	0.6	94%	0.4	88%
khoan	0.5	90%	0.9	94%	khoan	0	100%	0	96%
kiaN	0.6	92%	0.9	98%	kiaN	0	88%	0	86%
koan	0.1	92%	1	100%	koan	0	88%	0	86%
kun	1	100%	0.7	98%	kun	1	100%	0.1	100%
peng	0.4	96%	0.8	92%	peng	0.7	86%	0.3	84%
seng	0.4	88%	0.6	86%	poe	0.9	86%	0.2	88%
thau	0	100%	1	100%	seng	0.1	100%	0.9	100%
ti	0.3	88%	0.7	88%	thau	0	96%	0	90%
					ti	1	88%	0.1	92%

values everywhere (in fact it never fell below 70%) are of little avail in revealing inter-token differences, since where the correct response rates are low, the confidence level often remains high. The across-the-board high confidence level selected by the subjects in the perception experiment could be interpreted as either (1) the subjects believed that the difference between the two tones did exist but they might not be able to perceive it, or (2) the subjects generally assigned high values for the confidence level all the time as they were confident in whatever choices they made, which appears to be the case for the oldest male subject CJS.

Table 3.7: The overall correct response rates and the A prime statistics for each subject.

gender	subject	avg. correct response rate	A'
male	CJS	0.8286	0.8967
	ZJQ	0.5455	0.5834
female	CXY	0.2921	0.2059
	YMA	0.50	0.50
	WCM	0.60	0.6704
	WMA	0.315	0.2221

The overall ratio of correct responses along with the mean sensitivity index A' for each subject are listed in Table 3.7. The two columns pattern together and show that half of the subjects were able to distinguish the two sandhi 33 tones as their A' statistics are larger than 0.5. The female subject YMA performed at chance level, while the other two female subjects mixed the two sandhi variants and performed way below the chance level. In general, the male subjects outperformed the female subjects.

3.3 Discussion

In the production experiment, distinct patterns are observed for different age groups in terms of the effect of citation tone on the surface f0 values, revealing that there is an age-based acoustic variation. While young speakers do not maintain a pitch distinction between the two sandhi 33 variants, one derived from citation tone 55 and the other from citation tone 24, old speakers preserve the difference in the citation forms and consistently produce the sandhi 33 variant derived from citation 55 with higher pitch than that derived from citation 24.

In the perception experiment, the pattern of the correct response rates among different tokens was confounded with the lexical frequency effect and context preferences to some extent. As shown in Table 3.5, the word with citation tone 55 is preferred than the one with citation tone 24 for *choan* in the contexts of the stimuli sentences, hence higher correct response rates for words with citation tone 55 are expected, which is borne out by the data from subjects CJS, CXY, and WCM, as indicated by the green shades in the tables. For tokens such as *thau*, the reversed lexical preference is observed and thus words with citation tone 24 are predicted to have higher correct response rates, as corroborated by the data from subjects CJS, YMA, and WCM. However, there are cases conflicting with the expected pattern from lexical preferences, as shown by the red-shaded cells in the tables. Furthermore, despite the small sampling size, subjects' performance appears to correlate with their

age with reference to Table 3.1, suggesting that there might be another age effect in addition to the gender difference.

As incomplete neutralization is closely related to the phenomenon of near mergers, the age-based variations observed in both of the production and perception experiments may reflect an ongoing change towards a near- or complete merger, which is worthy of a longitudinal study. A possible implication of this tentative gender effect is in line with the prevalent observation made in sociolinguistics that women are more likely to lead the sound change. Assume that the complete neutralization is the trend and that the young speakers act towards a complete merger, the tonal neutralization in TSM seems to interact with not only age but also gender. This, however, is of course a provisional conclusion based on the limited subject size, and more subjects in both age groups are needed to verify this finding.

Chapter 4

Priming effects

There has been much debate about whether the circular tonal shift in TSM really exists, and if so, how these tonal variants are represented in the mental lexicon. In this chapter, an auditory-auditory priming experiment with a lexical decision task was conducted in order to investigate how citation and sandhi variants are represented in the listeners' mental lexicon. The relevant questions include whether priming effects can be induced by an autosegmental relation between the primes and the targets, e.g., an overlap in the underlying or surface tone, whether TSM listeners are more sensitive to the surface representations or underlying representations in spoken word recognition, and which form, citation or sandhi, is more underlying. In this experiment, the sandhi rule T51 \rightarrow T55 in TSM was investigated along with an exception to the general tone sandhi rules, such that both citation and sandhi variants can surface at the first of the targets in the priming experiment.

4.1 Participants

Forty Taiwanese subjects participated in this experiment in Taipei, the capital city of Taiwan. All of them are Mandarin and Taiwan Southern Min bilingual speakers, and none of them had any reported language impairments. Their ages ranged from 20 to 75 years old with an average of 45.68 years old at the time of the experiment. They were further divided into two generation groups, young (less than 40 years old) and old (greater than or equal to 40 years old). Ten male and twelve female subjects were in the young generation group, while nine male and nine female subjects were in the old generation group. All participants were asked to report the percentage of their daily usage of TSM, denoted as TSM_percent, which ranged from 0% “never used” to 100% “always used”. For a bilingual speaker living in the capital city of Taiwan where Mandarin is the dominant language, a value of 50% is considered high.

4.2 Stimuli

The design of the experimental stimuli is shown in Table 4.1. Thirty-six disyllabic words were selected as critical targets. Half of them, referred to as sandhi targets, has their first syllable undergo the tone sandhi rule and surfaces with sandhi tone 55. The other half of them, denoted as citation targets, are of the form /T51e24/ and exceptional to the general tone sandhi rules. This phrase structure involves a monosyllabic adjective or verb occurring before the modificational particle *e24* ‘of,

being’, and the first syllable is realized with its citation variant, e.g., the adjective ang24 means ‘red’, and /ang24 e24/ realized as [ang24 e33] means ‘something red.’ Following the experimental design in Chien et al. (2016b), in each critical trial, the prime and the target’s first syllable were segmentally (near-)identical, like a (quasi-)minimal pair, but might differ in tones. These words were devised with reference to two online TSM frequent word lists ¹ and all are frequently used in daily life conversation based on the author’s own judgment. An additional subjective familiarity rating task was conducted after the priming experiment, in which participants were asked to give subjective familiarity ratings for the thirty-six critical targets. The ratings range from 1 representing “never heard or said” to 7 “very often heard or said”. The purpose of this task was to make sure that all critical targets are frequently used in their daily conversation such that limited frequency effect would be expected.

Table 4.1: The stimuli design of the priming experiment.

Monosyllabic prime types	Disyllabic target types
s55 /T55/ (sandhi variant)	/T51e24/ → [T51 e33] (citation target)
u51 /T51/ (underlying- & surface-tone overlap)	
control /TX/ (unrelated control)	
s55 /T55/ (surface-tone overlap)	/T51TX/ → [T55 TX] (sandhi target)
u51 /T51/ (underlying-tone overlap)	
control /TX/ (unrelated control)	

Assume that citation tones are the underlying tones. Each of the targets was preceded by one of the following three types of primes, the first two of which shared

¹The links to the online Taiwan Southern Min word lists are http://twblg.dict.edu.tw/holodict_new/index/cimu_level1.jsp and <http://web.ffjh.tyc.edu.tw/tai5gi2/70.htm>.

phonological information with the target's first syllable: (1) *s55* prime: a monosyllabic word with citation tone 55 that corresponds to the sandhi variant of the target's first syllable in the citation target case and matches the surface tone of the target's first syllable in the sandhi target case, (2) *u51* prime: a monosyllabic word with citation tone 51 that matches both the underlying and the surface tones with the target's first syllable in the citation target case but overlaps only in the underlying tone with the target's first syllable in the sandhi target case, and (3) *control* prime: an unrelated monosyllabic control prime with a citation tone that is neither 55 nor 51. The monosyllabic primes are all real morphemes in TSM.

According to the stimuli design, in the case of citation targets, targets' first syllables remain in the citation tone 51 in that target phrases are exceptional to regular tone sandhi. With control primes as the reference, for instance, if *s55* primes have shorter/longer reaction time, the facilitation/inhibition of processing may result from the association between sandhi and citation tones invoked by the primes, which correspond to the sandhi variant of targets' first syllables. If it is the *u51* primes that induce shorter/longer reaction time, the facilitation/inhibition is likely to arise from the overlap in both the underlying and surface tones between primes and targets' first syllables. Same reasonin applies to the sandhi target case, where targets' first syllables with 51 as the citation tone undergo tone sandhi and are realized with the sandhi tone 55 on surface. Using control primes as the reference, for example, if *s55* primes induce shorter/longer reaction time, one could argue that facilitation/inhibition of target processing may be due to the overlap in surface tone between primes and tar-

gets' first syllables; and if u51 primes yield shorter/longer reaction time, one would suspect that facilitation/inhibition of processing is caused by the overlap in underlying tone between primes and targets' first syllables.

In addition to the critical targets, sixty disyllabic real words and ninety-six disyllabic nonwords were included as fillers, all preceded by real monosyllabic morpheme primes. Numbers of each of the five free tones are balanced across the critical targets, filler words, and nonwords. All stimuli were produced by a 69-year-old male native TSM speaker.

4.3 Procedure

The priming experiment was conducted using the PsychoPy program, an open-source application that is good for precise timing of responses. The procedure of the experiment is shown in Figure 4.1. Eight practice trials were presented first for participants to get familiar with the procedure and meanwhile to adjust the volume of the stimuli, and then the 192 main trials followed. Among them, the 36 critical trials were presented with a Latin Square design, such that each participant only heard a critical target once preceded by its corresponding prime, i.e., one of the three primes. The remaining 156 trials were shared across all participants. Total 192 trials (prime-target pairs) were randomly presented over headphones.

In each trial, a monosyllabic auditory prime was presented to subjects. Then a disyllabic audio target, either a word or a nonword, was presented 250 ms after the

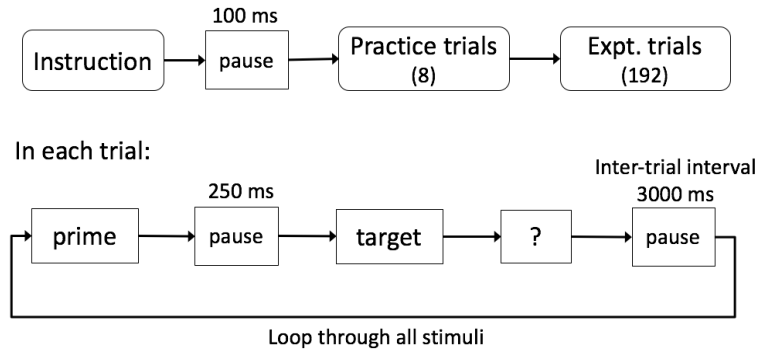


Figure 4.1: The procedure of the priming experiment.

offset of the prime. The subjects' task was to judge as quickly and accurately as possible whether the disyllabic target was a real word or not by clicking the left key, representing “yes”, or the right key, representing “no”. The inter-trial interval (ITI) was 3000 ms. The total duration of the experiment (including the lexical decision task and the familiarity judgment task) was about 25 minutes.

4.4 Data analysis and results

Statistical analyses were conducted on reaction times (i.e., the time duration between the presentation of a target to the key-press response from subjects.) obtained from the lexical decision task. Across all forty participants, the overall mean error rate was 11.74% (SD = 5.35) (902 trials/7680 trials). Only the critical trials were included for the reaction time analyses, and incorrect responses (3.68%) (53 trials/1440 trials) and responses over two standard deviations (5%) (72 trials/1440 trials) were excluded. The overall mean of self-reported TSM_percent is 39.25%, and the mean value for old subjects is 48.06% and that for young subjects is 32.05%.

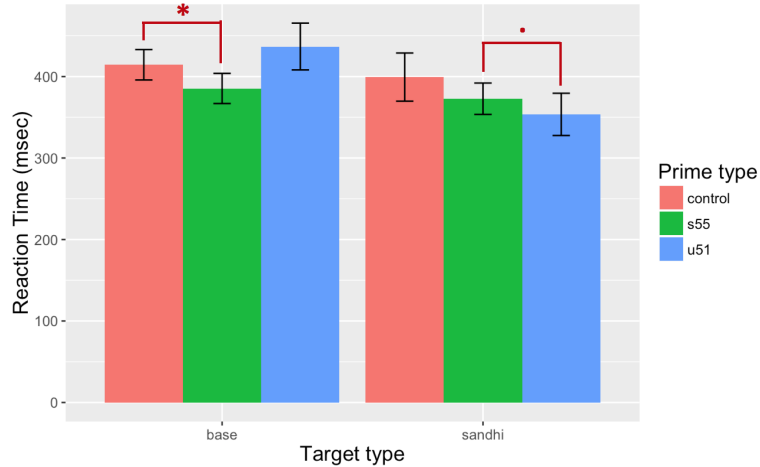


Figure 4.2: Mean reaction time of all subjects. Result of the citation target case is on the left, and that of the sandhi target case on the right.

Figure 4.2 shows mean reaction time of all the subjects for each prime type in two different target cases. The error bars represent one standard error above and below the mean. In the citation target case, u51 primes had the longest mean reaction time suggesting a potential inhibitory priming effect, whereas s55 primes yielded the shortest mean reaction time indicating a possible facilitatory priming effect. The mean reaction time with control primes as a reference lies in between the two. Reaction times with s55 primes and control primes are significantly different from each other ($\beta=-128.10$, $SE=55.01$, $t=-2.329$, $p=0.0202$). In the sandhi target case, u51 primes induced the shortest mean reaction time, followed by s55 primes, while control primes had the longest mean reaction time. However, the differences among the three are not significant. Reaction times with s55 primes and u51 primes are only marginally significantly different from each other ($\beta=163.67$, $SE=87.66$, $t=1.867$, $p=0.0623$).

Figure 4.3 shows the comparison between subjects of different generation groups.

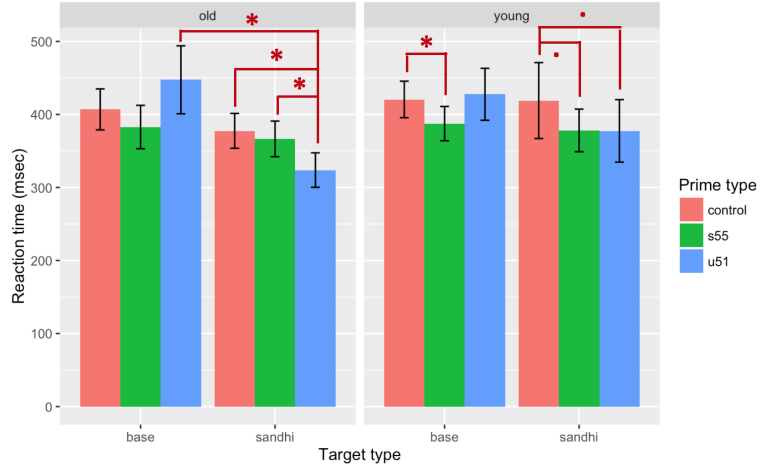


Figure 4.3: Mean reaction time of different generation groups. In each panel, result of the citation target case is on the left, and that of the sandhi target case on the right.

The result of the old age group is shown on the left while that of the young subject group is on the right. In the case of citation targets, i.e. on the left of each panel, the results of old subjects and young subjects pattern together. However, for the old age group reaction times with the three types of primes do not differ significantly; for young subjects, s55 primes (the green bar) elicited significantly shorter reaction times than control primes, the red bar ($\beta=-165.06$, $SE=75.71$, $t=-2.18$, $p=0.0299$), suggesting a facilitatory priming effect. In the case of sandhi targets, for old subjects, u51 primes (the blue bar) yielded the shortest reaction times, which were significantly shorter than those with s55 primes ($\beta=334.67$, $SE=147.86$, $t=2.263$, $p=0.024287$) and with control primes ($\beta=311.61$, $SE=147.86$, $t=2.107$, $p=0.035863$) respectively. The difference between reaction times with s55 primes and control primes is not significant. For young subjects, there are only marginally significant differences between reaction times with control primes and u51 primes ($\beta=-178.4$, $SE=102.8$, $t=-1.734$, $p=0.0837$)

and between reaction times with control primes and s55 primes ($\beta=-199.1$, $SE=102.8$, $t=-1.936$, $p=0.0536$). Note that for old subjects, the difference between reaction times with u51 primes in two target cases is also significant ($\beta=353.57$, $SE=169.78$, $t=2.083$, $p=0.0377$).

A series of linear mixed-effects models were applied to participants' log-transformed reaction times using the lme4 package in R. The dependent variable was the log-transformed reaction time (logRT), and there were prime type, target type, and generation as categorical independent variables, and TSM_percent and rating as continuous independent variables. Participant, item, and gender were considered as random variables. In total sixteen models were created and compared using the Likelihood ratio tests. The full model and the resulting final model are formulated in (1). The results showed that the interaction among prime type, TSM_percent, and generation was significant ($\chi^2=6.2308$, $df=2$, $p=0.04436$ *), and so were the two random variables, participant ($\chi^2=132.59$, $df=1$, $p=< 2.2e-16$ ***) and item ($\chi^2=23.855$, $df=1$, $p=1.039e-06$ ***).

- (1) a. Full: $\log RT \sim \text{targetType} * \text{primeType} * \text{TSM_percent} * \text{generation} * \text{rating} + (1|\text{participant}) + (1|\text{item}) + (1|\text{gender})$
- b. Final: $\log RT \sim \text{primeType} * \text{TSM_percent} * \text{generation} + (1|\text{participant}) + (1|\text{item})$

The data of the two generation groups were further separately analyzed. Series

Table 4.2: Results of a linear mixed models analysis for old subjects.

	Estimate	Std. Error	<i>t</i> -value
(Intercept)	5.57683	0.28894	19.301
primeTypes55	-0.45561	0.23253	-1.959
primeTypeu51	0.01194	0.23652	0.050
TSM_percent	0.17650	0.54384	0.325
primeTypes55:TSM_percent	0.86649	0.44386	1.952
primeTypeu51:TSM_percent	-0.35215	0.45195	-0.779

of linear mixed-effects models and Likelihood ratio tests were applied. The final model for old subjects was $\log RT \sim \text{primeType} * \text{TSM_percent} + (1|\text{participant}) + (1|\text{item})$, while the final model for young subjects was simply the intercept with two random variables: $\log RT \sim 1 + (1|\text{participant}) + (1|\text{item})$. Table 4.2 shows the results of the final model for old subjects. Generally speaking, as compared with control primes, reaction times with s55 primes were shorter (i.e., a facilitatory effect) while reaction times with u51 primes were longer (i.e., an inhibitory effect), reflecting what is observed in the citation target case in Figure 4.3. Meanwhile, reaction times increased when the TSM usage percentage increased, and the effect of prime types was modulated through the TSM usage percentage. With s55 primes, when TSM_percent increased a unit, reaction times increased by 0.87 log unit, but with u51 primes, the effect was reversed; that is, when TSM_percent increased a unit, reaction times decreased by 0.35 log unit. All these effects were only seen in the data of old subjects and not in the data of young subjects.

4.5 Discussion

Studies on auditory word recognition have shown that the processing of a target word is influenced by the prior presentation of a phonologically related prime. The aim of this experiment was to investigate whether priming effects can be induced by an autosegmental relation between the primes and the targets, e.g., an overlap in the underlying or surface tone, and to see how the priming effects, if any, can shed some light on the representation of citation and sandhi variants in the mental lexicon, such as the debate on which form, citation or sandhi, being the underlying form.

Regardless of some statistical insignificance, the experimental results show the following pattern, as observed in Figures 4.2 and 4.3: first, in the case of citation targets, for both young and old age groups, s55 primes that corresponded to the sandhi variant of targets' first syllables facilitated the recognition of target words, whereas an overlap in both the underlying and surface tones between primes and targets' first syllables inhibited target words' processing; Second, in the case of sandhi targets, however, an underlying-tone overlap between primes and targets' first syllables induced a facilitating priming effect for both generation groups, but the effect was larger for old subjects than for young subjects. An overlap in the surface tone between primes and targets' first syllables also had facilitation priming effect, and the difference between the facilitating priming effects of u51 and s55 primes was only significant for old subjects.

