Tone Sandhi in Ganyu Mandarin

By

Yiming Gu

M.A., Fudan University, 2016

Thesis

Submitted in partial fulfillment of the requirement for the
Degree of Master of Arts in the Department of Cognitive, Linguistic and
Psychological Sciences at Brown University

PROVIDENCE, RHODE ISLAND

MAY 2018

This thesis by Yiming Gu is accepted in its present form
by the Department of Cognitive, Linguistic and Psychological Sciences as satisfying the
thesis requirements for the degree of Master of Arts

Date ___________________ ___________________________

Chelsea Sanker, Advisor

Approved by the Graduate Council

Date ___________________ ___________________________

Andrew G. Campbell, Dean of the Graduate School

Acknowledgments

I would first like to thank my thesis advisor Prof. Chelsea Sanker for her continuous support for my study and research. I am grateful to her patient guidance throughout my study at Brown University.

I would also like to thank Prof. Uriel Cohen Priva and Prof. James Morgan for their insightful comments and encouragement. Professor Cohen Priva showed me how to prepare academic writing in a reader-friendly manner, which is extremely useful for a novice academic writer.

My sincere thanks also go to Prof. Pauline Jacobson, graduate advisor of the linguistics program, and Prof. Scott AnderBois. They have provided invaluable advice on my career path.

Contents

1 Introduction

Tone sandhi plays a crucial role in Ganyu Mandarin[1]. It is systematic, productive, widely attested in all sub-dialects, and seemingly complicated.

Tone sandhi in Ganyu is sensitive to several levels of the prosodic structure. The sentence provided in (1) could be grouped into two phonological phrases, with the subject ‘sjaw tʃjaŋ’ as the first one, and the predicate ‘p^{h}ən t^{h}jɤŋ jin jwɤ’ as the second one. Within each phonological phrase, all syllables except the last one have a surface realization which is different from their citation forms. I do not claim that citation forms are the underlying forms (cf. Chapter 4).

(1)	sjaw	tʃjaŋ	p^{h}ən	t^{h}jɤŋ	jin	jwɤ	
	LM	L^{C}M	HL	L^{C}M	L^{C}M	L^{C}M	(citation tones)
	[HL	**L^{C}M**]	[L	HH	HH	**L^{C}M**]	(surface realization)

Xiao Zhang PROG listen music
‘Xiao Zhang is listening to music.’

Ganyu Mandarin has four lexical tonal categories. Their citation tones are shown in Table 1.

Category	I	II	III	IV
My notation	L^{C}M	HH	LM	HL
Digit notation	213	55	324	51
Traditional name	yin level	yang level	rising	leaving
Example	jin	jɑŋ	ʃjɑŋ	k^{h}y

Table 1: Citation tones

The digit notation is taken from Sū and Wàn (2011). Although the 1-5 digit notation is commonly used in descriptive Sino-Tibetan linguistics, its limitation and redundancy has been discussed in Zhu (2010). My notation follows the practice in generative phonological analyses of Chinese tones. L, M and H correspond to Low, Mid and High pitches. ^{C} signals a short period of creakiness. Creakiness in tones has been observed in many Chinese dialects,

[1] Ganyu Mandarin is spoken by approximately 1 million people in Ganyu District and surrounding areas of Lianyungang Municipality, Jiangsu Province, China. It is classified as a Central Plains (Zhongyuan) Mandarin in Wurm, Li, and Baumann (1987). In Jiǎng (2005), it is classified as a Jiao-Liao Mandarin. Its onsets system is more similar to those found in Jiao-Liao Mandarin, and its rime system is more similar to those found in Central Plains Mandarin. Ganyu Mandarin has several sub-dialects, whose tone sandhi patterns may differ significantly. This paper focuses on Menhe sub-dialect, spoken by approximately 20 thousand people.

including Lianyungang Jiang-Huai Mandarin, a neighboring dialect of Ganyu Mandarin (T. Zhang & Zhu, 2012).

Tone I and Tone II are developed from different initial consonants, with Tone I corresponding to historically voiceless initial consonants and Tone II corresponding to historically voiced initial consonants (c.f *Sòng Běn Guǎng Yùn*, 'Extended Rhyme Book of Song Dynasty', compiled by Chén Péngnián, etc. in 1008). Many historical linguists also believe that Tone III and Tone IV are developed from the coda [ʔ] and [s] (Pan, 1982; Baxter & Sagart, 2014). Therefore, all tones in modern Chinese language are derived from historically segmental elements, and more specifically, consonants.

Some sample pitch realizations of each tones are shown in Figure 1[2].

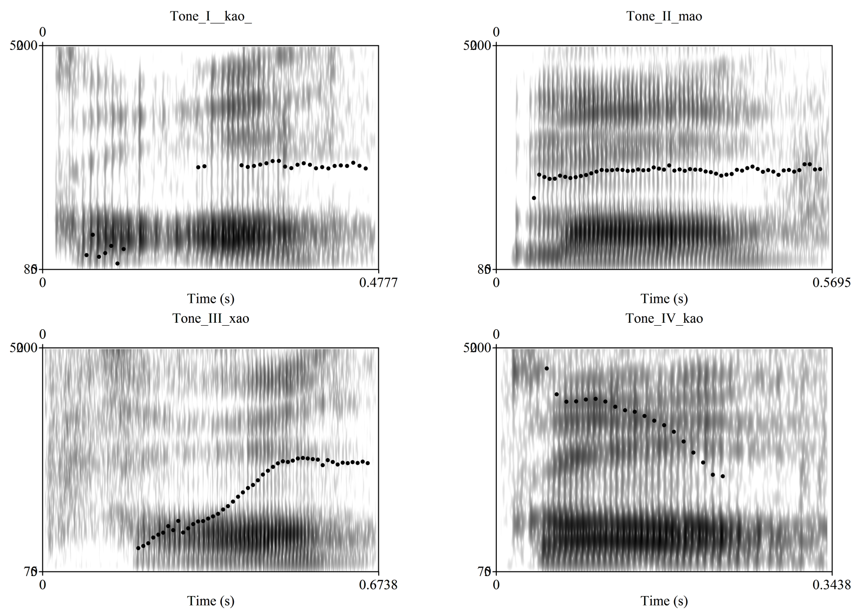


Figure 1: Citation tones

Because there is no real distinction between LM and LH, nor between $L^{C}M$ and $L^{C}H$. I will treat M as a phonetic realization of H when it is followed by a low pitch. In other words, a purely phonological representation of Tone I and Tone III could be $L^{C}H$ and LH. A detailed discussion is provided in Chapter 4.

[2] The tone contour figure was created from four canonical pronunciations of the tones. Data was recorded in Jan 2016, informant being Jieyu Gu, 63 year-old male speaker. The raw data for the whole thesis are collected from him. As a native speaker of the language, my own judgment is also frequently relied on.

1.1 Tone sandhi rules

In this section, I will provide descriptive rules for disyllabic and trisyllabic tone sandhi which occurs at the final foot of a phonological phrase. Although it is possible that people could simply remember all sandhi patterns on a one-by-one basis, two observations call for a structural explanation. First, non-final feet have different sandhi rules. Second, focused elements have different sandhi patterns. An explanation that is applicable to all these phenomena is thus more favorable.