The two findings summarized above appear to be contradictory in the sense that while an overlap in both the underlying and surface tones (i.e., u51 primes) caused inhibition in the citation target case, an underlying-tone overlap and a surface-tone overlap individually produced a facilitation effect. A possible explanation for this conflicting result may be due to some problem of the stimuli design — some subjects commented after the experiment that it often took them longer to process a citation target out of context since this kind of phrases generally serve as modifiers and precede whatever they modify. The delay in reaction time thus could result from some overanalysis such as coming up with a proper context for the citation target.

In the citation target case, the fact that corresponding sandhi variants induced a facilitating priming effect on the targets reveals that subjects are aware of the association between sandhi and citation tones invoked by the primes. In the sandhi target case, the facilitation effect caused by primes with a surface-tone overlap is not surprising since the primes and the target's first syllables are phonologically (near-)identical. The facilitating priming effect produced by an underlying-tone overlap suggests that the association between sandhi and citation tones can be triggered as well with an input of the underlying tone. The fact that reaction times with u51 primes (underlying-tone overlap) were significantly shorter than those with s55 primes (surface-tone overlap) for old subjects only indicates that old subjects are much more familiar with the association between sandhi and citation tones in terms of the tone sandhi rule $T51 \rightarrow T55$, as compared to young subjects. It is very likely that young subjects may have a weaker notion of citation tones due to insufficient input. For

example, if a monosyllabic word only or mostly occurs in a sandhi context, i.e., it is always or mostly realized with its sandhi variant, speakers with this limited input may never know what the citation variant is.

Now compare the result of the current study with that in Chien et al. (2016b), where only the sandhi target case is included and instead of a generational difference, a gap in the lexical familiarity is taken into consideration. Their results show that surface-tone primes (primes that overlap in the surface tone with the targets' first syllables) elicited significantly stronger facilitation effect than underlying-tone primes while only the former yielded significantly faster reaction times than control primes, and that the role of the surface representation seemed to be modulated by familiarity in the sense that the facilitation effect induced by surface-tone primes were stronger when participants were highly familiar with the tone sandhi targets. Their reported pattern is different from what is observed in the current study, which may very well stem from the difference in experimental design. Nevertheless, both studies confirm that primes that overlap in the underlying or surface tone with targets' first syllables can facilitate the processing of target words.

Chapter 5

Corpus Analysis

In the last chapter, I showed that the tonal neutralization case in TSM involves a generational difference in that speech produced by young speakers exhibits complete neutralization between the two sandhi 33 tones derived from two different citation tones, 55 and 24, whereas incomplete neutralization is observed from old speakers' production data, and that at least some old subjects, male speakers in particular, are able to perceptually distinguish the two sandhi 33 variants. While this result may suggest a possible linguistic change in progress, it also brings up the question as to whether the Tone Circle really exists.

The purpose of this corpus study is to examine the realization of the Tone Circle in spontaneous speech. Relevant issues include the realization of citation and sandhi variants of the same tone, whether citation and sandhi variants that are supposed to correspond to each other are realized the same, if the realization of tone sandhi shows any generational difference, and whether the tonal neutralization case shows complete or incomplete neutralization in spontaneous speech. To address these concerns,

spontaneous speech data was extracted from *The Taiwan Southern Min Spontaneous Speech Corpus* established by Fon (2004) (hereafter the TSM corpus).

5.1 The Spontaneous Speech Corpus of Taiwan Southern Min

5.1.1 Speakers and materials

The TSM corpus was primarily constructed and annotated for studying discourse prosody. Speech was elicited in the form of an interview where the speaker was encouraged to carry on longer monologues instead of two-way conversation with the interviewer when asked to talk about his or her personal life experiences, which embraced a wide number of different topics, such as college life, marriage, health problems, and traveling experiences, to name but a few. Speech collected in such a natural setting is spontaneous and can exhibit a greater range of variability inherent in human languages than speech produced in a reading task or any other carefully designed laboratory task. The interviews were conducted by members of the Phonetics Lab at National Taiwan University and TSM speakers from different regions of Taiwan were recruited for this corpus. As a long-term ongoing research project, the corpus is partially annotated and segmented into syllables in Praat (Boersma and Weenink, 2015) while most of the recordings were transcribed in a separate Word file with a Taiwanese-Romanization convention in which both Chinese characters and

romanized transcription were used.¹

The annotated part of the dataset includes eight hours of speech contributed by sixteen native speakers of Taiwan Southern Min, and all of them were from the same region, namely Taichung, the mid-Taiwan metropolitan area. In this case, no dialectal differences would be expected to confound any findings from this corpus. According to their age, the sixteen speakers can be divided into two generation groups, young and old, and there are four male and four female speakers in each generation group. Speakers in the young group were born in the 1980s (age range was 20-35 when recorded in years 2003 to 2005, so for now these speakers would be around 32-47 years old) while speakers in the old group were born in the 1940s (age range was 50-65 when recorded). Syllable segmentation and the alignment between audios and romanized transcription were first automatically conducted with the EasyAlign plugin (Chen et al., 2011) and were further manually checked and tested for cross-labeler agreement with a mean difference of 12.4 ms in terms of the alignment of syllable boundaries, which amounts to 5.7% of mean syllable duration (Wang and Fon, 2013).

In addition to the annotated eight hours of speech, I transcribed another 43-minute long speech in one of the interviews and manually segmented it into syllables in Praat. In total, the dataset for this corpus study contains about 8.7 hours (523.2773 minutes) of speech. The duration of each interview varies as listed in Table 5.1. Most of the

¹The Taiwanese-Romanization convention is mainly based on the online dictionary (Iûⁿ, 2003). For transcribing the recordings, Chinese characters were used if it was possible to identify the source characters for a given Southern Min expression, and romanization was used instead if otherwise. For transcription aligned with speech in Praat, the Chinese characters were transformed into corresponding romanization.

audio files are about 30 minutes or longer, and the content of each recording varies from subject to subject.

Table 5.1: Overview of the TSM Corpus.

Subject	Gender	Generation	Duration (sec)
HSY	f	old	1926.23
LXQ	f	old	1800.52
LYY	f	old	1521.93
XJC	f	old	2041.99
LCX	m	old	2062.31
LZZ	m	old	1862.17
XJK	m	old	1890.18
XWZ	m	old	1871.61
CXH2	f	young	1845.16
HRT	f	young	1906.79
XYR	f	young	2008.16
ZWH	f	young	3170.00
CQH	m	young	1906.56
CXH	m	young	1895.00
CZY	m	young	1870.00
HSJ	m	young	1818.04

5.1.2 Methods

As the annotated dataset came with transcriptions, syllable segmentation, and labeling of lexical tones (citation tones), I first manually checked and corrected the alignments of syllable boundaries and the citation tone labels in Praat, and then specified whether a syllable undergoes TS or not. Extraction of syllable-wise f0 values was done with the `get_f0a` pitch extraction script on server *harris* with f0 range set at [75 Hz, 300 Hz] for male speakers and [100 Hz, 500 Hz] for female speakers. F0 values

within a syllable were resampled at every 10th of the interval, and syllables that are too short, i.e., less than 50 msec in duration, were excluded for analysis. To compare the pitch patterns of different age and gender groups, I performed log-transformation of the f0 values into semitones with each speakers baseline as the reference using the formula: $St = \log_2(f0/baseline)*12$. The baseline here refers to the 10th percentile of the speakers pitch range. Moreover, I annotated whether a syllable is prepausal in order to tease out the phrasal position effect when comparing the citation and sandhi variants that are supposed to correspond to each other.

5.2 Results

Table 5.2 exhibits the frequency distribution of the citation and sandhi variants of each tone category in this corpus. The dataset has in total 19,650 citation variants and 53,920 sandhi variants where words that are too short in duration are excluded. The frequency of the sandhi variants in this corpus data is 2.74 times more than that of the citation variants, and the frequency ratio for each recording ranges from 2.22 to 3.66 as shown in Table 5.3. For simplicity, I left out the two checked tones and focused on the five unchecked ones for further analysis.

Figure 5.1 shows the mean f0 curves of both citation and sandhi variants of each unchecked tone for the two age groups. The upper panel is the result for the young speakers and the lower panel for the old speakers. The red lines represent the citation variants while the green lines stand for the sandhi variants. The citation tone values

Table 5.2: Distribution of the citation and sandhi variants of each tone category in this dataset.

Tone category	Citation count	Frequency	Sandhi count	Frequency
55	4497	0.36	7827	0.64
51	5279	0.28	13673	0.72
21	1893	0.27	5085	0.73
21c	1069	0.13	7290	0.87
24	2917	0.29	7251	0.71
33	3022	0.23	9894	0.77
53c	973	0.25	2900	0.75
Total	19650		53920	

are listed at the very top, and the shaded grey areas around each mean f0 curve represents the 95% confidence interval. For instance, in the leftmost top and bottom graphs, the red lines denoting the high-level citation variant 55 always lie above the green lines designating the corresponding mid-level sandhi variant 33. Comparing each top and bottom pair of graphs, one can observe that the data of these two generation groups patterns together in a great deal.

5.2.1 Fitting Smoothing Spline ANOVA models

For statistical analyses, each f0 contour was represented by eleven values for building *Smoothing Spline Analysis of Variance* (SS-ANOVA) models implemented in R (R Core Team, 2014) using the *gss* package (Gu et al., 2014). Figure 5.2 shows the result of using SS-ANOVA to compare pitch contours of different variants of the same tone category produced by all speakers. Red lines mark the citation variants while green lines denote the sandhi variants. The solid lines are the resulting smoothing splines indicating the mean f0 contour in each case. The shaded areas are the corresponding

Table 5.3: Frequencies of citation and sandhi variants in each subject’s interview.

Subject	Citation count	Sandhi count	Frequency ratio (sandhi/citation)
HSY	1091	3141	2.879
LXQ	1426	3358	2.355
LYY	651	1864	2.863
XJC	1898	4625	2.437
LCX	975	2163	2.218
LZZ	1626	3985	2.451
XJK	1196	3620	3.027
XWZ	1194	2924	2.449
CXH2	1213	4177	3.444
HRT	1045	3004	2.875
XYR	1296	3777	2.914
ZWH	1793	4767	2.659
CQH	1006	2477	2.462
CXH	1500	4572	3.048
CZY	1051	2948	2.805
HSJ	689	2518	3.655

95% Bayesian confidence intervals. An overlap between the confidence intervals represents insignificant difference between the f0 contours. As shown in the graph, for each tone category there is no overlap between the confidence intervals of the citation and the sandhi variants except for crossover of the two curves, suggesting that the two variants are significantly different from each other for all five tone categories.

In order to see whether there is a generational difference in the pitch contours, I separated the data and compared the f0 curves of two generation groups. Figure 5.3 shows the comparison of the f0 curves of the corresponding pair of citation and sandhi variants for each tone category between different age groups using SS-ANOVA. Red lines indicate the citation variants and green lines represent the sandhi variants. In each case, the solid lines are the resulting smoothing splines indicating the mean

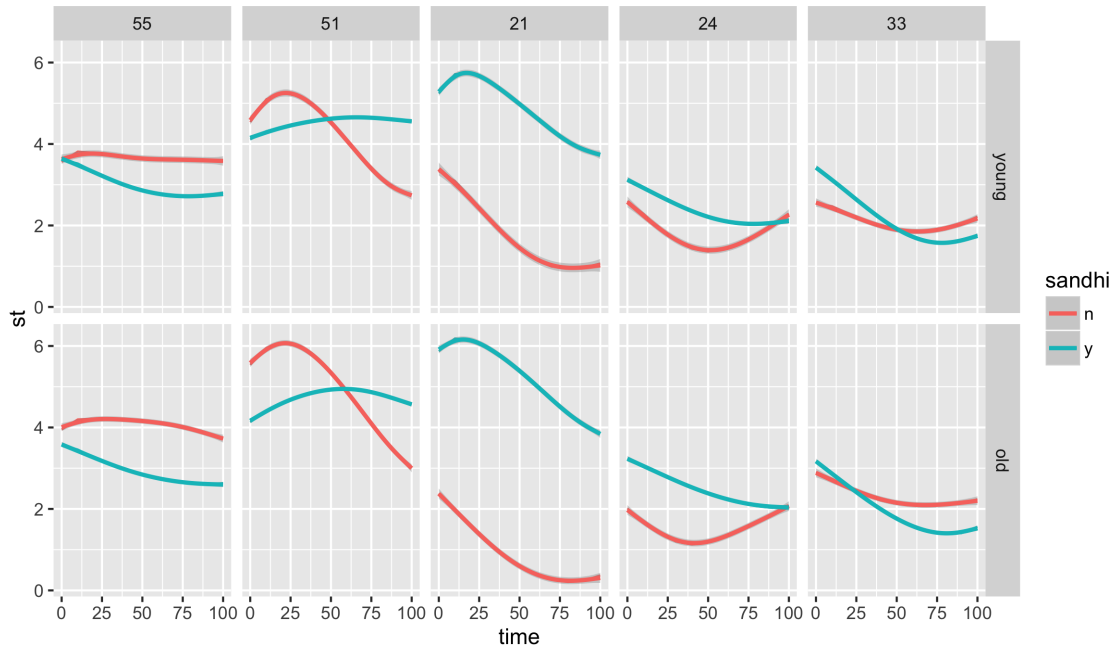


Figure 5.1: Mean f_0 curves of citation and sandhi variants of each unchecked tone for the two age groups. The upper panel shows the result for the young speakers and the lower panel for the old speakers. The red lines are the citation variants while the green lines are the sandhi variants. The shaded grey areas around each mean f_0 curve represents the 95% confidence interval.

f_0 contour for the old speakers whereas the dashed lines are for the young speakers, and the shaded areas are the corresponding 95% Bayesian confidence intervals. For instance, the tone category in Figure 5.3(a) has high-level 55 as its citation variant and mid-level 33 as its sandhi form; thus for both young and old speakers, the red line representing the citation variant always lies above the green line denoting the sandhi variant. In general, the f_0 curves between the two age groups are parallel to each other, similar to what is shown in Figure 5.2. Apart from the sandhi 33 in Figure 5.3(a) and the second half of the sandhi 33 in Figure 5.3(d), there is not much overlap between the f_0 contours produced by different age groups, suggesting

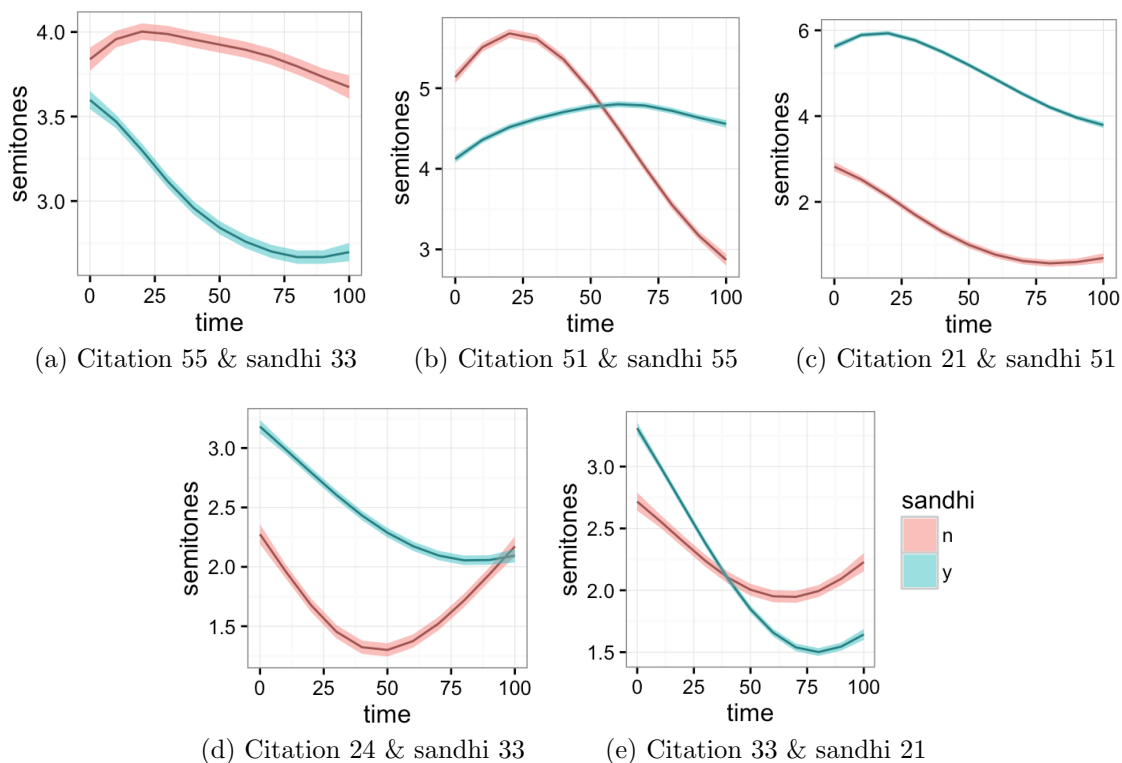


Figure 5.2: F0 curves of citation and sandhi variants of each tone category modeled by SS-ANOVA. Red lines indicate the citation variants and green lines represent the sandhi variants. The solid lines are the resulting smoothing splines indicating the mean f0 contour in each case, and the shaded areas are the corresponding 95% Bayesian confidence intervals.

that there is a significant difference in the f0 height in each of these cases. When comparing the corresponding f0 curves of the two age groups in each case, one can see that the f0 difference between the pair of citation and sandhi variants is larger for the old speakers than that for the young speakers. Or more specifically, the solid lines are above the dashed lines for high tones while the reverse is observed for low tones, as shown in Figures 5.3(a), (c), (d), and (e). This seems to suggest that the old subjects speak in a more extreme way than the young subjects in terms of the pitch range covered.

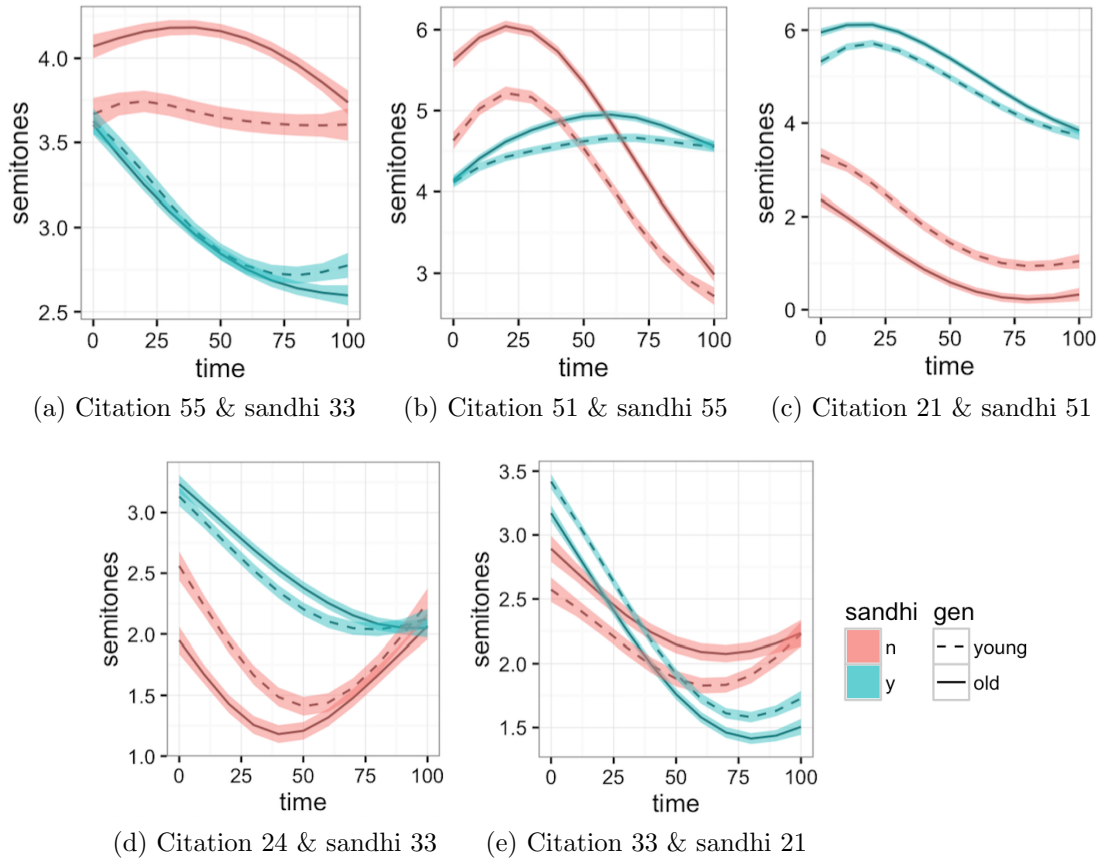


Figure 5.3: F0 curves of citation and sandhi variants of each tone produced by both age groups are modeled separately using SS-ANOVA. Red lines indicate citation variants and green lines represent sandhi variants. In each case, the solid lines are the resulting smoothing splines indicating the mean f0 contour for old speakers whereas the dashed lines are for young speakers, and the shaded areas are the corresponding 95% Bayesian confidence intervals.

Next, the realization of Tone Circle in TSM is examined by comparing pitch contours between different variants that are said to be realized the same on the surface, namely citation X and sandhi X where $X \in \{55, 51, 21, 33\}$. For instance, the citation variant 55 and the sandhi variant 55 derived from citation 51 are supposed to correspond to each other. Figure 5.4 shows the pitch contour comparison between variants that are said to have the same surface tonal contour. The solid lines are

data from the old speakers while the dashed lines are data from the young speakers. Different citation tone categories are represented with different colors; among them the red lines signify where the citation variant is realized as opposed to the sandhi variant. For example, in Figure 5.4(a) the sandhi variant of citation 51 is high-level 55 (the green lines), which is said to be realized the same as a citation 55 (the red lines) on surface.

In Figure 5.4(a), the pitch contour of sandhi 55 derived from citation 51 (the green line) is above that of citation 55 (the red line) for both old and young speakers. In Figure 5.4(b), where the surface contour is 51, the sandhi variant derived from citation 21 (the green line) is on top of the citation variant 51 (the red line) throughout the entire contour for the young speakers (the dashed lines). But for the old speakers, the two solid curves overlap in the second quarter and diverge afterwards where the citation 51 has much steeper slope. In the context of tone 21 on surface, similar pattern is observed. The pitch contour of sandhi 21 derived from citation 33 (the green lines) lies on top of that of citation 21 (the red lines) without overlap for both old and young speakers, as shown in Figure 5.4(c). The fact that the citation variant being lower or having a steeper slope than the sandhi variant is not surprising, because the two variants occur at different positions within a syntactically-determined domain: while citation variants can only appear at domain final position, sandhi variants can occur at any non-final position. As such a phrasal position effect is always expected, one cannot simply conclude that Tone Circle in TSM does not exist based on the fact, as shown here, that the f_0 contours of the corresponding citation and sandhi

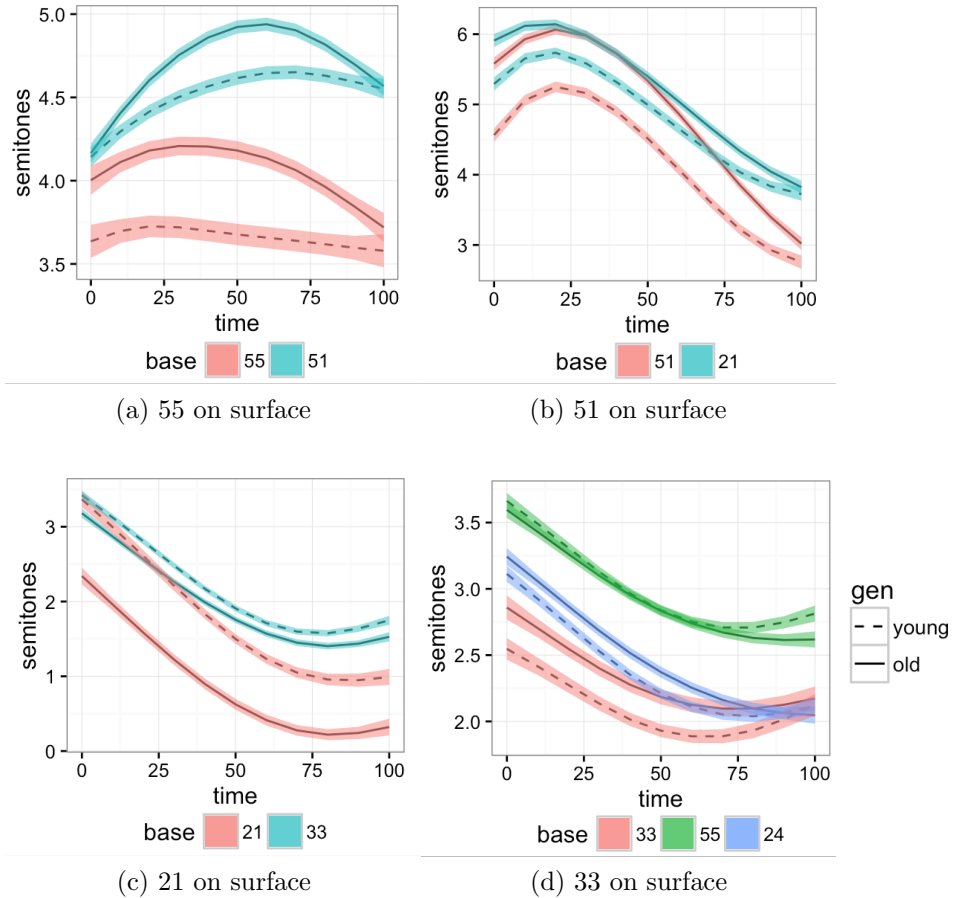


Figure 5.4: F0 curves of citation and sandhi variants that are supposed to correspond to each other produced by both age groups are modeled separately by SS-ANOVA. Red lines indicate citation variants and green (and blue) lines represent sandhi variants. In each case, the solid lines are the resulting smoothing splines indicating the mean f0 contour for old speakers whereas the dashed lines are for young speakers, and the shaded areas are the corresponding 95% Bayesian confidence intervals.

pair differ significantly from each other. In particular, it is unclear whether the f0 difference between the two variants (about 1 semitone) can be fully accounted for by this junctural effect. A potential evidence comes from the tonal neutralization case where both citation 24 and citation 55 are said to be realized the same as 33 on surface, as shown in Figure 5.4(d) along with the realization of citation 33. The

sandhi variants either derived from citation 55 (the green line) or from citation 24 (the blue line) almost always lie above the citation variant (in blue) for both age groups, which could possibly be due to the phrasal position effect. Assuming that Tone Circle really exists, a great deal of overlap in the pitch contours of sandhi variants derived from citation 55 and citation 24 should be expected. However, the two contours do not overlap and therefore are significantly different from each other, implying that for both old and young speakers the two different sandhi tones are not neutralized, or at least they are not completely neutralized. Meanwhile, the sandhi contour derived from citation 55 lies above the one derived from citation 24 for both generation groups, suggesting that the nature of f₀ height in their citation forms are preserved to some extent in the surface forms.

Since the conditions for tone sandhi application are morphosyntactic rather than phrasal, there is a certain dissociation between phrasing and sandhi environments, especially in spontaneous speech. As there are cases where citation variants are not followed by pauses, and there are other cases where sandhi variants occur before pauses for some reason or other, such as hesitation, the phrasal position effect on f₀ difference could possibly be tested by looking at prepausal sandhi environments and non-prepausal citation environments. I first compared non-prepausal citation and sandhi variants that are supposed to come out the same way using SS-ANOVA, and the result is shown in Figure 5.5. For both generation groups, there is no overlap between the confidence intervals of the citation and sandhi variants in all of the cases except where the two curves intersect, meaning that for non-prepausal syllables which

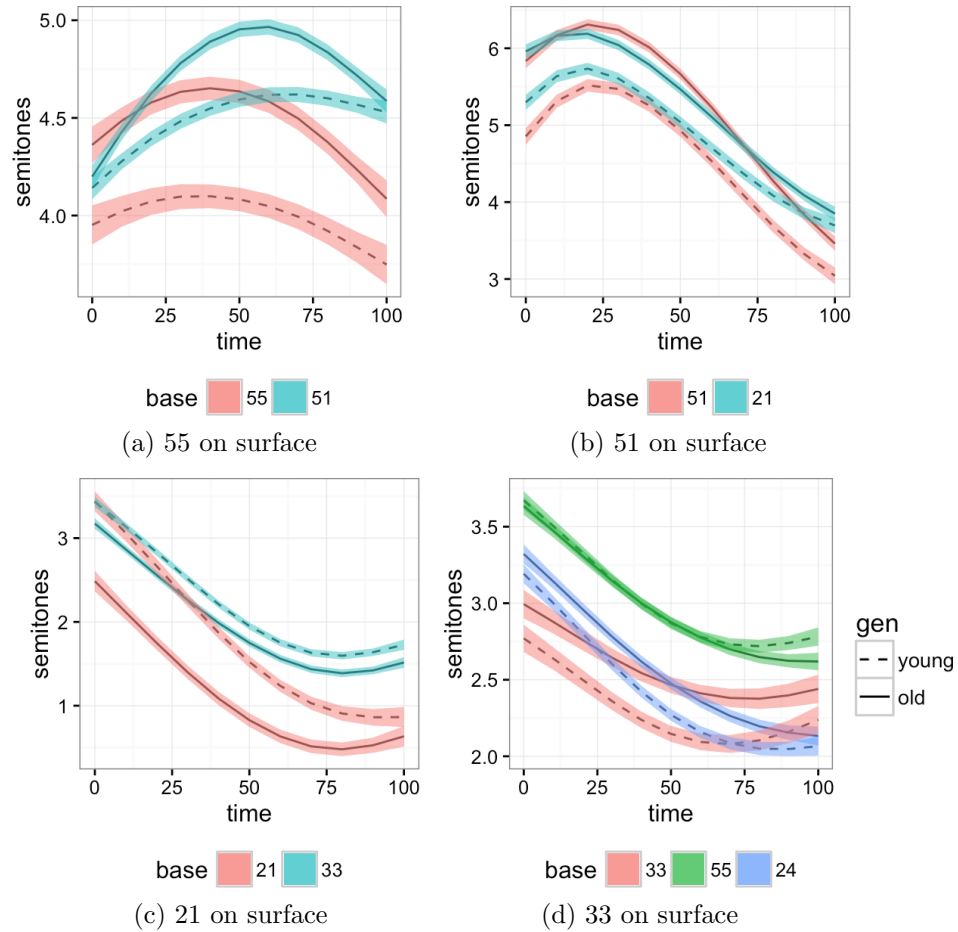


Figure 5.5: Non-prepausal f_0 curves of citation and sandhi variants that are supposed to correspond to each other produced by the young and old speakers are modeled separately by SS-ANOVA. Red lines indicate the citation variants and green (and blue) lines represent the sandhi variants. In each case, the solid lines are the resulting smoothing splines indicating the mean f_0 contour for the old speakers whereas the dashed lines are for the young speakers, and the shaded areas are the corresponding 95% Bayesian confidence intervals.

are subject to limited junctural effects, there is significant difference between the corresponding citation and sandhi variants that are supposed to surface the same. As shown in the graphs, the constantly changing difference lies in the f_0 height, in the slope of the f_0 curve, or most of the time in both. For instance, in Figure 5.5(c) there

is approximately one semitone difference between the citation and sandhi variants produced by the old speakers, whereas for the young speakers, the citation variant has much steeper slope than the sandhi variant derived from citation 33. Notwithstanding these dissimilarities, Figure 5.5(b) presents a case where the statistically significant difference between the f0 curves of the citation and sandhi variants may not be perceivable. On the other hand, the non-prepausal data provides another evidence against the tonal neutralization between the two sandhi 33 variants derived respectively from citation 55 (the green line) and citation 24 (the blue line), as shown in Figure 5.5(d), where there is about 0.5 semitone difference throughout the entire contour and the sandhi derived from citation 55 always lies above the one derived from citation 24, reflecting some preservation of their citation tone feature.

Figure 5.6 illustrates the comparison of f0 curves between prepausal citation and sandhi variants that are said to correspond to each other using SS-ANOVA. Apart from the variants with 21 tone on surface produced by the young speakers (the dashed lines) as depicted in Figure 5.6(c), no overlap throughout the entire contour is found between the confidence intervals of the corresponding variants, parallel to the general pattern of the non-prepausal data. Comparing Figures 5.6 and 5.5, one can observe that for the old speakers the difference in f0 height between citation and sandhi variants is larger for the prepausal data than for the non-prepausal data. And once again, as seen in Figure 5.6(d), for both generation groups, the contour for the sandhi variant derived from citation 55 has the highest f0, followed by that for the sandhi variant from citation 24, and lastly the curve for citation 33, suggesting that the

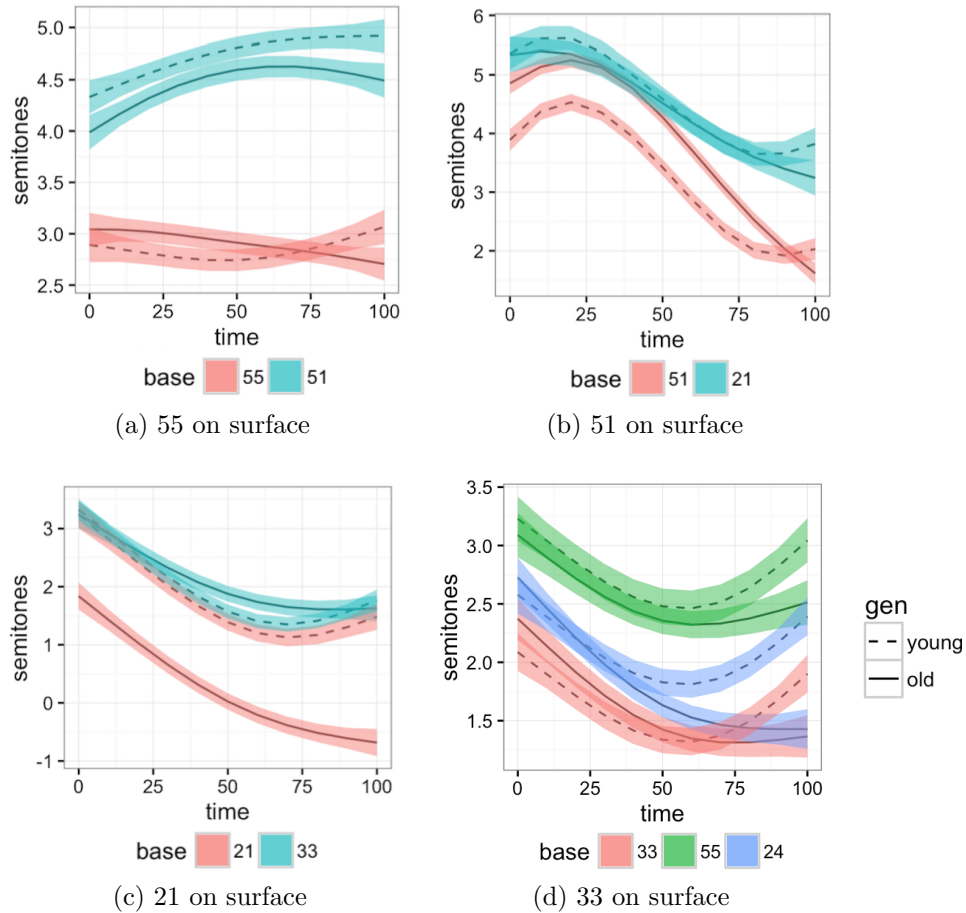


Figure 5.6: Prepausal f_0 curves of citation and sandhi variants that are supposed to correspond to each other produced by the young and old speakers are modeled separately by SS-ANOVA. Red lines indicate the citation variants and green (and blue) lines represent the sandhi variants. In each case, the solid lines are the resulting smoothing splines indicating the mean f_0 contour for the old speakers whereas the dashed lines are for the young speakers, and the shaded areas are the corresponding 95% Bayesian confidence intervals.

characteristics of the corresponding citation tones are retained to some extent. Note that the three f_0 contours are very similar in shape for young speakers, but for old speakers the sandhi contour derived from citation 55 differs from the other two to a certain degree in terms of slope and f_0 height.

5.2.2 Fitting orthogonal polynomial

In addition to SS-ANOVA, an orthogonal polynomial regression model is fitted for each pitch contour to reduce the dimensions, such that each contour is represented with three parameters. For statistical analyses, kernel density based global two-sample comparison test (*kde.test* function from *ks* package in R) is adopted to compare the result of orthogonal polynomial fitting by testing whether two samples of data are generated from the same distribution. This test can be used to compare two samples of one- to six-dimensional data, and the null hypothesis is that the two sets of data have the same density functions, meaning that they are generated from the same distribution. The higher the p-value, the better the fit between these two sets of data. Table 5.4 presents the result of comparing the dimension-reduced f0 curves of citation and sandhi variants of the same tone for both old and young speakers. And not surprisingly, for both generation groups the citation and sandhi variants of the same tone category differ significantly from each other in all but one pair, i.e., citation 33 and sandhi 21 from the old speakers being marginally significantly different from each other, which could possibly be due to some final lowering effect on citation 33 variants.

With the *rgl* package in R, one can visualize the three-dimensional orthogonal polynomial data in an interactive way by rotating the axes, or zooming in or out on the scene. Figure 5.7 shows a screenshot of such visualization where the contrast can be clearly observed. The data includes the citation and sandhi variants of T51,

Table 5.4: kde.test for orthogonal polynomial fitting f0 curves of different variants of the same tone.

Tone category	T55	T51	T21	T24	T33
Old	1.9009e-10 ***	4.8233e-36 ***	1.3459e-70 ***	1.7656e-24 ***	0.0962 .
Young	0.0093 **	4.9018e-51 ***	3.9816e-28 ***	3.5248e-12 ***	0.0078 **

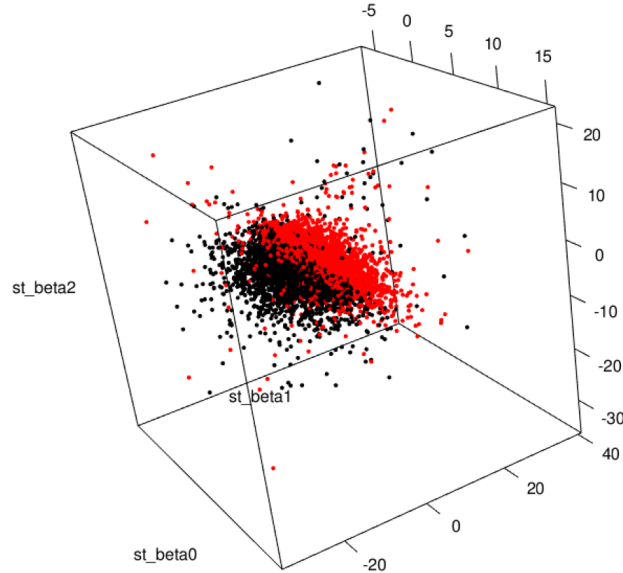


Figure 5.7: Visualization of the 3-D orthogonal polynomial data of citation and sandhi variants of T51 produced by the old speakers. Citation 51 variants are marked with black dots and sandhi 55 variants are represented with red dots. The three axes correspond to the three orthogonal polynomial parameters.

i.e., citation 51 (represented with black dots) and sandhi 55 (red dots), produced by the old speakers. The three axes correspond to the three orthogonal polynomial parameters. As shown here, the two variants seem to form their own cluster that can be distinguished by beta1 parameter of orthogonal polynomial, which is also verified by the kernel density based global two-sample comparison test with a p-value being 1.9009e-10.

Table 5.5: `kde.test` for orthogonal polynomial fitting f_0 curves of the same variant produced by different age groups.

Tone category	T55	T51	T21	T24	T33
Citation	0.0038 **	0.0266 *	0.1799	0.1164	0.1200
Sandhi	0.2711	0.0601 .	0.0830 .	0.0243 *	0.2326

Table 5.5 exhibits the result of comparing f_0 curves of the same variant produced by different age groups using `kde.test`. Contrary to the result of SS-ANOVA in Figure 5.3, where limited overlap (denoting statistical insignificance) occurs at the very beginning and/or the very end of the f_0 contours and only one statistically insignificant difference is found between the sandhi variants of T55 (viz. sandhi 33) produced by the old and young speakers, the output of `kde.test` on orthogonal polynomial fitting reports solely three out of ten comparisons with statistical significance. This is primarily due to some loss of information when the degrees of freedom are reduce; meanwhile, how well the orthogonal polynomial model fits a pitch contour is not taken into consideration.