Before introducing descriptive rules, I will provide working definitions to some important notions. The first definition is a standard definition of syllable which can be found in most introductory books to phonology.

Definition 1 (Syllable (Version 1)) Syllable *is an audibly recognizable sequence of sounds with exactly one nucleus, which is usually a vowel. Syllables are building blocks of (phonological) words.*

In Ganyu Mandarin, there are two sorts of syllables based on whether pitch is lexically encoded. Pitched syllables' surface tone realizations are derived from their underlying tonal information, but could be influenced by their environment. 'Pitchless-syllables' surface tone realization is independent from their underlying tonal information and are determined by their phonological environment (Duanmu, 2009). In Ganyu Mandarin, pitchless syllables are reducible to C-[ə], an extra-syllabic consonant, or even zero segmental representation which creates a tone sandhi boundary. In this paper, I will focus on cases where all syllables are pitched.

Definition 2 (Citation form) Citation form *is a syllable's surface pitch realization when it is uttered independently.*

From Definition 2 we could see that citation form is only a possible surface pitch realization. Cross-linguistically, it may or may not coincide with the syllable's underlying pitch information. For languages with very few sandhi phenomena, such as Beijing Mandarin, it is safer to assume that citation forms are mostly underlying forms. As will be shown in Chapter 4, I propose a set of underlying tones which differ from their citation forms for Ganyu Mandarin.

Definition 3 (Tone) Tone *could refer to the underlying* tone category*, which is composed of all syllables which have the same underlying lexical pitch information, or* tone realization*, which is the surface pitch realization at the syllable level.*

In Ganyu Mandarin, underlying tone categories are features of the syllable, surface tone realization is also borne by syllables. In other languages, tone-bearing unit(s) might be other phonological units. For example, in Shanghai Wu, underlying tone categories are features of the syllable, while surface tone realization is borne by phonological words, where only the underlying tone of the first syllable is retained and spreads across the whole word (Zhu, 2006).

Definition 4 (Tone sandhi) *By* tone sandhi *I mean the phenomenon where a syllable is realized in different pitch contours in different environments and such difference is usually related to the tonal information of neighboring syllables.*

If syllables in a specific tone category could be freely realized as different surface forms, regardless of its neighboring syllables, the phenomenon is not considered as tone sandhi. For example, in Cantonese, Tone I (*yīn píng*) could be realized as either HL or HH. But they are free variants (Matthews & Yip, 2013). So I will not call it tone sandhi.

Definition 5 (Tone sandhi domain) Tone sandhi domains *are levels of the prosodic structure at whose edges a sandhi process stops. A language may have more than one sandhi domains.*

The above definition of tone sandhi domain is based on a introductory discussion in Chapter 7 of Chen (2000). In Ganyu Mandarin, there are multiple sandhi domains, the biggest one being the phonological phrase, the smallest one being the foot. In this paper, I will not investigate the delimitation of phonological phrases. But sandhi domains of lower levels will be discussed at length. A detailed discussion of various possible sandhi domains is provided in Chapter 2.

1.1.1 Disyllabic tone sandhi

Table 2 lists all disyllabic tone sandhi processes which occur at the final foot without any focus. Tone sandhi with focus exhibits different patterns, and is also explainable by my theory, but will not be discussed in the paper.

1st syl. \2nd syl	I (L^CM)	II (HH)	III (LM)	IV (HL)
I (L^CM)	HH.L^CM	HH.HH	HH.LM	HH.HL
II (HH)	HL.L^CM	HL.HH	HL.LM	HL.HL
III (LM)	HL.L^CM	HH.HH	HL.LM	HH.HL
IV (HL)	LL.L^CM	LL.HH	LL.LM	LL.HL

Cell colors group cases where the 1st syllables are of the same tone after sandhi.

Table 2: Disyllabic tone sandhi

Observation 1 (Disyllabic tone sandhi) *In non-focused final-foot disyllabic tone sandhi, the second syllable has the same realization with its citation tone. For the first syllable, the following realizations have been observed:*

a. Tone I is uniformly realized as HH.

b. Tone II is uniformly realized as HL.

c. Tone III is realized as HL before Tone I and Tone III; it is realized as HH before Tone II and Tone IV.

d. Tone IV is uniformly realized as LL.

e. In the first syllable, there are three possible tone realizations: HH, LL and HL.

1.1.2 Trisyllabic tone sandhi

Table 3 lists all trisyllabic tone sandhi processes which occur in the final foot of pattern $[[\sigma\sigma]_{Min}\sigma]_{Maj}$ and without any focus[3]. This is the default pattern for trisyllabic words which do not have internal morphological boundary, such as trisyllabic loan words 'ɤ-lwɤ-sə, HL-HL-L^CH' ('Russia', II-II-I tones), although it is also possible to parse it as $[\sigma[\sigma\sigma]]$, which would be pronounced as 'ɤ-lwɤ-sə, HH-HL-L^CH'.

Observation 2 (Trisyllabic tone sandhi) *In non-focused final-foot trisyllabic tone sandhi of pattern $[[\sigma\sigma]\sigma]$, the following realizations have been observed:*

a. Except when the last two tones are II-II or II-IV, the last two syllables follow disyllabic sandhi rules.

b. If the first syllable is I, it changes uniformly into HH in all cases.

c. If the first syllable is IV, it changes into LL in all cases.

[3] The foot pattern is treated as the internal structure of a foot, as discussed under Definition 7 of Section 2.2

1st syl.	2nd syl. \3rd syl	I (L^{C}M)	II (HH)	III (LM)	IV (HL)
I (L^{C}M)	I (L^{C}M)	HH.HH.L^{C}M	HH.HH.HH	HH.HH.LM	HH.HH.HL
	II (HH)	HH.HL.L^{C}M	HH.HH.HH	HH.HL.LM	HH.HH.HL
	III (LM)	HH.HL.L^{C}M	HH.HH.HH	HH.HL.LM	HH.HH.HL
	IV (HL)	HH.LL.L^{C}M	HH.LL.HH	HH.LL.LM	HH.LL.HL
II (HH)	I (L^{C}M)	HL.HH.L^{C}M	HL.HH.HH	HL.HH.LM	HL.HH.HL
	II (HH)	HL.HL.L^{C}M	HL.HH.HH	HL.HL.LM	HL.HH.HL
	III (LM)	HL.HL.L^{C}M	HL.HH.HH	HL.HL.LM	HL.HH.HL
	IV (HL)	HL.LL.L^{C}M	HL.LL.HH	HL.LL.LM	HL.LL.HL
III (LM)	I (L^{C}M)	HL.HH.L^{C}M	HL.HH.HH	HL.HH.LM	HL.HH.HL
	II (HH)	HH.HL.L^{C}M	HH.HH.HH	HH.HL.LM	HH.HH.HL
	III (LM)	HH.HL.L^{C}M	HL.HH.HH	HH.HL.LM	HL.HH.HL
	IV (HL)	HH.LL.L^{C}M	HH.LL.HH	HH.LL.LM	HH.LL.HL
IV (HL)	I (L^{C}M)	LL.HH.L^{C}M	LL.HH.HH	LL.HH.LM	LL.HH.HL
	II (HH)	LL.HL.L^{C}M	LL.HH.HH	LL.HL.HH	LL.HH.HL
	III (LM)	LL.HL.L^{C}M	LL.HH.HH	LL.HL.LH	LL.HH.HL
	IV (HL)	LL.LL.L^{C}M	LL.LL.HH	LL.LL.LM	LL.LL.HL

Yellow cell color indicates cases where the last two syllables are inconsistent with disyllabic sandhi patterns.