Table 5.6: `kde.test` for orthogonal polynomial fitting f_0 curves between variants that are supposed to have the same surface contour.

Surface form	55 (b55/s51)	51 (b51/s21)	21 (b21/s33)	33 (b33/s55)	33 (b33/s24)	33 (s55/s24)
Old	0.0028 **	0.0001 ***	0.0231 *	0.0303 *	0.0193 *	0.1192
Young	0.0470 *	0.0002 ***	0.0067 **	0.0151 *	0.0215 *	0.0049 **

The output of the `kde.test` for orthogonal polynomial fitting f_0 curves between variants that are said to have the same surface contour can be found in Table 5.6, and is generally in line with the result of SS-ANOVA in Figure 5.4 with the exception of the only insignificant difference reported here which refers to the comparison between sandhi 33 variants derived from citation 55 and citation 24 produced by the old speakers.

5.2.3 Dealing with confounding

While spontaneous speech is more natural and dynamic than read speech and shows considerable variation in many respects such as prosodic and lexical patterns, it is this variability and the lack of systematic control that make it rather challenging to tease out any factors at play. Thus, a large scale database is often required to ensure maximum balance among all kinds of variabilities. As a final step, I will discuss two conceivable factors, i.e. the durational difference and the effect of morpheme frequency, and try to demonstrate that both of them have little effect on our end result.

First consider the possible effect of durational difference on the tonal variation. To assess whether this could be a potential confounder, I first check whether there is a difference in durational distribution between citation and sandhi forms. Here rhyme duration instead of syllable duration is measured, since the rhyme of the syllable carries the relevant f_0 contour, and outliers are removed with the R command

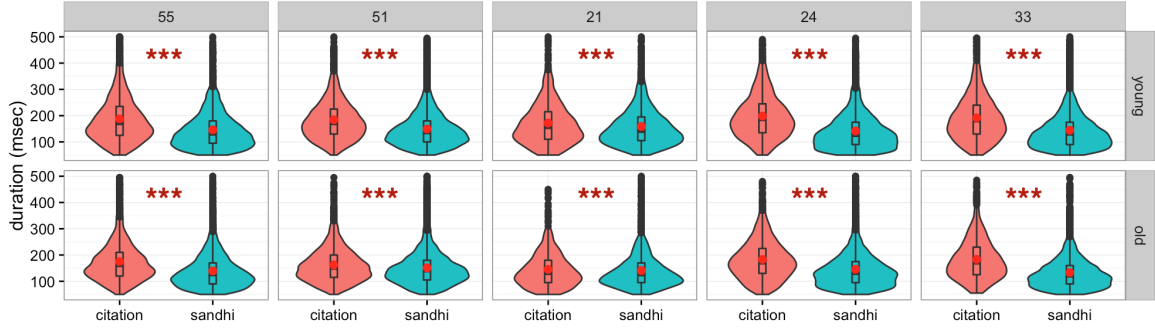
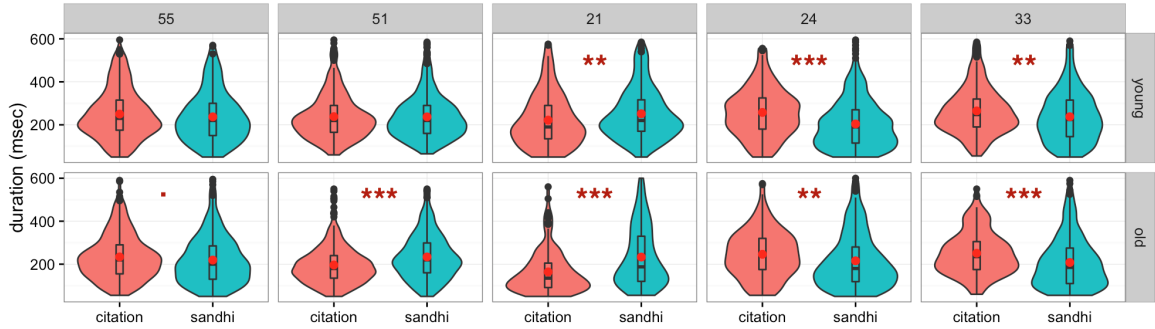


Figure 5.8: Distributions of rhyme duration of citation and sandhi variants grouped by tone category and generation groups. Red dots represent mean durations, and asterisks mark the statistical significance.

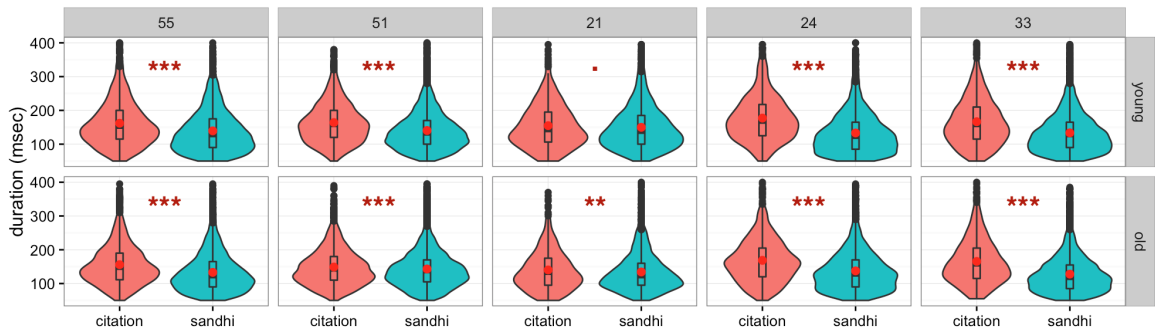
`boxplot.stats()$out`². Figure 5.8 shows the durational distributions of corresponding citation and sandhi variants in each unchecked tone category. The data of young subjects is in the upper panel, and that of old subjects in the lower panel. For both young and old subjects, the rhyme duration of citation variants is significantly longer than that of sandhi variants for each tone category, in line with our expectation that citation variants tend to have longer duration than sandhi variants in view of the phrasal position effect. This also implies that tone sandhi in TSM is not truly structure-preserving in light of this durational difference.

Once we are certain that the phrasal position effect has an influence on the durational distribution of citation and sandhi forms, the next step is to examine whether this difference in length still exists when we separate the data into prepausal and non-prepausal environments. Figure 5.9 shows the durational distributions of corresponding citation and sandhi variants in each unchecked tone category in prepausal

²The R command `boxplot.stats()$out` uses the Tukeys method to identify the outliers ranged above and below the $1.5 \cdot \text{IQR}$, where IQR refers to the inter-quartile range defined as the interval between Q1 and Q3.



(a) Prepausal environment



(b) Non-prepausal environment

Figure 5.9: Distributions of rhyme duration of citation and sandhi variants grouped by tone category and generation groups in prepausal and non-prepausal environments. Red dots represent mean durations, and asterisks mark the statistical significance.

and non-prepausal environments. In each graph, the data of young subjects is in the upper panel, and that of old subjects in the lower panel. Recall that the aim to divide the data into prepausal and non-prepausal environments is to exclude the phrasal position effect on f_0 difference; however, in each environment many of the corresponding pairs of citation and sandhi variants still differ significantly in the rhyme duration. This result suggests that the phrasal position effect cannot be simply ruled out with this crude approximation, and meanwhile confirms the difficulty in recognizing and controlling the contributing factors in spontaneous speech. Nonetheless, this further comparison of durational difference is crucial for cases where there is an overlap be-

tween the f0 contours of the corresponding citation X and sandhi X such as the one shown in Figure 5.6(c), since the two variants may still be significantly different from each other if there is significant durational difference between them, as corroborated with the T21 case in Figure 5.9(a).

Lastly, I try to examine whether morpheme frequency plays a role in determining the observed f0 contour pattern. First of all, to simplify matters I regard each monosyllabic word in TSM as a morpheme. Since there is no lexical/morpheme frequency information available for this corpus data, morpheme frequencies are simply calculated as the number of occurrences for each monosyllabic word divided by the total number of monosyllabic words in the corpus. Owing to the numerous cases of homophones (different words but have the same pronunciation, e.g., *ka55* can mean ‘home’ and ‘plus’ in TSM) and homographs (same word with different pronunciations, e.g., the same Chinese character meaning ‘home’ can be pronounced either as *ka55* as in *hap-ka* ‘the whole family’, or as *ke55* as in *jin-ke* ‘house, residence’) in Chinese languages, romanization transcription and tone category information together cannot fully represent every morpheme in TSM, I resort to the selection of Chinese characters put forth by the Taiwanese Ministry of Education in an effort to standardize the TSM content used in educational materials. In this respect, every morpheme in this corpus is characterized by a unique set of {Chinese character, romanization, tone category}. To achieve this in an efficient way, I randomly sample half of the interview recordings³ and transcribe them again with the corresponding Chinese characters. In

³This sample includes eight recordings from subjects CQH(m), CXH(m), CXH2(f), CZY(m),

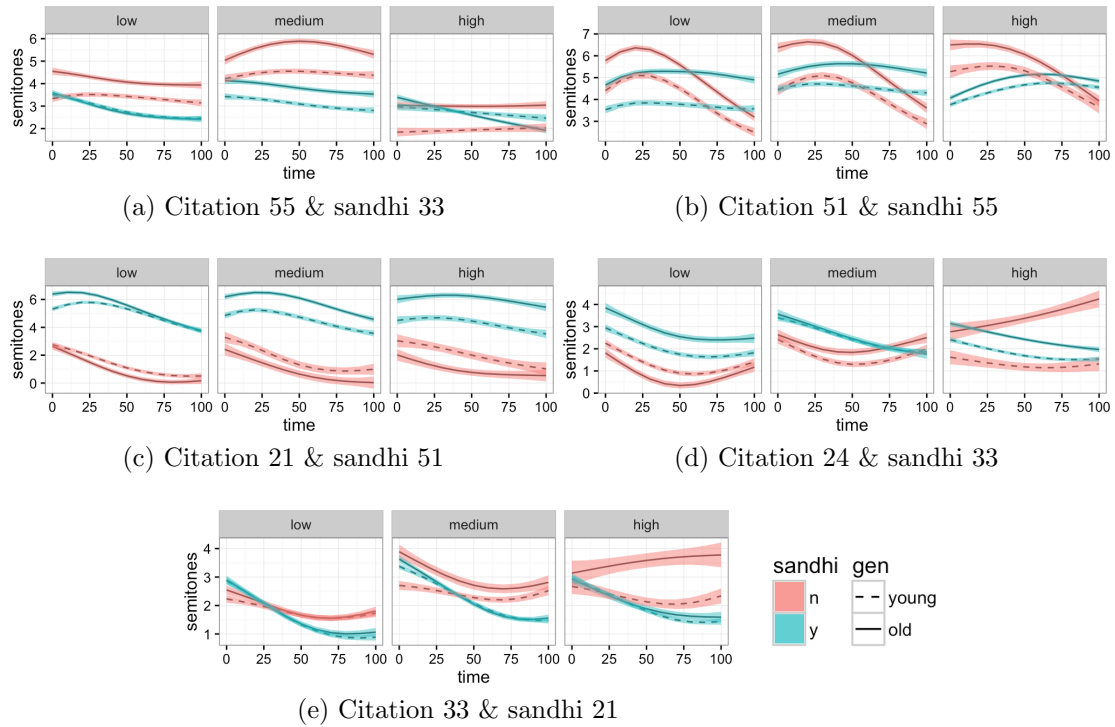


Figure 5.10: F0 curves of citation and sandhi variants of each tone grouped by morpheme frequencies are modeled separately using SS-ANOVA for both age groups. Red lines indicate citation variants and green lines represent sandhi variants. In each case, the solid lines are the resulting smoothing splines indicating the mean f0 contour for old speakers whereas the dashed lines are for young speakers, and the shaded areas are the corresponding 95% Bayesian confidence intervals.

total there are 2,534 unique morphemes in this corpus sample with the total word count equal to 396,770, and morpheme frequencies range from 1.63×10^{-5} to 0.0269 with an average of 3.946×10^{-4} .

To investigate the effect of morpheme frequency on the f0 contour pattern, I divide all the morphemes into three groups of similar size according to their absolute frequencies: (1) *low frequency*: including 115,863 words with a frequency value smaller than 10^{-4} , (2) *medium frequency*: including 115,863 words with a frequency value between 10^{-4} and 10^{-3} , and (3) *high frequency*: including 115,863 words with a frequency value greater than 10^{-3} . The first four speakers are in the young age group and the latter four in the old generation group.

than 0.001, (2) *medium frequency*: a total of 113,663 words with a frequency value in the range [0.001, 0.006), and (3) *high frequency*: containing 103,620 words with a frequency larger than or equal to 0.006. Given the effect of morpheme frequency, frequent words are generally pronounced in a more casual and relaxed manner such that a reduction in phonetic difference (e.g., the f0 difference between citation X and sandhi X that are supposed to correspond if the Tone Circle really exists) is expected. Figure 5.10 shows the SS-ANOVA modeling of the f0 curves of citation and sandhi variants of each tone grouped by morpheme frequencies. The solid lines refer to the result of old speakers and the dash lines are for young speakers. For each tone category, red lines represent citation variants whereas green lines stand for sandhi variants. Regarding tone category 55 in Figure 5.10(a), while citation 55 is expected to lie above sandhi 33, the opposite is observed in the high frequency condition for young speakers. In the rest of the four Figures, corresponding citation and sandhi variants unsurprisingly differ significantly from each other and show parallel pattern among three frequency groups for both generation groups. Though not obvious, the gradual reduction in f0 difference between citation 21 and sandhi 51 produced by young speakers may be a slight evidence for the morpheme frequency effect, as shown in 5.10(c).

Figure 5.11, on the other hand, illustrates SS-ANOVA modeling f0 curves of citation and sandhi forms that are supposed to correspond to each other grouped by morpheme frequencies for both age groups. For variants realized as 55 on surface shown in Figure 5.11(a), while citation and sandhi variants with high frequencies

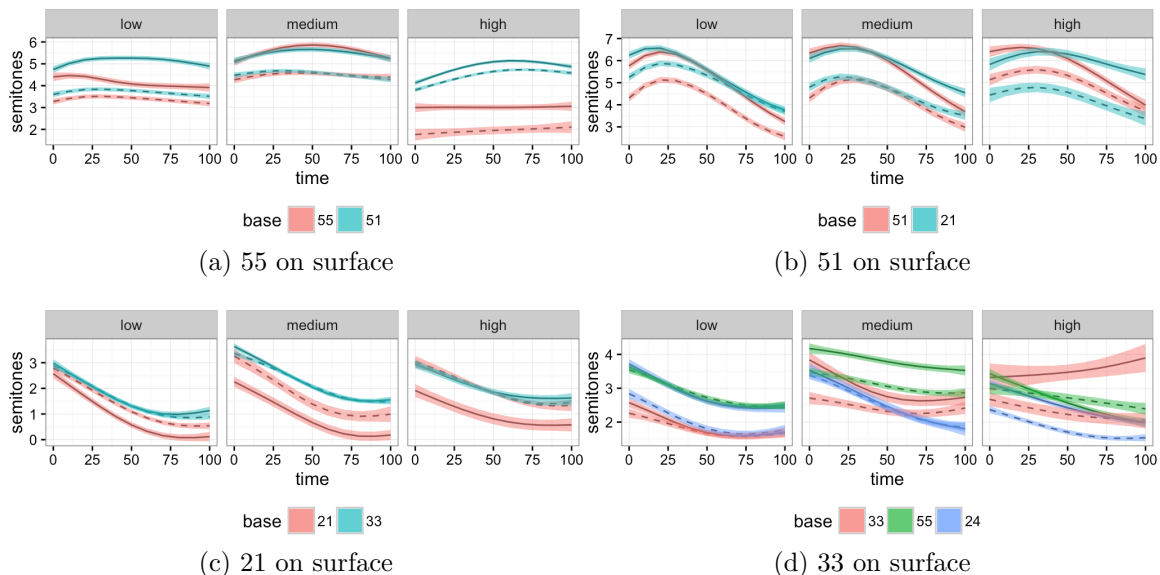


Figure 5.11: F0 curves of citation and sandhi variants that are supposed to correspond to each other grouped by morpheme frequencies are modeled separately by SS-ANOVA for both age groups. In each case, the solid lines are the resulting smoothing splines indicating the mean f0 contour for old speakers whereas the dashed lines are for young speakers, and the shaded areas are the corresponding 95% Bayesian confidence intervals.

differ significantly in f0 height for both age groups, this difference is completely lost for both generations in the medium frequency group and is greatly reduced by young speakers in the low frequency group while it is maintained by old speakers. For the surface falling tone 51 in Figure 5.11(b), corresponding citation and sandhi variants differ significantly either in f0 height or in contour slope or in both for most of the time except for where the two contours intersect. In Figure 5.11(c), while f0 contours of citation and sandhi variants differ significantly in all cases for old subjects, the two curves show very different patterns in the three frequency conditions: there are relatively small difference in slope in the low frequency group, significant difference

in f0 height for words with medium frequencies, and a complete overlap in the high frequency condition. Figure 5.11(d) compares two sandhi variants (i.e. the tonal neutralization case) and a citation variant, all of which are said to realized the same as 33 on surface. For old subjects, the two sandhi variants completely overlap with each other but differ significantly from the citation 33 in the low frequency condition, whereas significant difference among the three f0 contours is found in the other two frequency groups. Similar result is observed in the data of young subjects, only that it is the citation 33 and the sandhi 33 derived from citation 24 that are relatively close to each other, especially the latter half of the contours. This again confirms the result in the production experiment of tonal neutralization that the two sandhi 33 tones derived respectively from citation 55 and 24 are not neutralized, or at least are not in complete neutralization. In brief, Figures 5.10 and 5.11 together suggest that morpheme frequency plays a rather marginal role in determining the f0 contour pattern between citation and sandhi variants under such rudimentary systematic control.

5.3 Discussion

In this corpus study, I investigated the realization of the Tone Circle in a spontaneous TSM speech corpus. As expected, the realization of corresponding citation and sandhi variants of the same tone shows significant difference in f0 curves. In general, pitch contours produced by different generation groups are parallel to each other along with significant difference in f0 height, suggesting that there is a potential age-based

difference, in which old speakers tend to speak in a more extreme way (i.e. with larger pitch range) as opposed to young speakers. In regard to the realization of citation and sandhi variants that are supposed to correspond in their surface forms, significant difference among their pitch contours is found, and the result of the tonal neutralization case is in line with that of the production experiment described in Chapter 3, where the sandhi derived from citation 55 lies significantly above the one derived from citation 24, reflecting some preservation of their citation tone feature, only that incomplete tonal neutralization is observed in the corpus data for both young and old speakers (i.e. no signs for a generational difference) with an f_0 difference of at least 0.5 semitones, whereas in the previous tonal neutralization production experiment, incomplete neutralization is present only in the data of old speakers and complete neutralization is reported in the young subjects' data. A possible explanation for this is related to the time the two experiments took place. The young subjects in this corpus were between the ages of 20 to 35 when recorded, and thus they would be around 32 to 47 years old at the present time, which is between the age ranges of the old (52-66 years old) and young (25-35 years old) age groups in the production experiment. While this might serve as an evidence for a linguistic change in progress towards a complete merger, a carefully controlled apparent-time study is needed before such a conclusion can be reached.

With respect to the significant difference between citation X and sandhi X that are said to realized the same on surface, a potential contributing factor, i.e. the phrasal position effect, was examined by separating the data into prepausal and non-prepausal

environments. However, similar patterns are shown in both environments and significant difference between the corresponding citation X and sandhi X variants is found even for non-prepausal syllables, which presumably are subject to limited junctural effects as contrasted with prepausal syllables. Whether this could be regarded as an evidence against the existence of the Tone Circle requires further systematic controls on the corpus data. As a preliminary attempt, two conceivable confoundings, the durational difference and the effect of morpheme frequency, were taken into consideration, and the result shows that significant durational difference still exists in both prepausal and non-prepausal environments and that there is marginal effect of morpheme frequency on the tonal contours of citation and sandhi variants in the sense that the general pattern is consistent with the abovementioned result. This outcome suggests that a better and more careful control (e.g. separating content words from function words, and using slope estimation to factor out tonal coarticulation) is required to rule out the phrasal position effect, and meanwhile confirms the difficulty in recognizing and controlling the contributing factors in spontaneous speech.