Table 3: Trisyllabic tone sandhi of pattern $[[\sigma\sigma]\sigma]$

d. If the first syllable is II, it changes into HL in all cases.

e. Except in T*-II-II and* T*-II-IV, if the first syllable is III, it changes into HL before HH, and HH before HL and LL. Here* T *stands for any possible tone category.*

f. When the last two tones are II-II or II-IV, tone II as the middle syllable is realized as HH.

Table 2 and Table 3 are only a small part of the whole tone sandhi inventory. In sum, the above data call for a unified, systematic and structural explanation to all tone sandhi phenomena in Ganyu Mandarin.

1.2 Previous literature

Tone sandhi in Ganyu Mandarin is among the most complex ones in Mandarin dialects. Many neighboring Mandarin dialects exhibit no or a small number of sandhi phenomena (Li, 2004). However, Xiamen Southern Min, arguably one of the most remote cousins of Ganyu Mandarin, has tone sandhi patterns that are more similar to that of Ganyu Mandarin (Zwicky, Kaisse, & Chen, 1987; Chen, 2000). Standard (Beijing) Mandarin has relatively fewer tone sandhi phenomena, but prosodic structure plays a key role in Tone III sandhi of Standard Mandarin. Fortunately, tone sandhi in Southern Min and Standard Mandarin is relatively well-studied.

1.2.1 Ganyu Mandarin

Previous studies on tone sandhi in Ganyu Mandarin include Iwata (1989), (Liú, 2001) and (Sū & Wàn, 2011). These studies were descriptive in nature and described Qingkou vernacular, whose sandhi patterns are slightly different from Menhe vernacular, the sub-dialect that is investigated in this paper. Among these studies, Iwata (1989) is the most detailed one and studied several idiolects. However, it seems that he did not distinguish focused tone sandhi from non-focused ones. For example, Iwata noted that Tone I could be realized as 24 (LH) or 55 (HH) as the first syllable in a disyllabic word. The fact is that 24 (LH) is focused sandhi, while 55 (HH) is non-focused sandhi. This problem might be attributed to the method of elicitation: speakers are asked to read disyllabic or trisyllabic words. It is not uncommon for speakers to add focus stress in such a scenario.

Being descriptive in nature, none of the above studies has provided a description of the overall sandhi pattern of a phrase which contains more than three syllables. Their description only covers final-foot tone sandhi patterns. As a comparison, my theory could successfully predict sandhi realizations of phonological phrases of arbitrary lengths.

1.2.2 Other languages

Tone sandhi in Xiamen Southern Min is very similar to that in Ganyu Mandarin. The largest sandhi domain in both languages is phonological phrase (Zwicky, Kaisse, & Chen, 1987; Chen, 2000). And in both languages, within a phonological phrase, the last syllable's tone realization is identical to its citation realization, while tones of proceeding syllables are realized differently. Tsay and Myers (1996) and Hsieh (2005) and of the few attempts at explaining Xiamen tonal chain-shifts.

Chen (2000), among others, identified the sandhi domain of Standard (Beijing) Mandarin as the Minimal Rhythmic Unit, which is closely related to foot. Being a Mandarin dialect, foot also plays a crucial rule in Ganyu Mandarin. Tone sandhi in Beijing is also affected by prosodic categories higher than foot. But higher prosodic categories never stop tone sandhi, i.e. the initial syllable of a following phonological phrase could affect the previous syllable which is in another phonological phrase. While in Ganyu, phonological phrase boundaries stop sandhi process, i.e. a syllable in one phonological phrase would never affect a syllable from another phonological phrase (C. Shih, 1997).

1.3 Research questions and theoretical framework

As defined in Definition 4, tone sandhi is the modification of input/underlying tones according to their specific environments. Therefore, a successful explanation to Ganyu tone sandhi should address the following three issues related to tone sandhi:

(i) *Input*: The proposed underlying tone forms should be able to explain tonal realization in various environments.
(ii) *Environment*: The proposed hierarchical prosodic structure, at least part of the whole structure, should be useful in explaining tone sandhi.
(iii) *Modification*: Constraints for both prosodic arrangements and specific sandhi processes should have clear phonetic or structural motivations. When ranked in certain order, these constraints should be able to predict tone sandhi in all environments that the thesis is dealing with, i.e. cases where all syllables are pitched.

For the first issue, underlying tones are not directly observable. They might not occur superficially on any syllable. The identification of them requires us to categorize syllables according to their status at various levels in the prosodic structure. Only by combining observations of different surface forms could I find out the ultimate underlying forms of tones. In my analysis, the underlying tones of the four tonal categories are LLH, HHL, LHØ, HLØ, respectively, with Ø being an underspecified pitch target. This treatment can account for tone sandhi at various environments.

For the second and third issues, rules and constraints are sensitive to prosodic categories. The proposed prosodic structure should be based on observations of the various sandhi patterns.

The complexity of Ganyu tone sandhi makes rule-based explanations more complicated than a constraint-based explanation. In addition, in the literature of theoretical analyses of tone sandhi in Chinese languages, Optimality Theory is often adopted (cf. Chen 2000 for a comprehensive list of references in the treatment of tone sandhi phenomena in various

Chinese languages). In the following sections, I will examine the prosodic structure, underlying tones, and specific tone sandhi phenomena. In my analyses, I will adopt the classical Optimality Theoretic (OT) framework as proposed by A. Prince and Smolensky (1993).

Three major categories of OT constraints have been identified in my analysis. They are tonal saliency, metrical integrity and pitch realization. These constraints have phonetic and/or structural motivations. Different from many theoretical analyses of Chinese tone sandhi, such as Hsieh (2005) for Xiamen, which exhibits similar phenomena with Ganyu tone sandhi, my analysis does not use contrast-preserving constraints, nor does it assume Harmonic Serialism (McCarthy, 2010).

In analyzing pitch realization, a new notion of dependent pitch target is proposed, which can account for the short period of creaky sound as attested in Tone I's citation form in Ganyu, as well as falsetto pitch in Lianyungang.

The remaining part of the thesis is structured as follows: Chapter 2 reviews the notion of prosodic structure and its application in Ganyu phonology. Chapter 3 provides a short OT analysis of phonological phrase delimitation and foot construction. Phonological phrase, foot, and syllable are the three most important prosodic units in Ganyu tone sandhi. Chapter 4 deals with underlying tones. In Chapter 5, tone sandhi of various combinations and in different environments are analyzed with the help of pitch-sensitive OT constraints. Chapter 6 briefly concludes the thesis.

2 Prosodic structure

From Tables 2 and 3 we can see that tone sandhi is sensitive to the phonological status of the syllable. The possible output of tones in disyllabic and tri-syllabic sandhi is shown in (2). A full description of non-final foot's tone sandhi is provided in Section 5.2.4.

(2) Possible tonal realizations:

a. Final foot's final syllable: L^CH, HH, LH, HL

b. Non-final syllable: HH, HL, LL

c. Non-final foot's final syllable: HH LL

Such a difference in phonological status among syllables might be attributed to the organization of prosodic structure. There is no direct or exact evidence for different levels of prosodic structure. The identification of different prosodic layers would require different cues, and tone sandhi is one of them.

Therefore, in this section, I will propose a prosodic structure which is suitable to explain Ganyu tone sandhi. The proposal is based on previous (cross-linguistic) studies (E. Selkirk, 1986; Nespor & Vogel, 2012; H. Zhang, 2016; Zwicky, Kaisse, & Chen, 1987; Chen, 2000) as well as sandhi patterns observed in Ganyu.

2.1 Literature review

The outcome of tone sandhi is a specific linear arrangement of tones which is different from the linear combination of the underlying tone contours of syllables. Such an arrangement is operated with regard to a hierarchical prosodic structure, which also provides domains for other suprasegmental or segmental sandhi phenomena, such as liaison in French and flapping of /t/ and /d/ in American English (Nespor & Vogel, 2012).

The study of prosodic structure lies at the interface between syntax of phonology. In many mainstream syntactic theories, such as the Government and Binding Theory (GB) (Chomsky, 1986; Haegeman, 1994) and the Minimalist Program (MP) (Chomsky, 1998), researchers assume a Phonetic Form (PF) as the interface between the surface syntactic realization (S-form in GB, Spell-out in MP) and the phonetic implementation. In the study of prosodic structure, it is argued that there might be several intermediate structures between the surface syntactic realization and the final phonetic implementation. Prosodic structure lies in one of these intermediate levels. It is hierarchical and hosts syntax-sensitive phonological rules, i.e. phono-syntactic or phono-morphological rules (E. Selkirk, 1986).

Important works in prosodic structure include Nespor and Vogel (2012) (first published in 1982), E. Selkirk (1986) and H. Zhang (2016).

Based on Intalian *raddoppianmento sintattico* (RS), Napoli and Nespor (1979) found a closer relationship between syntax and phonology than Chomsky and Halle (1968) had predicted. According to Chomsky and Halle (1968), the phonological rules were blind to syntactic structure; at most they could make use of syntactic boundaries. RS, on the contrary, depends on the distinction between right and left branches in a syntactic tree. M. Y. Liberman (1975); M. Liberman and Prince (1977) pointed out in their metrical theory that English word and phrase stress are not primarily related to segmental properties, but related to a "hierarchical rhythmic structuring that organizes the syllables, words, and syntactic phrases of a sentence".

Based on previous studies and empirical evidence from phonological rules in several languages, Nespor and Vogel (2012) (first published in 1982) proposed that prosodic structure is responsible for those language-specific phonological rules, as shown in (3).

(3) $\boxed{\begin{array}{l}\text{syntax}\\ \hline \text{segments input}\end{array}} \xrightarrow{\text{prosodic structure}} \boxed{\text{segments output}}$

"Prosodic structuring" could also be thought of as a function which maps syntactically structured phonological segments into the phonological output, full-fledged with stress and sandhi realizations at all prosodic levels.

It is worth noting that Nespor and Vogel (2012) acknowledged that their formulation of phonological phrase and intonation phrase, two prosodic levels, could not be applied to Chinese dialects, as had been discussed in E. Selkirk and Shen (1990).

E. Selkirk (1986) abstracted prosodic levels into a single notion of prosodic structure, which is the main component of P-structure. The other components in P-structures are syntax-sensitive phonological rules, i.e. phonosyntactic rules. Selkirk's formulation of the whole syntax-phonology-phonetic implementation process is shown in the following figure:

In the above figure, the 'phonosyntactic' subcomponent contains syntax-sensitive phonological rules, while the 'phonological' subcomponent contains only pure phonological rules.

Subcomponents/syndromes	*Levels of representation*
	surface syntactic structure
‘phonosyntactic’	
	P-structure
‘phonological’	
	PI-structure
‘phonetic implementation’	
	phonetic representation

Selkirk recognized syllable, foot, prosodic word, phonological phrase, intonation phrase, and utterance as prosodic categories but later questioned the statuses of syllable, foot, intonation phrase, and utterance. The reason is that syllable and foot seem to play roles in other levels of linguistic representation as well. For example, syllabification could possibly be present in the surface syntactic structure and phonetic representation. Re-syllabification might occur after the implementation of all sorts of rules.

Therefore, Selkirk proposed that prosodic word, and phonological phrase are the most essential parts of the P-structure. Based on Chen’s earlier manuscript of Zwicky, Kaisse, and Chen (1987), she further maintained that these categories could be defined in terms of *ends* of specific syntactic constituents.

H. Zhang (2016) argued that the prosodic structure is conditioned by morphosyntax. But it could be divided into three levels based on their driving forces, as shown in (2).

(4) **Focus-discourse based hierarchy:** Utterance, Intonation Phrase

Morphosyntactic based hierarchy: Phonological Phrase, Clitic Group, Phonological Word

Phonology based hierarchy: Foot, Syllable, Mora

More specifically, phonological rules at utterance and intonation phrase levels are used to demonstrate discourse-level information such as focus and topic; phonological rules at the phonological phrase, clitic group and phonological word levels are used to demonstrate morphosyntactic grouping; while phonological rules at lower levels such as foot, syllable and mora are purely formal and used simply to re-organize the input for ease of phonological processing.

Zhang’s classification of the prosodic structure proves useful in my analysis of Ganyu tone sandhi. Because the paper will limit itself in the formulation of tone sandhi rules, I will not venture to investigate the phonosyntactic interface, i.e. I will not analyze how prosodic categories/levels, which are important to Ganyu tone sandhi, are conditioned by morphosyntax. Instead, I will only try to provide a systematic description of various changes to and realizations of the underlying tone information carried by each syllable as well as the interaction between relevant prosodic categories.

2.2 Prosodic structure for Ganyu tone sandhi

Possible tonal realizations in Ganyu tone sandhi, as shown in (2), are reproduced here as (5).

(5) Possible tonal realizations:

a. Final foot's final syllable: L^CH, HH, LH, HL

b. Non-final syllable: HH, HL, LL

c. Non-final foot's final syllable: HH LL

From (5b) we can infer that tone sandhi in Ganyu would be sensitive to neighboring syllables. For example, III-II sandhi's result is HH-HH; while III-III sandhi's result is HL-LH. From the comparison between (5a) and (5b), or between (5b) and (5c) we can infer that a syllable's status with regard to the container foot also plays a role in Ganyu tone sandhi. And from the comparison between (5a) and (5c), it is evident that the prosodic status of the foot is also important in determining the final output of tones.