Chapter 6

Tone Sandhi as Phrasal Phonological Rules

In previous chapters, I examined the phonological and phonetic nature of the tone sandhi system in TSM in terms of tonal neutralization, realizations of the citation and sandi forms in spontaneous speech, and the representations of citation and sandi tones in spoken word recognition. In this chapter, I focus on the circumscription of sandhi domains in this language. I will first provide a brief overview of Distributed Morphology in Section 6.1 and introduce some terminologies in order to outline the theoretical framework my proposal assumes. Then the definition of allomorphy in this framework is given in Section 6.2 to demonstrate that tone sandhi in TSM should not be analyzed as a case of allomorphy. In Section 6.3, I review a previous attempt adopting the same framework as my approach in Pak (2008), and thereafter my analysis of tone sandhi in TSM as phrasal phonological rules is presented in Section 6.4. This chapter concludes with a discussion offered in Section 6.5.

6.1 Overview of Distributed Morphology

The model of grammar underlying the theory of Distributed Morphology (DM) (Halle and Marantz, 1993, 1994, et seq.) is shown in Figure 6.1. Syntactic structures are first generated by a set of rules such as head movement in the syntax, and then are sent to be realized phonologically at PF (Phonological Form) and to be interpreted semantically at LF (Logical Form) by the spell-out operation.

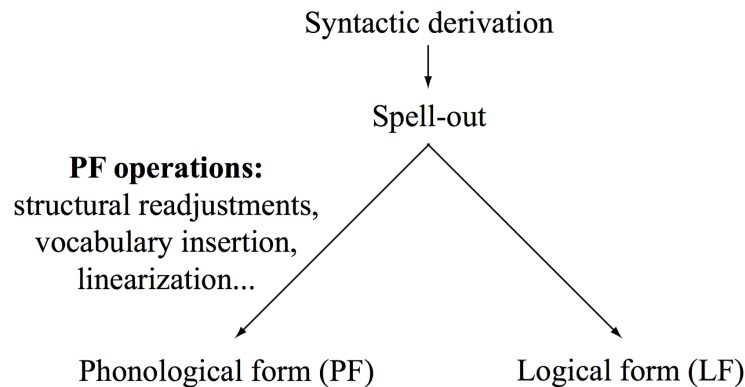


Figure 6.1: Architecture of Distributed Morphology.

In the framework of DM, there is no lexicon, and consequently terms like ‘lexical’, ‘lexicalized’, and ‘lexical item’ have no significance. The internal structure of words is produced first by a series of syntactic operations such as Merge and Move in the syntax proper, followed by a set of post-syntactic processes that take place along the PF branch. In this regard, morphology (i.e., the set of processes relevant for word formation, which in other theories are primarily ascribed to the lexicon) is now

distributed over other components of the grammar — hence the name “Distributed Morphology.” In other words, there is one single system responsible for generating word and phrase structures, namely, syntax (*Syntactic Hierarchical Structure All the Way Down*, as Halle and Marantz (1994) put it). By the same token, as there is no lexicon at all, there is no distinction between lexical and postlexical phonologies; rather, all phonology is placed in a single post-syntactic module.

Morphemes as basic objects of syntactic derivations are of two types: *Roots* and *functional morphemes*. Roots, e.g. $\sqrt{\text{CAT}}$ and $\sqrt{\text{SLEEP}}$, being category-neutral are categorized by category-defining functional heads such as *n*, *v*, *a*, etc. in syntactic structures. Functional morphemes, on the other hand, are bundles of abstract syntactico-semantic features such as [+pl] and [+past], and do not have phonology in the first place, but are realized phonologically via the operation called *Vocabulary Insertion* in the PF component of the grammar, corresponding to the so-called *Late Insertion* hypothesis, one of the core properties of the DM theory. In syntax, these two types of morphemes as the terminal nodes are built into complex objects that are hierarchically arranged but linearly unordered morphosyntactic structures, which are sent to the PF and LF branches at spell-out. In particular, following the syntactic phase theory of Chomsky (2000, 2001), it is assumed that syntactic structures are built up and processed one chunk at a time (i.e., in phases or cycles), and that what is spelled out at each phase is either a constituent or a predictable subpart of one (e.g. ‘left-over’ material from a previous phase). Under this direct spellout hypothesis, Pak (2008) proposes that phonological rules deal directly with the output of each phase

and apply at various points in the PF derivation such that they are interleaved with different kinds of linearization procedures.

On the PF side, operations including Structural readjustments, Vocabulary Insertion, and Linearization take place to convert abstract hierarchical structures into fully linearized phonologized strings (Embick and Noyer, 2007, among others), compatible with the idea that linear order is a requirement of the articulatory-perceptual interface rather than of the syntax proper. Specifically, objects that are referred to by PF operations are M-words (i.e. maximally complex heads) and Subwords (i.e. terminal nodes, either Roots or functional morphemes). Linearization operation is further decomposed into two steps: (1) within a given cycle, each branching node of a structure is visited and a left-adjacency statement between its two daughters is produced with respect to language-specific principles about e.g. headedness to select the right linear order. Notationally, the operator ‘*’ means ‘is left-adjacent to’ and can relate either M-words or phrases. (2) *Concatenation* as a second step of linearization indicates a binary relation of immediate precedence by looking inside each member of a *-statement and producing a corresponding statement of left-adjacency between the peripheral M-words on each side. The operator \frown encoding concatenation relates M-words only. The statement $X \frown Y$ denotes that X immediately precedes Y.

Pak (2008) further assumes that there are two separate steps involved in Concatenation: *Head-left* (or ‘early’) Concatenation and *Phrase-left* Concatenation, both of which are proposed for the tone sandhi rules in TSM, as described in Sections 6.3 and 6.4. The algorithms of both Concatenation rules are illustrated in (1). Head-left

Concatenation identifies pairs of M-words X, Y where (i) X is left-adjacent to Y, and (ii) X c-commands Y. The algorithm stops as soon as it encounters a *-statement X^*Y whose right-hand member Y is not internally complex. Phrase-left Concatenation, on the other hand, identifies pairs of M-words X, Y where (i) X is left-adjacent to Y, and (ii) X does not c-command Y.

- (1) a. *Head-left Concatenation* begins with a *-statement whose left-hand member is an overt M-word X (rather than a phrasal category or null head), and searches within its right-hand member until it identifies the overt M-word Y that appears immediately to the right of X, and creates a binary statement of left-adjacency between the two M-words ($X \frown Y$).
- b. *Phrase-left Concatenation* begins with a *-statement whose left-hand member is a phrasal category; it then searches within the left-hand member until it identifies the right-most M-word, and within the right-hand member until it identifies the left-most M-word.

6.2 Allomorphy in DM

In its most general sense, the term *allomorphy* covers any variations in shape that different morphemes undergo. It is often used ambiguously to describe two different types of alternations: one that more phonological, and one that is morphological. To be clear, it is only the second of these that is referred to as allomorphy.

For example, the plural form for regular nouns in English appears in three different surface forms: /z/ as in *dog-s*, /s/ in *cat-s*, and /əz/ in *fox-es*. The alternation between the three is phonological in that it is possible to account for the distribution of alternants by phonology given /z/ as the underlying form of the plural morpheme, and accordingly only one morpheme in memory is needed instead of three distinct ones. However, English has irregular plurals such as *ox-en* for *ox*, *deer-Ø* for *deer*, and *geese-Ø* for *goose* (along with stem change), where there is no plausible phonological relation among the alternants. In this case, we say that the plural morpheme [+pl] has three distinct underlying forms, or allomorphs, associated with it morphologically, namely /-z/, /-en/, and Ø, which means there are three different objects in memory, termed as *Vocabulary Items* (VIs) in DM, as listed in (2).

(2) Vocabulary Items for [+pl] in English:

- a. [+pl] ↔ -en/___ { $\sqrt{\text{OX}}$,...}
- b. [+pl] ↔ -Ø/___ { $\sqrt{\text{DEER}}$, $\sqrt{\text{SHEEP}}$, $\sqrt{\text{GOOSE}}$,...}
- c. [+pl] ↔ -z

In the example of the English plural, the syntax generates a structure that contains the plural node [+pl]. In the PF computation, the VIs in (2) compete for insertion into this node. When Roots like $\sqrt{\text{OX}}$ and $\sqrt{\text{DEER}}$ are present, the Vocabulary Insertion process inserts -en and -Ø into the [+pl] node respectively; in other cases, the default -z is inserted. Note that the three items have the same morphosyntactic

feature but differ in their stated insertion contexts and phonological features, and thus this kind of allomorphy is called *contextual allomorphy*. In the case of English plural, the allomorphy is *grammatically conditioned* as the determining factor is a specific morpheme/Root in the context of the node showing allomorphy, viz., the identity of the particular noun that the node [+pl] is attached to. Contextual allomorphy can also be *phonologically conditioned*, in which case the choice of a particular allomorph of some morpheme is determined by phonological factors.

- (3) Locality conditions for contextual allomorphy:
- a. Linear adjacency: Morphemes can interact for allomorphy only when they are immediately linearly adjacent, i.e. *concatenated*: $X \smallfrown Y$.
 - b. Phase: Morphemes can interact for allomorphy etc. only when they are in the same phase domain.

In DM, the engine that drives contextual allomorphy is Vocabulary Insertion, where individual VIs compete for insertion at a given node, and the most specific VI that can apply assigns the phonological content to that node. Possible patterns of contextual allomorphy are constrained by the interaction of two independent locality conditions (Embick, 2010), as listed in (3). The linear locality constraint states that contextual allomorphy, where a node X which VIs compete for can see another node Y that is related to the context of node X, is possible only when X and Y are concatenated. The phase-cyclic notion of locality follows directly from Chomsky's assumption that

syntactic structures are built up and processed in phases that are spelled out cyclically. The idea is that only the nodes that are in the same phase can interact for PF purposes such as allomorphy.

Clearly, the allomorphy in DM described above is different from the allomorphic treatments of tone sandhi in TSM discussed in Section 2.4, where multiple allomorphs with diacritics marking tone category stored in the lexicon, either generated by pre-compiled tone sandhi rules (Hayes, 1990) or essentially listed in the lexicon (Tsay and Myers, 1996), are inserted in the relevant syntactic contexts postlexically, e.g., insert corresponding citation allomorphs at the right edges of non-lexically-governed maximal projections and sandhi allomorphs elsewhere. For one thing, two distinct allomorphs need to be posited for all morphemes in the language, as each monosyllabic word in TSM is generally a morpheme and is subject to tone sandhi. Apart from the storage of all the allomorphs in memory, this approach seems redundant for the fact that as long as words belong to the same citation tone, they undergo the same sandhi phenomenon and that the two tonal alternants, citation and sandhi, occur in complementary distribution. For another, in order to determine which allomorph to insert, the postlexical insertion frame must have access to nonlocal information about the syntactic structure such as the head governing the node for allomorphy is lexical or functional, which demands significant relaxation of the locality conditions on allomorphy, and thus can rule in unattested cases of allomorphy conditioned by material that is not local, say more than one word away.

In order to account for tone sandhi in TSM, the phonology must be able to ‘see’

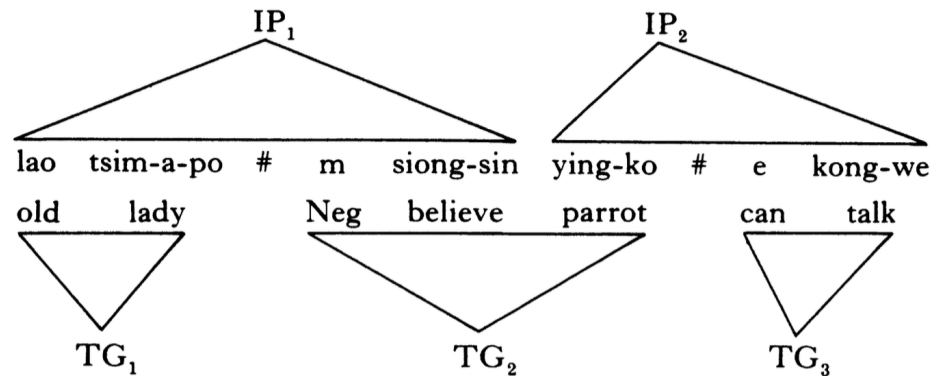
certain aspects of the syntactic constituent structure. In the DM-based articulated derivational model of the syntax-phonology interface developed in Pak (2008), phrasal phonological rules, i.e., phonological rules that apply across words but not across the board, are cross-linguistically *sensitive to the syntactic constituent structure* in some very general sense — words are not just randomly grouped together into phonological domains, but instead are grouped systematically in ways that preserve basic aspects of the underlying syntax. Specifically, *all phonological rules apply directly to the syntactic structure* as it happens to exist at the given point in PF, and domains for rule application thus ‘come for free’, which appears to be a good fit for the tone sandhi phenomena in TSM. In the rest of the chapter, I will first review one such alternative proposed in Pak (2008), and then explicate my proposal in detail, followed by the discussion.

6.3 Tone sandhi as Phrase-left Concatenation

In the case of tone sandhi in TSM, Pak (2008) assumes that the sandhi variant is the underlying form while the citation variant is the output of the rule, following Tsay and Myers (1996). She proposes that sandhi domains should be analyzed as a Phrase-left Concatenation rule — the later of the two steps involved in Concatenation. A Phrase-left Concatenation rule works by identifying a *-statement whose left-hand member is a phrasal category (i.e., internally complex) and searching within the left-hand member until the rightmost M-word is identified, and searching within the

right-hand member until the leftmost M-word is identified. To put it simply, phrase-left Concatenation rules apply internal to a spellout domain, namely CP, between two concatenated M-words X and Y, where (i) X precedes Y, and (ii) X does not c-command Y. One advantage of this approach lies in the fact that phonological rules applying after phrase-left Concatenation are put into effect at the junctures between XP constituents but not within XPs, which is congruous with the first observation made by Chen (1987). By way of illustration, Pak refers to the domain-mismatch case in (20), reproduced below as (4), and proposes the corresponding set of phrase-left Concatenation statements as in (5).

(4) ‘The old lady doesn’t believe that parrots can talk.’



- (5) a. $\text{tsim-a-po} \hat{\ } m$
 b. $\text{ying-ko} \hat{\ } e$

Pak claims that the tone sandhi domains are exactly those listed in (5) rather than the three TGs marked in the lower level of (4). That is to say, there are sandhi boundaries between *tsim-a-po* and *m* and between *ying-ko* and *e*. Under this analysis,

the pause between the two IPs does not break apart a sandhi domain, and accordingly, (4) shows no misaligned boundaries, which is compatible with her model's prediction that cases of domain-mismatch should be ruled out. As Pak focuses on justifying syntax-phonology mismatches instead of the tone sandhi behavior, only one TSM example is considered. It is unclear how her analysis can account for the fact that an utterance-final tone-bearing syllable, as well as a monosyllabic word in isolation, is always realized with its citation form. Likewise, a fundamental issue of this approach is related to the long-lasting debate about which form is underlying. The result of the priming experiment described in Section 4.4 showing that native speakers, especially the old speakers, have stronger association with the citation form than with the sandhi form seems to argue in favor of the opposite of her assumption. With this in mind, let's look at how this model applies to other TSM examples discussed in Section 1.3.

Start with the distinction between sentential adverbs (S-adverbs) and VP adverbs (VP-adverbs). While each of the S-adverbs in (6) and *ka-chai* in (8b) form separate sandhi domains, the VP-adverb *chiah* in (7) and *loan-chu* in (8a) phrase together with the following (modal) verb.

(6) ki-sit # au-pai # kho-leng # i e khi Bi-kok #
 actually next time maybe he will go USA
 'Actually, next time maybe he will go to USA.'

(7) keh-tng-kang # chiah e tng#-lai0
 the second day just will return
 '(She) will not return until the next day.'

- (8) a. i loan-chu = kong
 he mindlessly talk
 ‘He is talking mindlessly.’
- b. gua ka-chai # che chit pan ki
 I fortunately take this Cl flight
 ‘Fortunately, I am taking this flight.’

Provided that an adverb is the head of an AdvP, a S-adverb adjoining to IP/TP does not c-command its following M-word and thus forms its own sandhi domain, a correct result. Likewise, a VP-adverb adjoining to VP/V’ does not c-command the ensuing verb, and therefore tone sandhi is expected to be blocked between the VP-adverb and its following verb, and the model’s prediction is falsified by the attested facts. One can argue that VP-adverbs in TSM may be in a much closer relation with the verb, such as adverb incorporation into the lexical verb. Rivero (1992) discusses a relatively small set of Modern Greek sentences, in which a verb and its modifying adverb constitute a phrase in one instance but occur joined into a single word, a composite, in another, e.g. férome kaká vs. kakoférome ‘behave badly’. Rivero reports that this incorporation is applicable to manner adverbs but not to temporal adverbs, and then suggests it will only be possible with adverbs that are VP-internal and thus construable as arguments. Smirniotopoulos and Joseph (1997) further show that in terms of productivity and semantic compositionality, this process is not the result of a syntactic rule but that of a lexical rule. To regard a VP-adverb and its following verb in TSM as a case of adverb-incorporation, further investigation on TSM adverbs would be needed, such as whether this phenomena is productive for all

VP-adverbs at all times and in what ways these adverbs differ from the other adverbs such that incorporation into verbs is allowed or required.

Consider the case in which tone sandhi applies between a noun or an adjective and its preceding modifier adjunct, e.g. the adjective ‘lazy’ in (9a) and *moa-a* ‘sesame seeds’ as an adjunct of ‘big’ in (10b). By the same token, this is contrary to what is predicted by Pak’s model, where a sandhi boundary is expected here due to the lack of a c-commanding relation between the two adjacent M-words. In this case, one can argue for some compounding between the two being at play here. For the two nominal structures in Chinese, [M N] and [M *de*¹ N], where N is a nominal, M the modifier of N, and *de* a particle in Mandarin, as shown in a TSM counterpart in (9), Huang (1984) and Duanmu (1998), among others, assume that [M N] is always a compound at X⁰ while [M *de* N] is a XP phrase. Notwithstanding the controversial status of the marker *e*, a tone sandhi boundary always follows its preceding modifier, either a noun as in (9b), an adjective in (11), or a clause as in (10a).

- (9) a. pin-toaN = hak-seng
 lazy student
 ‘lazy student(s)’
- b. pin-toaN # e hak-seng ‘lazy student(s)’
 lazy E student
 ‘lazy student(s)’
- (10) a. [moa-a # toa]_{CP} # e sio-piaN
 sesame seed big E bun
 ‘buns with big sesame seeds’

¹The modificational marker *e* in Taiwan Southern Min is the equivalent of *De* in Mandarin.

- b. [moa-a = toa]_{AP} # e sio-piaN
 sesame seed big E bun
 ‘buns as big as sesame seeds’ (tiny buns)
- (11) toa-han # e cha-bO gin-a # kio-cho a-khim #
 elder E female child call as A-Khim
 ‘The elder daughter is called A-Khim.’
- (12) a. i ka tang-oh # kai-siao chit e lu-peng-iu
 he to schoolmate introduce one Cl girlfriend
 ‘He introduced a girlfriend to his schoolmate.’
- b. i kai-siao chit e lau peng-iu # hO in bO
 he introduce one Cl old friend to his wife
 ‘He introduced an old friend to his wife.’
- c. i chiong hit pun chheh # sang hou tang-oh
 he obj-marker that Cl book give to schoolmate
 ‘He gave that book to his schoolmate.’

In double-object constructions like (12), tone sandhi applies between the verb and the object following it, either D.O. or I.O, but is blocked between two objects. Objects as verb arguments are inside the VP and the verb alone is clearly not a phrasal category; consequently no Concatenation statement can be generated between the verb and its following object and no sandhi boundary is predicted to occur between them, a correct result. For the same reason, a negative particle as the head of NegP, e.g. *m* ‘not’ in (4), phrases together with the following verb, adjective, or adverb, which lies inside the NegP. The non-application of tone sandhi between the two objects can be ascribed to the fact that both objects are internally complex DPs, and that the former one does not c-command the latter. When one of the objects is fronted before the verb,

that object phrases separately from the following verb as in (12a), (12c) and (25a) due to the fact that the object adjoining to IP does not c-command its following verb.

Except under focus or in contrastive use², pronouns generally do not form their own sandhi domain but phrase together with the M-word to its right, e.g. *i* ‘he’ in (12), pronouns *i* ‘she’, *in* ‘them’, and *lang* ‘someone’ in (13). This is in conflict with Pak’s analysis, in which a sandhi boundary is expected after the pronoun since the pronoun DP does not c-command the succeeding M-word. One can argue that in terms of sandhi phrasing, pronouns, either as subjects or objects, seem to ‘cliticize’ to the subsequent word. A similar case is found in Igbo, where a type of pronominal elements being dependent, short, and weak appear as proclitics only at the subject position before verbal elements, as shown in (14). Non-pronominal subjects, on the other hand, always form a single domain, such as ‘the elder daughter’ in (11), the CP-clause subject in (15b), and ‘sesame seeds’ as the subject of the CP in (10a). This can be perfectly explained with Pak’s account since there is no c-commanding relation between the subject and the following M-word.

- (13) *i kau-tai in nng e lang # m-thang sui-pian # ka lang khui*
 she urge them two E people cannot arbitrarily for someone open
mng #
 door
 ‘She urges them not to arbitrarily open the door for strangers.’

- (14) *ó- rìrì níí* (Anyanwu, 2012, (1c) on p.378)
 3sg CL eat-past food

²When pronouns are in contrastive focus, they could be realized either in their citation form or in their sandhi form but with greater vocal effort such as intensity or longer duration.

‘S/he ate food.’

- (15) a. i [sia = khah kin]_{VP}
 he write more fast
 ‘He writes faster.’
- b. [i sia]_{CP} # khah kin
 he write more fast
 ‘It would be faster for him to write.’
- (16) a. chO chit chhut = liok-iaN-phiN # lai khoaN
 rent one Cl video movie to watch
 ‘rent a video movie to watch’
- b. liok-iaN-phiN # chO chit chhut = lai khoaN
 video movie rent one Cl to watch (same as (a))
- c. liok-iaN-phiN # chit chhut # chO loa-choe chiN
 video movie one Cl rent how much money
 ‘How much does a (cassette of) video movie rent for?’

In (16a), one can observe that a DP of the form (demonstrative) + numeral + classifier + NP forms a single sandhi domain, and tone sandhi applies among the elements within it. Assume that the corresponding DP structure is $DP[D_{NumP}[Num_{CIP}[Cl NP]]]$, in which the NumP is the complement of the determiner, the CIP is the complement of the numeral, and the NP is the complement of the classifier. This embedded relationship suggests that no Concatenation statement is generated, and hence no sandhi boundary is found within a DP, as desired. When the NP is preposed to sentence-initial position and adjoins to IP as in (16b), it itself forms a single sandhi domain as the noun head does not c-command the following verb ‘buy’. The rest of the DP remains in situ and phrases together with the following infinitival

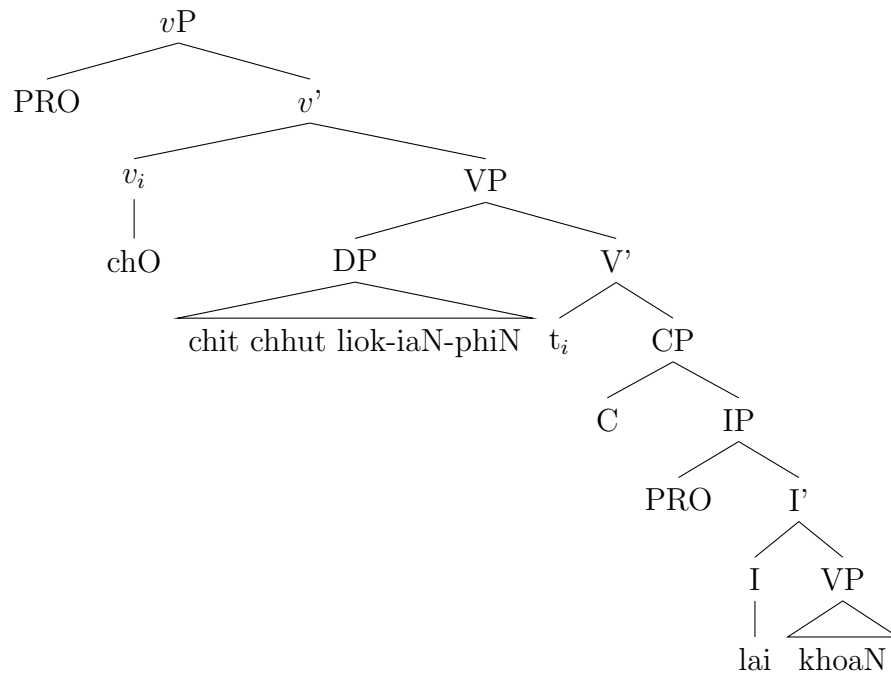
clause. The lack of a sandhi boundary after the classifier *chhut* is incompatible with Pak’s account to the extent that *chhut* certainly does not c-command the Infl *lai*. (16c) shows a case where the entire DP is preposed and the NP is further fronted to sentence-initially and both the NP and the “DP-leftover” *chit chhut* induce a sandhi break. The sandhi domain boundary after the NP is expected as the noun does not c-command the determiner, and the classifier similarly does not c-command the following verb, so according to Pak’s analysis a sandhi domain boundary is expected to occur after the DP-leftover, as borne out by the facts.

- (17) a. Chhun-Kiau # siong-sin = $_{CP}$ [A-Bin # boe lai]
 Chhun-Kiau believe A-Bin will not come
 ‘Chhun-Kiau believed that A-Bin would not come.’
- b. gua jin-ui = $_{CP}$ [A-Bin # kong # e khah u to-li]
 I think A-Bin say E more have sense
 ‘I think what A-Bin said makes more sense.’

A greater challenge for Pak’s approach lies in that some apparently biclausal structures in TSM can form a single sandhi domain, e.g. sentences with a matrix verb taking a purposive infinitival clause complement as in (16b), and those with a matrix-clause verb that introduces indirect speech, thought, or belief taking a finite clausal complement as in (17). Following Whelpton (1995), I assume that a purposive infinitive clause is adjoined to V’ and the corresponding structure for (16a) is shown in (18). A seemingly possible workaround would be to posit these clausal complements as what Pak referred to as reduced, sub-CP structures, which undergo spellout to-

gether with the matrix verb. Structures of this kind are found in Huave’s application of a High Tone Plateau rule and Luganda’s H-Tone Anticipation rule, but they do not typically include nonfinite purposive clause and finite complements of ‘say/know’ verbs, which is exactly what is shown in the TSM data.

- (18) chO chit chhut liok-iaN-phiN # lai khoaN
 rent one Cl video movie to watch



Another kind of biclausal structures nevertheless involves non-application of tone sandhi across clausal boundaries, as illustrated in (19). Here the CP, as a clausal adjunct, is located at the same position as a S-adverb, and thus the rightmost M-word within that CP does not c-command the first M-word to its right, so a sandhi boundary is expected right after this CP adjunct, a correct result.

- (19) a. [_{CP} sia chit phiN bun-chiong # chin-cheng,] # gua su-iau khah
 write this Cl essay before I need more
 che gian-kiu
 many research
 ‘Before writing this essay, I need to do more research.’
- b. gua chin thiam, # [_{CP} sO-i # lan lai tng#-khi0
 I very tired so we let go back
 ‘I’m tired, so let’s go home.’
- c. [_{CP} i na-si = boe = hiau tai-oan-oe] # li to ka i hoan-ek
 he if Neg know Taiwanese you then to him translate
 ‘If he doesn’t know Taiwanese, you translate for him.’
- d. [_{CP} i sui-lian # ni-ki # ia toa] # (put-ko #) sin-the # ia
 he although age quite big however body quite
 ho
 good
 ‘Although he is quite old, he is pretty healthy.’

For coordinate XP constructions, tone sandhi is blocked between XPs such as the APs in (20). Assume that *kah* ‘and’ or an empty operator is the head of an &P, and that the first AP *khoai-lok* is the specifier of the &P while the second AP is the complement. Since the first adjective ‘happy’ does not c-command the &P head *kah*, a sandhi boundary is expected to follow the AP, as desired. Also, no Concatenation statement is produced between *kah* and the second adjective since the left-hand member, i.e., the head of &P, is not a phrasal category. The sandhi domain boundary after the second AP is due to the presence of the modificational marker *e*.

- (20) in koe-tioh chin khoai-lok # (kah) peng-cheng # e seng-oah #
 They live very happy and quiet E life
 ‘They lead a very happy and quiet life.’

In sum, Pak's analysis of sandhi domains in TSM specified by a Phrase-left Concatenation rule covers most but not all of the TSM data. In particular, it fails to account for the lack of a sandhi domain boundary between the DP-leftover and the following infinitival clause when the NP base-generated in the DP is preposed to sentence-initial position, as shown in (16b), and requires several structural assumptions such as treating VP-adverbs preceding verbs as cases of adverb-incorporation, modifier adjuncts followed by nouns/adjectives as compounding between them, pronouns as clitics to subsequent M-words, and clausal complements as reduced sub-CP structures that undergo spellout together with the matrix verb. Last but foremost, it is unclear how her analysis can account for the fact that an utterance-final tone-bearing syllable, as well as a monosyllabic word in isolation, is always realized with its citation form. In the next section, I present my proposal based on the same theoretical framework as an alternative analysis of the tone sandhi rules in TSM.

6.4 Tone sandhi as Head-left Concatenation

Unlike Pak's approach, I assume citation forms being underlying and sandhi forms as the derived output from tone sandhi rules for the following reasons: first, a monosyllabic word pronounced in isolation is always realized with its citation form; second, the result of the priming experiment shows that native TSM speakers, especially old speakers, have stronger association with citation forms than with sandhi forms. Last but not least, in light of the potential tonal neutralization case, where the sandhi

forms for both citation 55 and 24 are said to be realized the same as 33, I argue that it is more reasonable to postulate citation variants as the underlying form and sandhi variants as derived; otherwise, speakers would have to be able to perfectly tell the two sandhi 33 variants apart.

Assume that spellout is triggered at each CP. I suggest that tone sandhi domains in TSM are determined by a head-left Concatenation rule, which identifies pairs of M-words X, Y where (i) X is left-adjacent to Y, and (ii) X c-commands Y. The proposed framework works as follows: within each spellout domain, the right edge of a sandhi domain is not inserted when there is a head-left Concatenation statement, i.e., for a Concatenation statement $A \frown B$, a sandhi domain boundary is not located between A and B, but instead occur after B when no further Concatenation statement exists between B and its following M-word; then a Chaining operation takes place to establish linear order across all the M-words between a pair of sandhi domain boundaries, and lastly, the phonological context-free rule in (21) applies from left to right within that chain of M-words.

(21) *Tone Sandhi Rule (TSR)*

$T \rightarrow T'$, where T is citation tone and T' is sandhi tone.

Following Pak's analysis of Huave and Luganda, I assume that matrix-clause preverbal subjects, preverbal objects, S-adverbs, and clausal adjuncts as in (19) occupy Spec,CP in TSM, since they all phrase separately from one another and from their following M-

word, and that sentences with multiple preverbal constituents have recursive layers of CP structure at the top edge of the clause. In this case, subjects and fronted objects as well as S-adverbs being in the CP ‘edge’ consistently form their own spellout domains. In other words, since spellout is triggered at every CP node, each clause-edge constituent is spelled out separately and thus is expected to form its own sandhi domain, as borne out by the facts.

In the case of VP-adverbs, no Concatenation statement is generated between a VP-adverb and its ensuing verb since the left-hand member of the *-statement, namely the VP-adverb, is not an overt M-word, and thus a sandhi domain boundary is expected to occur between the VP-adverb and its following verb, which is an incorrect result. Note that this is different from the case of Huave and Luganda in Pak (2008) since the manner adverbs in these two languages are usually right adjoined to the verb and thus phrase together with the verb as per the definition of head-left Concatenation. Recall that Pak’s phrase-left Concatenation approach in the previous section has exactly the same issue. In order to produce a Concatenation statement between the VP-adverb and its following verb, one would need to presume that the VP-adverb does not project at all before its left adjunction to VP/V’. Once the direct adjunction of the VP-adverb to the verb is supposed, no phrase-left Concatenation statement but a head-left Concatenation statement would be generated between them since the VP-adverb now c-commands its following verb, and obligatory tone sandhi application between the adverb and the following verb is predicted as desired.

In like manner, to account for the tone sandhi application between a modifier ad-

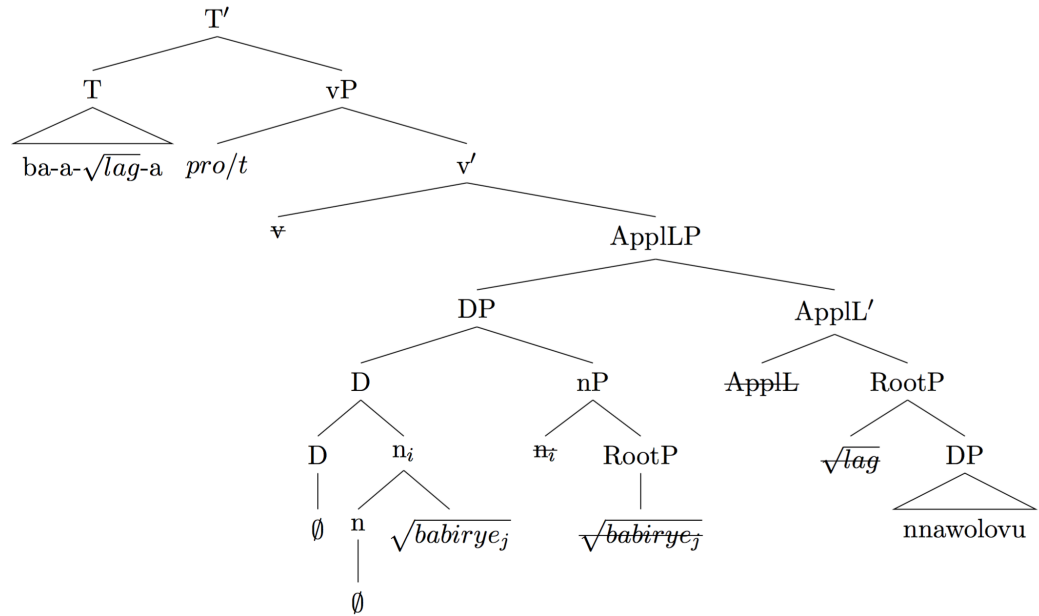
junct and its following modified head, one can postulate that a modifier adjunct does not project and is directly adjoined to the X'/XP level such that a c-commanding relationship exists between them, and consequently a head-left Concatenation statement between the two is generated, which ensure that no sandhi domain boundary is located between them. For phrases involving the modificational marker *e*, a sandhi boundary is always expected before *e* since no head-left Concatenation statement is generated between *e* and the internally complex XP preceding it.

The fact that tone sandhi applies between a verb and its first object to the right can be attributed to the fact that the verb c-commands the object, and hence there exists a head-left Concatenation statement between them. In Pak (2008) Luganda's Low Tone Deletion (LTD) analyzed as a Head-left Concatenation rule shows a case like this, as shown in (22), where a low applicative structure is assumed and the objects *Babirye* and *nnawolovu* 'chameleon' are internally complex DPs as in (23). Here only one Concatenation statement is produced: T[baalaga]∧D[Babirye], hence LTD only applies between *baalaga* and *Babirye* and the direct object *nnawolovu* forms a separate domain. For a sentence of the structure (S)-V-IO-DO, the indirect object groups together with the verb, while the direct object phrases separately as in (24a). No head-left Concatenation statement can be produced between the IO and the DO because the left-hand member of the *-statement between the two objects is an internally complex DP, not an overt M-word, and thus a sandhi domain boundary is predicted between the IO and the DO, as borne out by the attested data in Luganda and in TSM (e.g. (12)). When the IO is fronted, the verb groups with its following DO for

the same reason stated above, as shown in (24b).

- (22) (àbàlènzì) (bá-á-lág-á Bábíryè) (nnàwólòvù) (Pak, 2008, p.203 (55))
 2.boy 2-pst-show-ind 1.Babirye 1a.chameleon
 ‘The boys showed Babirye a chameleon.’

- (23) ba-a-lag-a Babirye nnawolovu (Pak, 2008, p.198 (51))
 2-pst-show-ind 1.Babirye 1a.chameleon
 ‘They showed Babirye a chameleon.’



- (24) a. (n-à-gúl-ír-á lújúújù) (káàwà) (Pak, 2008, p.203-204)
 1s-pst-buy-appl-ind 1a.drunkard 1a.coffee
 ‘I bought the drunkard some coffee.’
- b. (lùjúújù) (n-à-mú-gúl-írá káàwà)
 1a.drunkard 1s-pst-obj1-buy-appl-ind 1a.coffee
 ‘The drunkard, I bought him coffee.’

As presumed in previous section, pronouns in TSM cliticize to the word on their right and thus do not phrase their own sandhi domain. On the other hand, the tone sandhi

application among the elements within a DP of the form demonstrative + numeral + classifier + NP naturally derives from the layers of c-commanding relationship between the adjacent pair of elements. The sandhi boundary after the entire DP can be ascribed to the internally complex DP failing to be the left-hand member of a head-left Concatenation statement.

Considering the biclausal structures in TSM such as those in (16a), (16b) and (17). As described in previous section, I assume that the purposive infinitive *lai khoaN* ‘to watch’ is a reduced sub-CP, which is spelled out together with the matrix verb, and that the corresponding structure for (16a) is (18). A head-left Concatenation statement is produced between the verb and the DP, which accounts for the tone sandhi application between them. The sandhi domain boundary between the DP and the purposive clause is also predicted since the DP is not an overt M-word and thus fails to be the left-hand member of a *-statement.

The issue at hand lies in (16b), where the sandhi domain boundary after the DP is left out when the NP is moved to the sentence initial position, which has been problematic for all previous analyses and requires a workaround of one kind or another. To account for the tone sandhi application between the headless adnominal QP *chit chhut* and the purposive clause, a head-left Concatenation statement needs to be generated between them, which means that some transformation of the headless QP to an overt M-word is needed. Following Lin’s suggestion that QPs alone are regarded not as fully articulated noun phrases but are interpreted as manner adverbials in some cases, such as *chit liap* in (16b), I assume that *chit chhut* in (16b) should be

reanalyzed as a VP-adverb that does not project before its left adjunction to VP/V'. In this regard, *chit chhut* is an overt M-word c-commanding the purposive clause; thus a head-left Concatenation statement between them is generated and obligatory tone sandhi is predicted to apply between the two, as desired. An indirect evidence to support the reanalysis comes from (25), which is an example comparable to (16) but with a full DP *hit nng pun sio-soat* 'those two novels' as the direct object. While nothing new is about (25a) and (25c), as opposed to the lack of a sandhi domain boundary after *chit chhut* in (16b), the DP-leftover in (25b), i.e. *hit nng pun* 'those two Cl', is bounded by a sandhi domain boundary, which is accounted for by the fact that no head-left Concatenation statement can be generated between the DP-leftover and the Infl *lai*.

- (25) a. boe hit nng pun sio-soat # lai khoaN
 buy that two Cl novel to read
 'buy those two novels to read'
- b. sio-soat # boe hit nng pun # lai khoaN (same as (a))
 novel buy that two Cl to read
- c. sio-soat # hit nng pun # boe loa-choe chiN
 novel that two Cl buy how much money
 'How much does it cost to buy those two novels?'

The obligatory tone sandhi between 'think/say' verbs and the subject of the embedded CP in (17) can be explained by the fact that the phase specifier, i.e. the subject, is spelled out on a subsequent cycle along with the matrix-clause verbs.

Lastly, given that a coordinate XP construction can be represented as [_{&P} XP [_{&'}

& XP]], the non-application of tone sandhi between the first XP and *kah* ‘and’ in (20) arises from the fact that the internally complex XP cannot serve as the left-hand member of a head-left Concatenation statement. *Kah* as the head of &P c-commands the following XP in its complement, and therefore a head-left Concatenation statement is produced between them, which justifies the lack of a sandhi domain boundary after *kah*.