From the above observations, I have identified three levels in the prosodic structure which are informative in Ganyu tone sandhi: phonological phrase, foot, and syllable. Now I will provide a language-specific working definition to each of them[4], together with prosodic word, which does not play an important role in pitched syllable tone sandhi, but are of essential importance in pitchless syllable tone sandhi.

Definition 6 (Phonological phrase) Phonological phrase *is a prosodic category which corresponds to syntactic phrases, i.e. non-terminal syntactic nodes. But a same syntactic*

[4] Although these notions are used cross-linguistically, the identification and behaviors of them show considerable variations in different languages.

structure might be realized as different arrangements of phonological phrases. (Based on H. Zhang (2016))

In a "NP-VP" sentence, if the NP is not a clitic, the sentence would contain two primary phonological phrases, as shown in (1). However, if the NP is a clitic such as personal pronoun, the sentence usually contains only one primary phonological phrase[5].

Phonological phrase is the biggest tone sandhi domain in Ganyu Mandarin. Within a phonological phrase, the last foot is the most salient foot and shows different sandhi patterns from previous feet. Within a phonological phrase, only the last pitched syllable is realized as its citation tone, unless non-citational tone sandhi is blocked by pitchless syllables within the phrase. Saliency of a prosodic category's sub-categories is derived from directionality in aligning these sub-categories, which will be discussed in Section 5.1.1 from an Optimality Theoretic perspective.

Definition 7 (Foot) Foot *is a binary arrangement of syllables. In Ganyu, foot is composed of exactly two or three syllables. Additionally, there are degenerate feet which contain only one syllable. The identification of foot should take syntactic arrangements of words into consideration. (Based on Feng (2009))*

Major foot *is the outer level of a tri-syllabic foot which is composed of a minor foot and a single syllable, and* minor foot *is the internal level of a tri-syllabic foot which is composed of two single syllables.*

A tri-syllabic foot could contain two foot levels: the minor foot and the major foot. Minor foot is the primary binary combination of syllables. Major foot is also binary, but one of its component could be a minor foot. Therefore, at the surface level there could be tri-syllabic feet, which is actually underlyingly binary but the binarity exhibits at two levels. The distinction between $[[\sigma\sigma]_{Min}\sigma]_{Maj}$ and $[\sigma[\sigma\sigma]_{Min}]_{Maj}$ will prove useful in the analysis of trisyllabic tone sandhi, especially when Tone II occurs at minor foot boundaries. Because the final pitch target of Tone II is L, it would often create a rising pitch contour at foot boundaries, which violate *R-BetweenFt, as discussed in Section 5.1.1. In most conditions, the $[[\sigma\sigma]\sigma]$ or $[\sigma[\sigma\sigma]]$ pattern is not treated as a prosodic word which contains

[5] Ganyu predicates are mostly one-place or two-place. There is asymmetry between subject NP and object NP. Identifying phonological phrases is a mostly syntactic task and I remain neutral as to whether phonological phrases are recursive.

a disyllabic foot and a degenerate foot because the middle syllable behaves as other non-foo-final syllables in tone sandhi. But the existence of an internal minor foot boundary is important for *R-BetweenFt.

Another possible treatment is to make both major and minor feet obligatory, as adopted in S.-h. Shih (2017a). Chen (2000) replaced foot with the notion of Minimal Rhythmic Unit to emphasize that a foot could be formed across word boundaries, because in Chinese many words are monosyllabic. I will still use the term foot in this paper because a narrower and language-specific definition is provided here.

H. Zhang (2016) argued that there was no foot in Mandarin Chinese. He said that only in languages with metrical binary contrast could foot exists. However, although Mandarin Chinese does not show strict strong-weak patterns in foot construction; foot does exist as a binary arrangement of syllables, and could be manifested in many tone sandhi processes, such as the famous third-tone sandhi in Standard Mandarin (C. Shih, 1997).

From a broader perspective, feet tend to be binary cross-linguistically. A. Prince and Smolensky (1993) put forward the Ft-bin constraint: feet are binary at some level of analysis (σ,μ). E. O. Selkirk (1980),A. S. Prince (1980) and Martínez-Paricio and Kager (2015) argued for a layered internal structure of feet with one level of recursion. Typologically speaking, there is no absolute boundary between binary and ternary systems. Treating trisyllabic groups as flat ternary feet makes it difficult to capture typological predictions relating to the binary foot inventory (Hayes, 1995). Ganyu is an example where binary and ternary feet coexist.

The idea of prosodic word is more subtle. Zhang claimed that it is the immediate parent of foot in many languages except Standard Mandarin. However, there are many monosyllabic words, especially V-NP pairs in Mandarin. There are two terminal syntactic nodes, but the V-NP pair behaves as a single foot. For words of more than 3 syllables, a phonological word, if it exits, could contain more than one foot. Thus, in Mandarin Chinese, I could probably re-defined phonological word as follows:

Definition 8 (Prosodic word) Prosodic words *are terminal non-clitic syntactic elements which correspond to at least one foot. If a syntactic word contains only one syllable, it is often part of a disyllabic or trisyllabic phonological words; in such cases the prosodic words contain exactly one foot.*

Feng (2009) also observed that "In prosodic morphology, the smallest prosodic unit that can be used is 'foot.' Therefore, a prosodic word must contain at least one foot. And if a foot consists of two syllables, then a prosodic word naturally must contain at least two syllables." Since most lexical words in Mandarin Chinese are either monosyllabic or disyllabic, prosodic words are usually identical with feet. As far as I have observed, in Ganyu tone sandhi, prosodic word does not play a role in domain delimitation. Therefore, I will omit the category of prosodic word whenever appropriate in the discussion of Ganyu tone sandhi.

After introducing several levels of representation in the whole phonological structure, such as P-structure and PI-structure, syllable now becomes overloaded[6]. Therefore I have to provide an updated definition to it. However, such a definition only applies to Ganyu Mandarin. Syllable's roles in different languages may differ dramatically.

Definition 9 (Syllable (version 2)) Syllable *is an overloaded term and could mean one of the following notions:*

a. In the lexicon, syllable is a pair of segmental and pitch/tone information $\langle$*seg, pitch*$\rangle$*, such pairing has to obey phonotactic rules*[7]*.*

b. In PI-structure, syllable is a re-organized version of the $\langle$*seg, pitch*$\rangle$ *pair in the lexicon, which obeys all phonological rules and is good for phonetic implementation.*

c. (Version 1) *Phonetically, syllable is an audibly recognizable sequence of sounds with exactly one nucleus, which is usually a vowel. The segmental sounds are paired with tone information.*

The above notions are interrelated and can be the subject matter of a theory of syllabification.