6.5 Discussion

To sum up, this chapter introduced two alternative analyses of the circumscription of tone sandhi domains in TSM adopting an articulated derivational model of the syntax-phonology interface within the distributed morphology framework, where phonological rules apply directly to the syntactic structure and thus domains for rule application are directly constrained by the underlying syntax, as opposed to the indirect-reference framework assumed by previous approaches such as Chen (1987) and Lin (1994).

The current proposal presented in Section 6.4 and Pak’s analysis sketched in Section 6.3 both cover a wide range of TSM data and share the same set of structural assumptions (including adverb-incorporation, compounding between modifier adjuncts and the following nouns/adjectives, cliticization of pronouns, and reduced sub-CP structures for clausal complements); however, they differ in the types of linearization operations involved — while Pak suggests that tone sandhi domains are determined by a Phrase-left Concatenation rule, I propose to use a Head-left Con-

catenation rule instead — and in the presumed underlying form. As Pak assumes the sandhi form as the underlying and the citation variant as the derived, extra effort is necessary to account for the fact that an utterance-final tone-bearing syllable, as well as a monosyllabic word in isolation, is always realized with the citation form. At last, regarding the residual problem of Pak’s analysis, i.e., the lack of a sandhi domain boundary between the DP-leftover and the following infinitival clause when the NP base-generated in the DP is preposed to sentence-initial position, as shown in (16b), I offer a solution that reanalyzes that DP-leftover as a VP-adverb directly left-adjoined to VP/V’ without further projection, following Lin’s suggestion, and provide an indirect evidence for support. Therefore, I argue in favor of my proposal of circumscribe sandhi domains in TSM with a Head-left Concatenation rule.

Chapter 7

Discussion

This chapter concludes the research by bringing together the results of Chapters 3, 4, 5, and 6 and presenting the overall discussion in relation to the research questions. Section 7.1 combines the results of the production and perception experiments of the tonal neutralization in TSM along with the result of the corpus study, and discusses the possibility of an ongoing sound change towards a complete tonal merger between the two sandhi 33 variants. Section 7.2 discusses whether the Tone Circle really exists by examining the realization of the Tone Circle in the corpus data, and considers whether tone sandhi domains in this language correspond to prosodic domains. Lastly, Section 7.3 concerns the interaction between different components of the grammar in the case of tone sandhi phenomena in TSM.

7.1 Incomplete tonal neutralization

The results of the production experiment in Section 3.1 and the spontaneous corpus data described in Chapter 5 lead to the same conclusion that the two sandhi 33 variants derived respectively from citation tones 55 and 24 are incompletely neutralized in that they tend to preserve the citation tone feature to some extent, i.e., produce the one derived from citation 55 with higher pitch than that from citation 24. Meanwhile, despite the small sampling size, the perception experiment reports that some subjects, old male speakers in particular, are able to distinguish the two sandhi variants perceptually. This is similar to the incomplete neutralization between German voiced /d/ and voiceless /t/ stops at syllable-final position. Production and perceptual experimental results of Port and Crawford (1989) show consistent phonetic difference between the two sets of final voiceless stops, e.g., longer duration of closure voicing and/or shorter stop closure in *Rad* ‘wheel’ than in *Rat* ‘advice’. They therefore argue that the final devoicing in German is represented with a combination of a categorical phonological rule that signals the discrete change in the voicing feature and a phonetic implementation rule that justifies continuously variable values of phonetic features.

The kind of systematic subphonemic differences between representations that are taken to be identical, such as those between the two sandhi 33 variants in TSM and those between the two sets of final voiceless stops in German, is unexpected

within the traditional theories of the phonetics-phonology interface (Pierrehumbert, 1990, among others), where phonological contrasts are assumed to be discrete and binary while the results of phonetic implementation are gradient and variable, and any phonetic variations on the surface are considered as context- or performance-induced variation, for instance, the less orthographic influence from the experimental design, the smaller the durational difference is found between German final voiceless stops, as reported by Fourakis and Iverson (1984). However, counterexamples are found; For instance, in the case of final devoicing in Dutch, subphonemic durational differences persist even with control of possible orthographic influence (Warner et al., 2004). Similarly, the result of the spontaneous corpus study, where no orthographic influence is present, shows systematic subphonemic differences in f_0 between the two sandhi 33 variants.

Cases of incomplete neutralization are often related to the phenomenon of near mergers, the situation where speakers consistently report that two classes of sounds are ‘the same’, yet consistently differentiate them in production at better than chance level (Labov et al., 1991). However, in the tonal neutralization case in TSM, speakers generally are unaware of tone sandhi and inconsistent production patterns are produced by different generations — complete neutralization is found in the data of young speakers (age range: 25-35) in the production experiment, while incomplete neutralization is observed not only in the production data of old speakers (age range: 52-66) but also in the spontaneous corpus data for both young (age range: 20-35 when recorded, currently 32-47) and old (age range: 50-65 when recorded, currently

62-77) age groups. If we compare these subject groups according to their current age range, an age-based variation in tonal neutralization is revealed, suggesting an ongoing sound change towards a complete tonal merger between the two sandhi 33 variants. Coupled with the result of the perception experiment, where most female subjects perform at/below chance level in distinguishing the two sandhi variants, this possible sound change in progress may be led by women speakers, a common observation made in sociolinguistics; notwithstanding, this is only a tentative explanation for the observed data and a carefully controlled apparent-time study covering more subjects and more age groups is in need before any conclusion can be reached.

7.2 Realization of Tone Circle

In general, the realization of citation and sandhi variants in the spontaneous corpus is compatible with the assigned tonal values in the sense that relative difference in pitch height is well captured, as reflected in Figure 5.2. For instance, the contour of citation 55 lies above that of sandhi 33 as in Figure 5.2(a), and the f0 curve of citation 21 lies down below that of sandhi 51 as in Figure 5.2(c). However, most variants with level tones, either high level 55 or mid level 33, exhibit some f0 declination effect, possibly due to tonal coarticulation. For citation and sandhi variants of the same tone, the result of the corpus study unsurprisingly shows that their f0 contours differ significantly, and that pitch contours produced by different generation groups are in parallel but with significant difference in f0 height, as shown in Figure 5.3, suggesting

a potential age-based difference in the sense that old speakers tend to speak in a more extreme way (i.e. with larger pitch range) as opposed to young speakers.

Regarding citation and sandhi variants that are taken to be identical on the surface, significant difference is found between the f₀ contours for both generation groups in the corpus study, and this systematic difference persists even when the phrasal position effect is approximately controlled by separating the data into prepausal and non-prepausal environments, especially in the non-prepausal environment, where presumably limited junctural effect is present. In light of the result of our preliminary attempt to control two conceivable confounding factors, i.e. the durational difference and the morpheme frequency effect, further systematic controls on the corpus data are needed to determine whether the observed patterns could serve as an evidence against the existence of the Tone Circle. A few ways to gain better control on the data include separating content words from function words, using slope estimation to factor out tonal coarticulation and f₀ declination, etc. Nevertheless, on account of the persistence of such systematic subphonemic differences in the experimental data and corpus data, I'm inclined to argue that the Tone Circle is just a phonological idealization that cannot be realized in ordinary speech.

Previous studies that associate tone sandhi domains in TSM with prosodic domains, such as Kuo (2013), find support in acoustic production data. For example, Peng (1997) observes f₀ final lowering and final lengthening in a domain-final position (i.e. where citation tones occur) as opposed to domain-initial and domain-medial positions (i.e. where sandhi tones occur). Pan and Tai (2006) reports that smoother

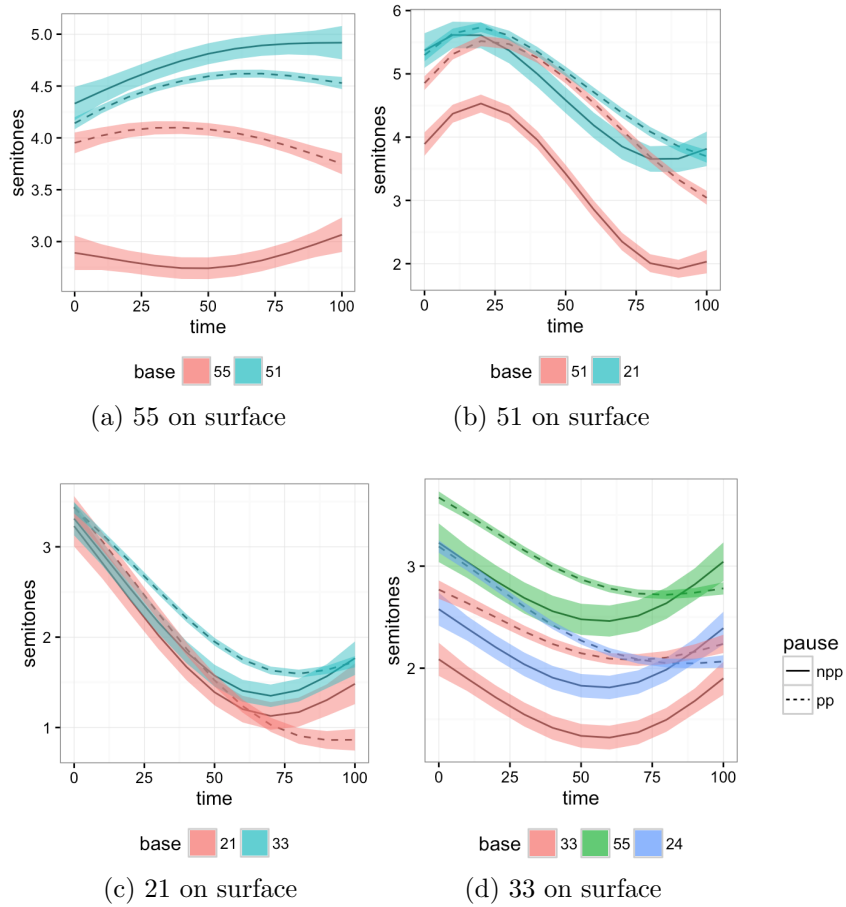


Figure 7.1: F0 curves of citation and sandhi variants that are taken to be identical grouped by pausing environments are modeled separately using SS-ANOVA. Red lines indicate citation variants and green/blue lines represent sandhi variants. In each case, the solid lines indicate the non-prepausal environment whereas the dashed lines are for the prepausal environment. The shaded grey areas around each mean f0 curve represents the 95% confidence interval.

f0 slope is found in a domain-final position than in a non-domain-final position. In the corpus study, the pre-boundary lengthening is also revealed in Figure 5.8, which shows the durational distributions of citation and sandhi variants grouped by tone category and generation groups. In each tone category, the rhyme duration of citation variants (which occur in domain-final positions) is significantly longer than that of

sandhi variants (that surface in domain-initial and domain-medial positions) for both age groups.

With respect to the slope difference in the f0 contours that occur in different positions, i.e. citation and sandhi variants, I compare the prepausal and non-prepausal data in the corpus study. In view of the fact that the data of two age groups are in parallel but differ significantly in f0 height, I randomly pick the data of young subjects in the comparison between pitch contours of citation and sandhi variants that are supposed to correspond in prepausal and non-prepausal environments, as illustrated in Figure 7.1, where red lines indicate citation variants and green/blue lines represent sandhi variants. Considering each pair of citation X and sandhi X in different pausing environments, for non-level surface forms, the slope of citation X is steeper than that of sandhi X in both prepausal and non-prepausal environments, as in Figures 7.1(b) and 7.1(c), which contradicts with the previous findings stating that citation forms (occur at domain-final positions) have smoother f0 slope than sandhi forms (surface in non-domain-final positions). This contradictory result can be very likely attributed to the natural continuity of spontaneous speech, in which fewer pauses are inserted between phrases or sandhi domains such that many citation variants do not precede pauses and are in non-prepausal environments. With this in mind, let's compare the same variant (lines that are in the same color) in different pausing environments. Again, focus on the non-level variants, as in Figures 7.1(b) and 7.1(c). Relatively smoother f0 slopes are found in prepausal environments (the dashed lines), except for the citation 21 case, in which the two pitch contours are

parallel until the final quarter of the rhyme duration, where a final rising is observed in the non-prepausal environment, which could be due to tonal coarticulation of the following tone. Whether this difference in slope is significant or not requires further statistical analysis. Despite that, the abovementioned observations altogether are in favor of Hayes's argument that tone sandhi domains in TSM do not refer to a prosodic domain.

7.3 Interaction between components of grammar

Within the theory of prosodic phonology assumed in Chen (1987) and Lin (1994), tone sandhi in Xiamen/ TSM is regarded as postlexical phrasal rules whose domains of application correspond to the level of the Phonological Phrase, a domain being larger than words but smaller than intonational phrases or utterances, and are constrained by syntactic structure to a certain extent. However, exactly how tone sandhi can get access to nonlocal information about the syntactic structure is not specified in such an indirect-reference model. The alternative proposed in Hayes (1990) regards tone sandhi in Xiamen as a precompiled rule stored in the lexicon that lexically inserts corresponding allomorphs in the relevant syntactic contexts but selects the corresponding allomorphs for insertion into syntax in the postlexical component. Tsay and Myers (1996), on the other hand, suggest that tone sandhi in TSM as a case of allomorphy couched in the Precompilation Theory only involves postlexical allomorph selection between citation and sandhi allomorphs that are listed in the lexicon. As

pointed out in Pak (2008), the tonal alternation is treated phonologically in these approaches, but tone sandhi assumed to be context-free applies blindly in the lexicon without knowing in which syntactic position, domain-final or non-domain-final, it occurs in. In this respect, the phonological rule is divorced from its morphosyntactic context, and thus nothing rules out a hypothetical situation where a sandhi allomorph is inserted at domain-finally rather than in non-domain-final positions. Furthermore, in order for lexical insertion frames to ‘see’ linearly adjacent words as well as to have access to nonlocal syntactic information, such proposals requires that we significantly relax the standard locality conditions on allomorphy, and thus admitting a number of unattested alternations, such as insert allomorph X at the right edge of a clausal adjunct, if the following word is a vowel-initial adjective.

Cases of syntax-phonology domain mismatches, such as the one in (20), where domains of the tone sandhi phrasal rules in Xiamen are not equivalent to syntactic constituents, have been used to support the idea that underlies an indirect-reference model, that is, syntactic effects in phonology can be mediated by a prosodic structure. At the same time, these examples are claimed to serve as evidence against direct-reference models assumed in this dissertation, in which phonological domains are syntactic objects at different stages of the derivation, rather than derived prosodic constituents. Note that this reasoning is only valid under the assumption that phonological domains must be syntactic constituents in direct-reference models. However, such an assumption is neither grounded nor necessary in that direct-reference theories can in principle use any type of information that is available in the syntax to define

phonological domains, and that phonological rules apply directly to the spelled-out content of each cycle minus what has already been spelled out on previous cycles, which then gives rise to some objects that are not syntactic constituents as phonological domains.

In the DM-based direct-reference model assumed in this dissertation, syntactic structures are built up and processed in cycles and syntax-sensitive phrasal phonological rules such as French liaison and Hausa final vowel shortening, apply directly to the syntactic structure as it happens to exist at the given point in PF, and syntactically determined domains for rule application thus ‘come for free’, which appears to be a good fit for the tone sandhi phenomena in TSM. I therefore propose that tone sandhi in TSM should be analyzed as a phrasal phonological rule within this framework. Specifically, tone sandhi domains in this language are determined by a head-left Concatenation rule, which identifies pairs of M-words X, Y where (i) X is left-adjacent to Y, and (ii) X c-commands Y. Within each spellout domain, the right edge of a sandhi domain is not inserted when there is a head-left Concatenation statement, i.e., for a Concatenation statement $A \frown B$, a sandhi domain boundary is not located between A and B, but instead occur after B when no further Concatenation statement exists between B and its following M-word; then a Chaining operation takes place to establish linear order across all the M-words between a pair of sandhi domain boundaries, and lastly, the phonological context-free rule of the following form,

$T \rightarrow T'$, where T is a citation tone and T' is a sandhi tone,
applies from left to right within that chain of M-words.

Chapter 8

Conclusion

In this dissertation, I aimed to thoroughly investigate various aspects of the tone sandhi phenomena in Taiwan Southern Min, including the tonal neutralization between two sandhi 33 variants derived respectively from citation 55 and 24 variants, priming effects induced by citation or sandhi primes, the realization of the Tone Circle, the circumscription of sandhi domains, and how tone sandhi relates to syntactic information in the syntax-phonology interface.

The findings and implications of this dissertation can be summarized as follows. First, the results of Chapter 3 showed an age-based acoustic variation in the sense that while complete neutralization was found in the data of young speakers, old speakers preserved the difference in the citation forms and consistently produced the sandhi 33 variant derived from citation 55 with higher pitch than that derived from citation 24. Coupled with the result of the perception experiment, a sound change in progress towards a near-merger or complete tonal merger was suggested between the two sandhi 33 variants, which might be led by female speakers. Certainly, a care-

fully controlled apparent-time study covering more subjects and more age groups is in need before any conclusion can be reached. Chapter 4 illustrated that an overlap in both the underlying and surface tones caused inhibition in the citation target case, while an underlying-tone overlap and a surface-tone overlap individually produced a facilitation priming effect. Meanwhile, the reaction time results demonstrated that old subjects had stronger association with citation variants than with sandhi variants, indicating that citation form is more underlyingly represented. Chapter 5 examined the realization of the Tone Circle in a corpus study and suggested that the Tone Circle is merely a phonological idealization that cannot be realized in ordinary speech in light of the systematic subphonemic difference in f_0 between citation X and sandhi X that are taken to be identical even with some control of conceivable confounding factors. In addition, a potential age-based difference in terms of the pitch range used by speakers of different generations was revealed. Lastly, Chapter 6 proposed an alternative account for the circumscription of sandhi domains in a DM-based direct-reference model and analyzed tone sandhi in this language as a head-left Concatenation rule.

While this dissertation provides a comprehensive picture of the tone sandhi phenomena in Taiwan Southern Min, there remain a number of open questions that I wish to answer with future research. First, although the spontaneous speech corpus covers a wide range of speakers, no evidence is found to support the ongoing sound change towards a near- or complete tonal merger reported in the tonal neutralization study. This is primarily due to the fact that the corpus was constructed about fifteen years ago and thus is lack of younger generation that is comparable to the young

age group in the production experiment. Additionally, as noted in the discussion, further systematic controls can be applied to the corpus data, such as calculating slope estimation to factor out tonal coarticulation and f0 declination effects, and perhaps using TF-IDF (term frequency inverse document frequency) in place of simple morpheme frequency to examine potential frequency effects. Other questions such as “whether the observed systematic subphonemic difference in f0 between citation X and sandhi X can be perceived by native speakers?” and “how exactly the gradient phonetic properties are encoded in the phonological component of the grammar” remain to be answered in future work.

Furthermore, although the result of the priming experiment suggests that citation form is more underlyingly represented in view of the fact that old subjects had stronger association with citation variants than with sandhi variants, it remains unknown as to how native speakers acquire this complex tonal alternation which shows phonetic arbitrariness and phonological opacity, and how these tones are represented in phonology. Studies like Li and Chen (2015) using non-invasive electroencephalography (EEG) to investigate the effect of tonal variation on the mental representation and neural processing of lexical tones may provide a future research direction to probe into the mental representation of citation and sandhi variants in Taiwan Southern Min.

Last but not least, it is interesting to see how the current proposal for the circumscription of sandhi domains in TSM can be applied to further tone sandhi cases in other Chinese dialects, including other Southern Min dialects such as Fuzhou and

Chaozhou, and perhaps Shanghainese and other Wu dialects (domains of application are said to be defined with reference to stress-feet), and in what ways Chinese dialects may vary with respect to the sets of structural assumptions to be imposed, the types of linearization operations involved, and what kinds of syntactic objects count as phases. I hope to have provided a working model for investigating some of these questions in future work.

Appendix A

Orthography for Taiwan Southern Min transcription

As there is no official written form for Taiwan Southern Min, various attempts have been made in the past trying to establish a standard writing system for this vernacular either for cultural or for instructional purposes. Currently available writing schemes for Taiwan Southern Min can be classified as character-based orthography such as the collection of Chinese characters selected by Taiwanese Ministry of Education, or script based on Latin, including *P'eh-ōe-jī* ‘vernacular writing’ (POJ) Romanization system¹, *Taiwan Language Phonetic Alphabet* (TLPA)², and *Tâi-lô* Romanization³.

Among them, as one the oldest transcription systems for the TSM variant and the

¹Once as a method for Western missionaries to spread church teachings as well as document the locally spoken language in mid-19th century, POJ is also referred to as Church Romanization or Missionary Romanization.

²TLPA is a revision of POJ devised in late 1990s by the Linguistic Society of Taiwan. It includes three systems for transcribing Taiwan Southern Min, Taiwan Hakka and the Formosan languages, the ethnic languages of the aboriginal tribes of Taiwan, respectively.

³*Tâi-lô*, short for *Tâi-uân Bân-lâm-gí Lô-má-jī Phing-im Hong-àn* ‘Taiwan Southern Min Roman Script Phonetic Transcription Scheme’, is a combination of the POJ and TLPA Romanization schemes and is promulgated by Taiwanese Ministry of Education in 2006.

form in which a large body of historical documents and literature has been produced, POJ remains the most popular romanization scheme, especially for online users (Lin, 2015).

The orthography used in this dissertation is primarily based on POJ along with some modifications in order to avoid using special characters such as o^\cdot for the mid back open vowel [ɔ] and $\text{n}^{\text{̃}}$ for nasalization on vowels. While POJ resorts to diacritics for tone markings (e.g. a, á, à, ah, â, ā, ǎh representing a word with vowel a in tones 55, 51, 21, 21c, 24, 33, and 53c respectively), here pairs of numerical values are used to directly reflect the pitch of the tones. All contrastive vowels are enumerated in (1) along with the corresponding IPA, and consonants are shown in Table A.1. There are two syllabic nasals as listed in (2). Only the nasals and voiceless stops are allowed as syllable codas.

(1) Vowel sounds:

- a. Monophthongs: i [i], e [e], a [a], o [ə], u [u], O^4 [ɔ]
- b. Diphthongs: ai [ai], au [au], ia [ia], io [io], iu [iu], oa [ua], oe [ue], ui [ui]
- c. Triphthongs: iau [iaʊ], oai [uai]
- d. Nasalized vowels: indicated by N^5 , e.g. iaN [iã]

(2) Syllabic nasals: m [m̩], ng [ŋ̩]

⁴This mid back open vowel is denoted as o^\cdot , a dot on the upper right of o, in Peh-oe-ji.

⁵Nasalization is denoted as a raised n to the vowel, $\text{V}^{\text{̃}}$, in Peh-oe-ji.

Table A.1: The list of consonants in Taiwan Southern Min.

		Bilabial		Dental-Alveolar		Alveopalatal	
		voiceless	voiced	voiceless	voiced	voiceless	voiced
Nasal			m		n		
Stop	unaspirated	p	b	t			
	aspirated	ph		th			
Affricate	unaspirated			ch	j	chi	ji
	aspirated			chh		chhi	
Fricative				s		si	
Lateral					l		

		Velar		Glottal
		voiceless	voiced	voiceless
Nasal			ng	
Stop	unaspirated	k	g	h(coda) [ʔ]
	aspirated	kh		
Affricate	unaspirated			
	aspirated			h
Fricative				
Lateral				

Appendix B

Stimuli of tonal neutralization production experiment

The study included 10 target minimal pairs of monosyllabic words, all of which are embedded in minimal pairs of sentence, as listed below. Target words that are realized as sandhi 33 on surface are marked in bold and annotated with the citation tone.

- (1) a. beh kio i **poe24** kau tang-si
want ask him accompany until when
'How long do you still want him to accompany you?'
- b. beh kio i **poe55** kau tang-si
want ask him fly until when
'How long do you still want him to travel by flight?'
- (2) a. thao-ching tloh si toa- **ping24** -soaN
ahead right is big flat mountain
'There's Mount Taiping (a name of a mountain in Taiwan) right ahead.'
- b. thao-ching tloh si toa **bing55** soaN
ahead right is big ice mountain
'There's a big iceberg right ahead.'

- (3) a. i **kiaN24** kah khi cchoan beh khi0-lai0
 (s)he walk till breath pant cannot up
 ‘(S)he is out of breath while walking.’
- b. i **kiaN55** kah khi cchoan beh khi0-lai0
 (s)he frightened till breath pant cannot up
 ‘(S)he is too scared to catch his breath’
- (4) a. i beh **koaN24** i-a
 (s)he want tall chair
 ‘(S)he wants a tall chair.’
- b. i beh **koaN55** i-a
 (s)he want donate chair
 ‘(S)he wants to donate chairs.’
- (5) a. khoaN i **khoan24**-han kao to-ui
 see/depend (s)he authority to where
 ‘It depends on the scope of his authority.’
- b. khoaN i **khoan55**-han kao to-ui
 see (s)he limits of tolerance to where
 ‘Let’s see to what point (s)he can still be tolerant.’
- (6) a. che si **kun24**-tui e oah-tang
 this is group E activity
 ‘This is a group activity.’
- b. che si **kun55**-tui e oah-tang
 this is military E activity
 ‘This is a military activity.’
- (7) a. tan in **sing24**-kao liao-au
 wait they close a deal after
 ‘(Let’s) wait until they close the deal.’
- b. tan in **sing55** kao liao-au
 wait they first submit after
 ‘(Let’s) wait until they submit it.’

- (8) a. u **ti24**-thao ti hia
 there is hoe at there
 ‘There is a hoe over there.’
- b. u **ti55**-thao ti hia
 there is jerk at there
 ‘There is a jerk over there.’
- (9) a. i tloh si **siu24** phio e lang
 (s)he exactly is sell ticket E person
 ‘(S)he is the person who sells the tickets.’
- b. i tloh si **siu55** phio e lang
 (s)he exactly is collect ticket E person
 ‘(S)he is the ticket-taker.’
- (10) a. i **thao24** chit kai tloh hO lang liah khi0-lai0
 (s)he first one time just *passive ptcl* someone catch up(direction
 compl.)
 ‘(S)he got caught right at the first time.’
- b. i **thao55** chit kai tloh hO lang liah khi0-lai0
 (s)he steal one time just *passive ptcl* someone catch up
 ‘(S)he got caught stealing right at the first time.’

Bibliography

Ogbonna Anyanwu. Pronominal subject clitics in igbo. *Theory and Practice in Language Studies*, 2(2):377, 2012.

Douglas Bates, Martin Maechler, Ben Bolker, and Steven Walker. *lme4: Linear mixed-effects models using Eigen and S4*. URL <https://github.com/lme4/lme4/> <http://lme4.r-forge.r-project.org/>. R package version 1.1-7.

Paul Boersma and David Weenink. *Praat: doing phonetics by computer [Computer program]*, 2015. URL <http://www.praat.org/>. Version 5.4.06, retrieved 21 February 2015.

Yuen Ren Chao. A system of tone letters. *Le Maitre Phonétique*, 30:24–27, 1930.

Yuen Ren Chao. *Mandarin primer: An intensive course in spoken Chinese*. Harvard University Press, 1948.

Jan Charles-Luce and Daniel A. Dinnsen. A reanalysis of catalan devoicing. *Journal of Phonetics*, 15:187–190, 1987.

Mao-Hsu Chen, Jean-Philippe Goldman, Ho-hsien Pan, and Janice Fon. Taiwan min

- easyalign: an automatic phonetic alignment tool under praat. In *Proceedings of the Workshop on New Tools and Methods for Very-Large-Scale Phonetics Research*, pages 109–112, 2011.
- Matthew Y Chen. The syntax of xiamen tone sandhi. *Phonology*, 4(01):109–149, 1987.
- Matthew Y. Chen. Argument vs adjunct: Xiamen tone sandhi revisited. 1992. Manuscript.
- Matthew Y Chen. *Tone sandhi: Patterns across Chinese dialects*, volume 92. Cambridge University Press, Cambridge, UK, 2000.
- Szu-Wei Chen, James Myers, and Jane Tsay. Testing the allomorph selection hypothesis in taiwanese tone sandhi. In *Proceedings of the 22nd North America Conference on Chinese Linguistics (NACCL-22) and the 18th Annual Meeting of the International Association of Chinese Linguistics (IACL-18)*, pages 283–300, 2010.
- Chin-Chuan Cheng. Domains of phonological rule application. In Jerrold M. Sadock and Anthony L. Vanek, editors, *Studies Presented to Robert B. Lees by His Students*, pages 39–59. Linguistic Research, Edmonton, 1970.
- Robert L Cheng. Tone sandhi in taiwanese. *Linguistics*, 6(41):19–42, 1968.
- Robert L Cheng. Some notes on tone sandhi in taiwanese. *Linguistics*, 11(100):5–25, 1973.

- Yu-Fu Chien, Joan A Sereno, and Jie Zhang. Priming the representation of mandarin tone 3 sandhi words. *Language, Cognition and Neuroscience*, 31(2):179–189, 2016a.
- Yu-Fu Chien, Joan A Sereno, and Jie Zhang. Priming the representation of taiwanese tone sandhi words. In *5th International Symposium on Tonal Aspects of Languages, Buffalo, NY*, 2016b.
- Bien-ming Chiu. The phonetic structure and tone behaviour in hagu (commonly known as the amoy dialect) and their relation to certain questions in chinese linguistics. *T'oung Pao*, 28(1):245–342, 1931.
- Noam Chomsky. *Barriers*, volume 13. MIT press, 1986.
- Noam Chomsky. Minimalist inquiries: The framework (mitopl 15). *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, pages 89–155, 2000.
- Noam Chomsky. Derivation by phase (mitopl 18). *Ken Hale: A Life in Language*, pages 1–52, 2001.
- Lisa Davidson. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance a. *The Journal of the Acoustical Society of America*, 120(1):407–415, 2006.
- San Duanmu. Metrical and tonal phonology of compounds in two chinese dialects. *Language*, pages 225–259, 1995.
- San Duanmu. Wordhood in chinese. *New approaches to Chinese word formation:*

- Morphology, phonology and the lexicon in modern and ancient Chinese*, pages 135–196, 1998.
- David Embick. *Localism versus globalism in morphology and phonology*, volume 60. Cambridge, Mass: MIT Press, 2010.
- David Embick and Rolf Noyer. Distributed morphology and the syntax-morphology interface. *The Oxford handbook of linguistic interfaces*, pages 289–324, 2007.
- Janice Fon. A preliminary construction of taiwan southern min spontaneous speech corpus. Technical Report NSC-92-2411-H-003-050, National Science Council, Taiwan, 2004.
- Marios Fourakis and Gregory K Iverson. On the incomplete neutralization of german final obstruents. *Phonetica*, 41(3):140–149, 1984.
- Stephen D Goldinger. Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of experimental psychology: Learning, memory, and cognition*, 22(5):1166, 1996.
- Stephen D Goldinger. Echoes of echoes? an episodic theory of lexical access. *Psychological review*, 105(2):251, 1998.
- Government Information Office. The republic of china yearbook 2012, 2012.
- J Brown Grier. Nonparametric indexes for sensitivity and bias: computing formulas. *Psychological bulletin*, 75(6):424, 1971.

- Chong Gu. *Smoothing spline ANOVA models*, volume 297. Springer Science & Business Media, 2013.
- Chong Gu et al. Smoothing spline anova models: R package gss. *Journal of Statistical Software*, 58(5):1–25, 2014.
- Kenneth Hale and Elisabeth Selkirk. Government and tonal phrasing in papago. *Phonology*, 4(1):151–183, 1987.
- Morris Halle and Alec Marantz. *Distributed morphology and the pieces of inflection*, pages 111–176. The MIT Press, 1993.
- Morris Halle and Alec Marantz. Some key features of distributed morphology. *MIT Working Papers in Linguistics*, 21:275–288, 1994.
- Bruce Hayes. The prosodic hierarchy in meter. *Rhythm and meter. Phonetics and phonology*, 1:201–260, 1989.
- Bruce Hayes. Precompiled phrasal phonology. *The phonology-syntax connection*, ed. by Sharon Inkelas and Draga Zec, pages 85–108, 1990.
- Hsin-I Hsieh. The psychological reality of tone sandhi rules in taiwanese. In *6th meeting of the Chicago Linguistic Society*, pages 489–503, 1970.
- Cheng-Teh James Huang. *Logical relations in Chinese and the theory of grammar*. PhD thesis, Massachusetts Institute of Technology, 1982.

- Cheng-Teh James Huang. Phrase structure, lexical integrity, and chinese compounds. *Journal of the Chinese Language Teachers Association*, 19(2):53–78, 1984.
- Ún-giân Iûⁿ. “Taiwen huawen xianshangzidian jian-zi jishu ji shiyongcingxing tantao [Construction and utilization of Taiwanese-Chinese online dictionary]”. In *Proceedings of the 3rd International Conference on Internet Chinese Education*, pages 132–141, 2003.
- Ray S Jackendoff. *X syntax: A study of phrase structure*. MIT press, 1977.
- Jola Jakimik, Ronald A Cole, and Alexander I Rudnicky. Sound and spelling in spoken word recognition. *Journal of Memory and Language*, 24(2):165–178, 1985.
- Jörg D Jescheniak and Willem JM Levelt. Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4):824, 1994.
- Ellen M Kaisse. Toward a typology of postlexical rules. *The phonology-syntax connection*, pages 127–143, 1990.
- Paul Kiparsky. Phonological representations. In Osamu Fujimura, editor, *Three dimensions of linguistic theory*. TEC, Tokyo, 1973.
- Paul Kiparsky. Some consequences of lexical phonology. *Phonology*, 2(1):85–138, 1985.

- Chen-Hsiu Kuo. *Perception and Acoustic Correlates of the Taiwanese Tone Sandhi Group*. PhD thesis, University of California, Los Angeles, 2013.
- William Labov, Mark Karen, and Corey Miller. Near-mergers and the suspension of phonemic contrast. *Language variation and change*, 3(1):33–74, 1991.
- Xiaoqing Li and Yiya Chen. Representation and processing of lexical tone and tonal variants: Evidence from the mismatch negativity. *PloS one*, 10(12):e0143097, 2015.
- Jo-wang Lin. Lexical government and tone group formation in xiamen chinese. *Phonology*, 11(02):237–275, 1994.
- Philip T Lin. *Taiwanese Grammar: A Concise Reference*. Greenhorn Media, 2015.
- T-H Jonah Lin and W-W Roger Liao. Purposives in mandarin chinese and their syntactic properties. *Unpublished ms., National Tsing Hua University, Taiwan*, 2008.
- James Meyers and Gregory R Guy. Frequency effects in variable lexical phonology. *University of Pennsylvania Working Papers in Linguistics*, 4(1):14, 1997.
- James Myers and Jane Tsay. Neutralization in taiwan southern min tone sandhi. *Interfaces in Chinese phonology: Festschrift in honor of Matthew Y. Chen on his 70th birthday*, pages 47–78, 2008.
- Allard Jongman Olga Dmitrieva and Joan Sereno. Phonological neutralization by

- native and non-native speakers: The case of russian final devoicing. *Journal of Phonetics*, 38:483–492, 2010.
- Shu-chen Ou and Y. E. Hsiao. Cong yunlv yinyunxue kan taiwan minnanyu de qingsheng xianxiang. neutral tone in taiwan southern min: a prosodic phonology perspective. *Chinese Phonology*, (6):865–895, 1997.
- Marjorie Pak. *The postsyntactic derivation and its phonological reflexes*. PhD thesis, University of Pennsylvania, 2008.
- Ho-hsien Pan and Yi-hsin Tai. Boundaries and tonal articulation in taiwanese min. In *Proceedings of Speech Prosody*, volume 51, 2006.
- Shu-hui Peng. Production and perception of taiwanese tones in different tonal and prosodic contexts. *Journal of Phonetics*, 25:371–400, 1997.
- Shu-Hui Peng. Lexical versus phonological representations of mandarin sandhi tones. *Papers in laboratory phonology V: Acquisition and the lexicon*, pages 152–167, 2000.
- Janet Pierrehumbert. Phonological and phonetic representation. *Journal of phonetics*, 18(3):375–394, 1990.
- David B Pisoni and Paul A Luce. Acoustic-phonetic representations in word recognition. *Cognition*, 25(1-2):21–52, 1987.
- Robert F Port. The discreteness of phonetic elements and formal linguistics: response to a. manaster ramers. *Journal of Phonetics*, 24(4):491–511, 1996.

- Robert F. Port and Penny Crawford. Incomplete neutralization and pragmatics in german. *Journal of Phonetics*, 17(4):257–282, 1989.
- Robert F Port and Adam P Leary. Against formal phonology. *Language*, pages 927–964, 2005.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014. URL <http://www.R-project.org/>.
- Maria-Luisa Rivero. Adverb incorporation and the syntax of adverbs in modern greek. *Linguistics and philosophy*, 15(3):289–331, 1992.
- Catherine Rudin. Phonetic evidence for a phonological rule: G-deletion in turkish. *Research in Phonetics*, 1:217–232, 1980.
- Elisabeth Selkirk. On derived domains in sentence phonology. *Phonology*, 3:371–405, 1986.
- Yen-Liang Shue, Patricia Keating, Chad Vicenik, and Kristine Yu. Voicesauce: A program for voice analysis. *Energy*, 1(H2):H1–A1, 2010. URL <http://www.ee.ucla.edu/~spapl/voicesauce>.
- Louisa M Slowiaczek and Daniel A Dinnsen. On the neutralizing status of polish word final devoicing. *Journal of Phonetics*, 13(3):325–341, 1985.
- Jane Smirniotopoulos and Brian D Joseph. On so-called ”adverb-incorporation” in

- modern greek. In *Greek Linguistics' 95: Proceedings of the 2nd International Conference on Greek Linguistics*, pages 117–128, 1997.
- Richard Sproat. Looking into words. In *Studies in lexical phonology*, pages 173–195. Elsevier, 1993.
- Tim Stowell. Subjects, specifiers, and x-bar theory. *Alternative conceptions of phrase structure*, pages 232–262, 1989.
- Chih-Chen Tang. *Chinese Phrase Structure and the extended X'-theory*. PhD thesis, Cornell University, 1990.
- Jane Tsay and James Myers. Taiwanese tone sandhi as allomorph selection. In *Annual Meeting of the Berkeley Linguistics Society*, volume 22, pages 395–405, 1996.
- Jane Tsay, Jan Charles-Luce, and Yung-Song Guo. The syntax-phonology interface in taiwanese: acoustic evidence. In *Proceedings of the 14th International Congress of Phonetic Sciences*, volume 3, pages 2407–2410, 1999.
- Suhchuan Jane Tsay. *Phonological pitch*. PhD thesis, The University of Arizona, Tucson, AZ, 1994.
- Shu-Chuan Tseng. Lexical coverage in taiwan mandarin conversation. *International Journal of Computational Linguistics and Chinese Language Processing*, 1(18):1–18, 2013.

- Tung-Ho Tung, YL Chao, and YS Lan. A south min dialect of taiwan. *Monographs Series*, 1967.
- H Samuel Wang. *Experimental studies in Taiwanese phonology*. Crane Publishing Company, 1995.
- Sheng-Fu Wang and Janice Fon. A taiwan southern min spontaneous speech corpus for discourse prosody. *The Proceedings of Tools and Resources for the Analysis of Speech Prosody, Aix-en-Provence, France*, pages 20–23, 2013.
- Yuedong Wang. *Smoothing splines: methods and applications*. CRC Press, 2011.
- Yuedong Wang, Chunlei Ke, and Morton B Brown. Shape-invariant modeling of circadian rhythms with random effects and smoothing spline anova decompositions. *Biometrics*, 59(4):804–812, 2003.
- Natasha Warner, Allard Jongman, Joan Sereno, and Rachèl Kemps. Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from dutch. *Journal of phonetics*, 32(2):251–276, 2004.
- Matthew James Whelpton. *The syntax and semantics of infinitives of result in English*. PhD thesis, University of Oxford, 1995.
- Jie Zhang, Yuwen Lai, and Craig Sailor. Modeling taiwanese speakers knowledge of tone sandhi in reduplication. *Lingua*, 121(2):181–206, 2011.

Zheng-sheng Zhang. *Tone and tone sandhi in Chinese*. PhD thesis, The Ohio State University, 1988.

Xiaolin Zhou and William Marslen-Wilson. The abstractness of phonological representation in the chinese mental lexicon. *Cognitive Processing of Chinese and other Asian languages*, pages 3–26, 1997.

Elizabeth Zsiga. *Features, gestures, and the temporal aspects of phonological organization*. PhD thesis, 1993.