Example (6) is a concrete example of how information provided by the lexicon could be accommodated by the (native) prosodic structure. Boldfaced tones are realizations which

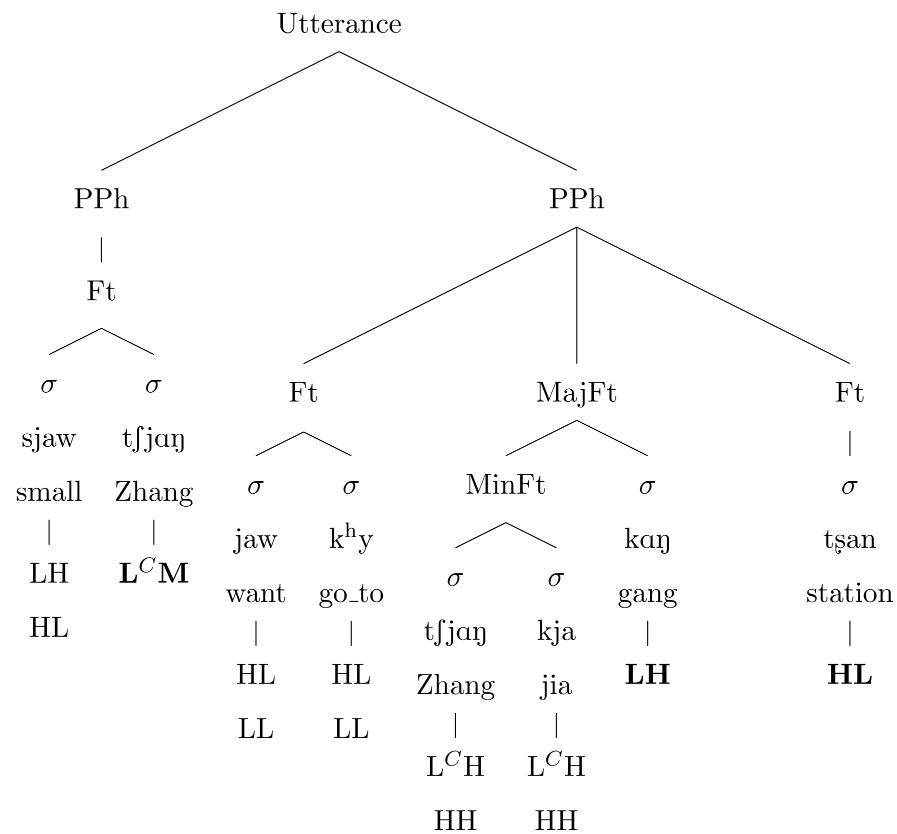
---

[6] Overloading is a term borrowed from theoretical computer science. It refers to situations where a single term could have several (possibly related) meanings in different context.

[7] Molinier (1920), (originally written around 1328-1337) provided a classical yet powerful definition of syllable in Occitan upon which my definition of syllable is based. "Silaba votz es literals. Segon los ditz gramaticals. En un accen pronunciada. Et en un trag: d'una alenada." (A syllable is the sound of several letters, [...] pronounced in one accent, and uninterruptedly in one breath). This old definition emphasized the interplay between syllable structure and tone, which can be found in more recent works such as M. Yip (2002). A more conservative definition would strip syllable out of any suprasegmental information and the notion of syllable used here would correspond to "syllable-tone" pair in that version of definition.

are identical with the citation tones. In (6), I omitted the intermediate categories between foot and phonological phrase because they do not play a role in Ganyu pitched-syllable tone sandhi.

(6)



Translation: 'Xiao Zhang wanted to go to Zhangjiagang Station.'

In (6), The lone syllable [tʂ̩an] forms an independent foot. Such a foot is called a degenerate foot in the literature (Crowhurst & Hewitt, 1995; Poser, 1989; Chen, 2000; H. Zhang, 2016). Such feet are only treated as foot because they are daughters of a phonological phrase in the prosodic structure to ensure that each prosodic category is a total partition of the utterance (E. Selkirk, 1986). Syntactically motivated degenerate feet are allowed and could trigger the lengthening of the previous syllable which would otherwise be shorter in duration. A detailed phonological discussion of degenerate feet is beyond the scope of the paper.

Example (7) shows how foot construction is confined by morphosyntax. Monosyllabic words could only form its own sandhi domain (as a degenerate foot) if it is not able to combine with other words to form a foot.

(7) a. [tʃjɑŋ kja kɑŋ] [tʂ̻an]
L^{C}H L^{C}H LH HL (citation tones)
HH HH **LH** **HL**
'Zhangjiagang Station'

b. [tʃjɑŋ kja kɑŋ] [ʂ̻an-ts]
L^{C}H L^{C}H LH HL-LH (citation tones)
HH HH HH **HL**-∅
'Zhangjiangang fans'[8]

c. *[tʃjɑŋ kja] [kɑŋ ʂ̻an]
L^{C}H L^{C}H LH HL (citation tones)
HH HH HH **HL**
'Zhangjiangang fans' (Not permitted)

In (7b), although the final foot is composed of one syllable and an additional consonant, which is a reduced form of the full syllable [tsɜ], it still maintains its foothood. Consequently, 'Zhangjiagang' could only have the status as a non-final foot, and undergoes non-final foot tone sandhi rules.

Table 4 provides an example of counts of different prosodic categories found in different linguistic segments. The tone of syllables used in foreign proper names is determined by the specific Chinese characters that have been used in transcribing them. The clitic could be realized as either L or H depending on its syntactic or structural status[9].

gloss	surface segments	surface pitch	syl. #	foot #	phon. ph. #
1.sg (clitic)	wo	H or L	1	0	0
1.sg (non-clitic)	wo	LH	1	0	0 or 1
1.sg.pos	wo.t	LH.Ø	2	1	0 or 1
1.sg pos	wo.ti	LH.L	2	1	0 or 1
Wang's village	wɑŋ.tʂ̻wɑŋ	HH.Ø.L^{C}H	3	1	0 or 1
Wang's village	wɑŋ.kj.tʂ̻wɑŋ	HH.Ø.L^{C}H	3	1	0 or 1
Wang's village	wɑŋ.kja.tʂ̻wɑŋ	HL.HH.L^{C}H	3	1	0 or 1
Spain	si.pan.ja	HH.HH.HH	3	1	0 or 1
Alabama	a.la.pa.ma	HH.HH.HH.LH	4	2	0 or 1
Philadelphia	fej.la.tej.ɜ˞.fej.ja	LL.HH.HH.HH.HH.HL	6	3	0 or 1

Table 4: Examples of prosodic category counts

[8] [-ts] is a common nominal suffix. The meanings of [ʂ̻an-ts] and [ʂ̻an] are identical.

[9] Degenerated syllables such as bare initials are still temporarily treated as a syllable in Table 4. A more proper discussion of their prosodic status is beyond the scope of the paper.

3 Prosodic constraints

Ganyu tone sandhi is sensitive to prosodic boundaries. Conversely, tone sandhi even provides many, although not all, evidence for prosodic boundaries at various levels. Therefore, in order to prepare for the discussion of pitch-sensitive constraints in Chapter 5, I am in a position to discuss pitch-neutral constraints which are primarily responsible for how prosodic structures in Ganyu Mandarin are determined.

3.1 Literature review

McCarthy and Prince (1993) put forward the generalized alignment constraints, as shown in (8).

(8) *Generalized alignment*

Align (Cat1, Edge1; Cat2, Edge2) $=_{\text{def}}$ $\forall$ Cat1 $\exists$ Cat2 such that Edge1 of Cat1 and Edge2 of Cat2 coincide, where

Cat1, Cat2 $\in$ Prosodic category $\cup$ Grammatical category

Edge1, Edge2 $\in$ {Right, Left}

In addition, E. Selkirk (1995) proposed the Strict Layer Hypothesis for prosodic structure grouping. There are four general constraints entailed by the Strict Layer Hypothesis , as shown in (6).

(9) *Strict layer hypothesis*

a. Layeredness: no C^i dominates a C^j, iff $j > i$.

b. Headedness: any C^i must dominate a C^{i-1}

c. Exhaustivity: no C^i dominates C^j, iff $j < i - 1$

d. non-recursivity: no C^i dominates C^j, iff $j = i$

(C^n = some prosodic category)

Observing both the strict layer hypothesis and the General Alignment constraints at the same time would inevitably creates conflict. A well-studied phenomenon is the conflict between syllable parsing (by feet) and the binary requirement of syllables. There are two possible results: we could either posit the existence of unparsed syllables or violate the binarity of feet to some extent. In the latter case, any trisyllabic foot (and even for disyllabic feet) are treated as having two layers: the major foot and the minor foot (S.-h. Shih, 2017b). In the treatment of single-syllables, Green (1995) made a further distinction between major syllable and minor syllable. In this paper, I will not discuss pitchless syllables. In my analysis, trisyllabic foot where all syllables are pitched are treated as observing FOOT-MAX(3σ). The insights provided by major/minor feet distinction and the Binarity Principle for feet (McCarthy & Prince, 1993) are captured by two general constraints: FOOT-MAX(3σ) and FOOT-MIN(2σ).

3.2 My proposal for Ganyu prosodic constraints

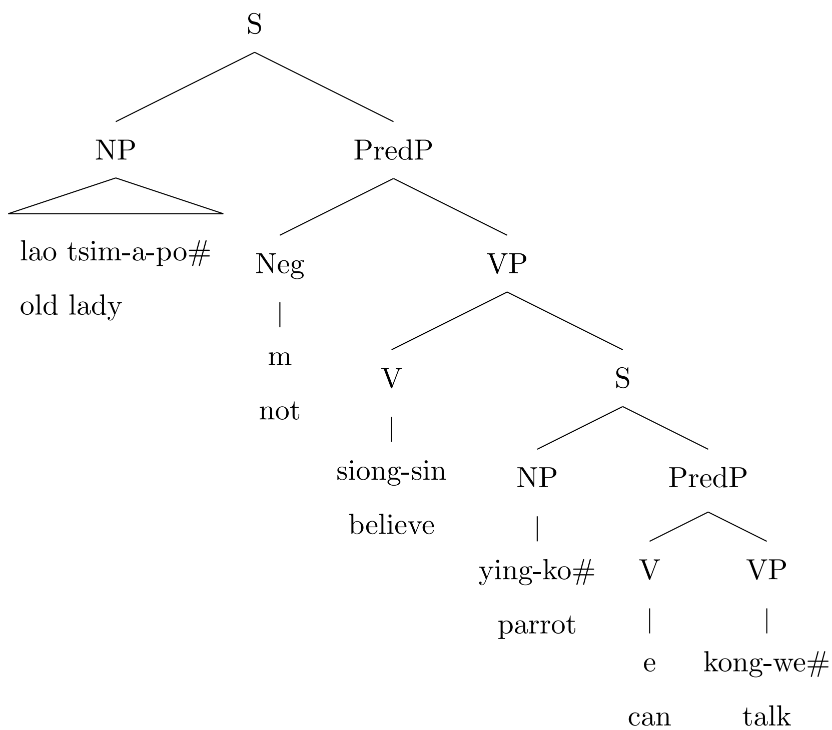
3.2.1 Phonological phrase delimitation

Chen (2000) formulated the derivation of phonological phrase from the perspective of Government and Binding Theory (Chomsky, 1986; Haegeman, 1994). In term of E. Selkirk (1986)'s end-setting typology, Chen made a first approximation of Xiamen phonological phrase, as shown in (10), which says that Xiamen phonological phrase boundaries are identified with the right edge of a maximal phrasal projection. X stands for N(noun), V(erb), etc.

(10) Xiamen p-phrase: {Right, X^{max}} (preliminary version)

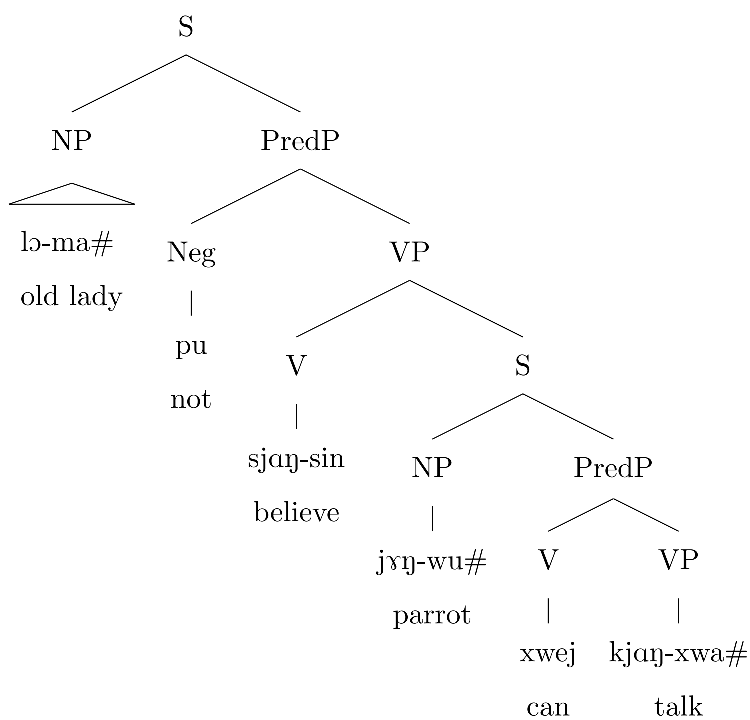
Chen provided an example which could be well translated into Ganyu and the translation follows exactly the same phonological phrase demarcation rule, as shown in (11).

(11) Xiamen (taken from Chen (2000), and adding a # at phonological phrase ends):



"the old lady doesn't believe that parrots can talk"

Ganyu:



Chen mentioned that the fact that tone sandhi applies across *siong-sin* "to believe" and *ying-ko* "parrot" demonstrates that Xiamen p-phrasing is blind to the left-edge of X^{max}.

Lin (1994) revised (Chen, 2000) (originally put forward in 1987) as:

(12) Xiamen p-phrase: {Right, X^{max}}, X^{max} not lexically governed.

Note that this is exactly what Zwicky, Kaisse, Hale, and Selkirk (1987) proposed for Papago tonal phrasing parameter:

(13) $]_X^{max}$, X^{max} not lexically governed.

Lex^{max} refers to the maximal projection of a lexical category, while X^{max} refers to the phrasal maximal projection (Kornai & Pullum, 1990). In this respect, Ganyu Mandarin clearly resembles Xiamen Southern Min.

Putting aside difficult syntactic problems, I use a fairly universal constraint to align phonological phrase with syntactic phrase in the explanation of Ganyu tone sandhi.

(14) ALIGN(PHP, R; X^{MAX}, R)

In other words, the right edge of Ganyu phonological phrase is identified with the right edge of a phrasal maximal projection in the syntax.

3.2.2 Foot construction

Now I will try to parse Ganyu syllables into feet. First, I have identified four alignment constraints, which are responsible for the alignment of prosodic word and foot. The mutual dependence of syntactic and prosodic categories in foot construction has been captured in these constraints.

(15) Phonosyntactic constraints:

a. ALIGN(PRW,R; FT,R): the right edge of a prosodic word must also be the right edge of a foot.

b. ALIGN(FT, R, LEX, R): the right edge of foot must also be the right edge of a syntactic word.

c. ALIGN(FT, L, LEX, L): the left edge of foot must also be the left edge of a syntactic word.

d. ALIGN(W, R; PRW; R): align a syntactic word with a prosodic word.

Constraint (15a) aligns feet with prosodic words. Constraint (15b-c) aligns syntactic words with feet boundaries. There is a two-way alignment in that both edges of a foot should also be edges of syntactic words. Constraint (d) aligns the right boundaries of syntactic words and those of prosodic words.

In (15), we dealt with constraints on the interaction of syntax and phonology. Now I will show some purely phonological constraints. In 2.2, it is argued that in Ganyu a non-degenerate foot contains at most three syllables and at least two syllables. These requirements are also important in foot construction. They are realized as constraints (16e-f).

In Definition 8, I argued that there is at least one foot in a prosodic word. Together with the constraint that syllables should be parsed into feet and we should not add new syllables to rescue foot formation (Chen, 2000; E. Selkirk, 1995), the collection of purely phonological constraints in foot is shown in (16).

(16) Purely phonological constraints:

e. Ft-Min(2σ): there are at least two syllables in a foot.

f. Ft-Max(3σ): there are at most three syllables in a foot.

g. Parse(σ, Ft): parse syllables into feet from left to right.

h. Dep(σ): do not add new syllables.

i. PrW-Min(1Ft): there is at least one foot in a given prosodic word.

Based on empirical data from Ganyu Mandarin, the above constraints are ranked as follows:

Ranking:

PrW-Min(1Ft), Parse(σ, Ft), Ft-Min(2σ), Ft-Max(3σ) $\gg$, Align(PrW,R; Ft,R), Align(Ft, L, Lex, L), Align(Ft, R, Lex, R) $\gg$Dep(σ) $\gg$ Align(W, R; PrW; R)

Tables 5 and 6 are concrete examples of how syllables can be parsed into feet [10].

[10] It is also possible to parse the phrase into two prosodic words: $[_{Pw}[_{Ft}\sigma\sigma]][_{Pw}[_{Ft}\sigma]]$. This realization is used only in emphasis or cases where the speakers are unfamiliar with the proper name. It involves the discussion of degenerate feet, which has been avoided in the paper. The author has a tentative solution, i.e. 'ghost syllable', to such phenomena. Readers who are interested are welcome to contact me for a detailed explanation.

$\sigma\sigma$ [kja.tʃjɤw]	PrW-Min	AlignR(Ft,Lex)	AlignL(F,Lex)	AlignR(W, PrW)
a. $[_{Pw}[_{Ft}\sigma]][_{Pw}[_{Ft}\sigma]]$	*!			
☞ b. $[_{Pw}[_{Ft}\sigma\sigma]]$				*!

Table 5: ‘California’, lit. ‘Ca state’.

Constraints in (15) and (16) that are not reflected in Table 5 are either not violated or uniformly violated. Table 5 shows that although (a) and (b) do not violate foot-syntactic word alignment, the requirement of syntactic word-prosodic word alignment makes (b) more acceptable.

Based on Feng (2009), prosodic word was defined in Definition 8. Consequently, even treating [kja] and [tʃjɤw] as two separate degenerate feet cannot rescue Candidate (a) in Tableaux 5 from being not optional.

$\sigma\sigma\sigma$ [jɤw.t^ha.tʃjɤw]	Ft-Max	Ft-Min	AlignL(Ft,Lex)	AlignR(W, PrW)
☞ a. $[_{Pw}[_{Ft}\ \sigma\sigma\sigma]]$				*
b. $[_{Pw}[_{Ft}\sigma\sigma][_{Ft}\sigma]]$		*!	*	
c. $[_{Pw}[_{Ft}\sigma][_{Ft}\sigma\sigma]]$		*!	*	*

Table 6: ‘Utah’, lit. ‘Utah state’.

Similarly, we can also parse feet into phonological phrases. Within a phonological phrase, the final foot has a specific status, where the last pitched syllable can and must be realized as the same form of its citation tone.

In addition, the parsing of a major foot into a minor foot and a syllable requires two additional constraints: AlignR(MinorFt;Morpheme) and Parse(MinorFt,Ft; Left-to-Right). The first constraint penalizes boundary mismatch between morphemes and minor feet. The latter constraint penalizes right-to-left minor foot parsing. If a trisyllabic foot does not have any internal morphological structure, then it is by default parsed as $[[\sigma\sigma]\sigma]$ as Parse(MinorFt,Ft; Left-to-Right) would penalize the other possibility. However, whenever there is internal morphological boundary, the internal morphological boundary should serve as the boundary of minor foot. Therefore, AlignR(MinorFt;Morpheme) ranks higher than Parse(MinorFt,Ft; Left-to-Right).

In sum, in this chapter, I discussed constraints that are responsible for phonological phrase boundary and foot boundary delimitation. However, none of the constraints introduced in this chapter has anything to do with a syllable's concrete tone/pitch value. Constraints that are sensitive to concrete pitch values will be introduced in Chapter 5, after an analysis of Ganyu underlying tones in Chapter 4.

4 Underlying tones

The representation of underlying tones in Ganyu should address two questions. First, how many tonal features do we need? Second, how many targets does an underlying tone have?

In Ganyu, there are 7 possible tone shapes, as shown in (17).

(17) a. In surface realization of pitchless syllables: L, H

b. In surface realization of pitched syllables: $L^{C}H$, HH, LH, HL, LL

The second question is how to view (17b): are contour tones better regarded as independent pitch units or linear combinations of different pitch targets? My analysis assumes the latter.

Most Chinese dialects have contour tones. Chao's (1930) five-number system has been widely used in descriptive linguistics, where not only contour tones, but also level tones, are represented by a pair or triple of letters. In Chao's system, 1 signals the lowest pitch, while 5 represents the highest pitch. This practice has provided theoretical linguists with important insights. Duanmu (1990, 1994, 2007), among others, have claimed that a stressed syllable in many Chinese dialects consists of two pitch targets, and contours tones should not be analyzed as a single tonal unit (Duanmu, 1994; Hyman & Valinande, 1985; Woo, 1969). Duanmu's claim is supported by tone sandhi phenomena which have been well attested in Chinese dialects, including Ganyu Mandarin. In addition, M. J. Gordon (2002) and M. Gordon (2007) provided a phonetic basis for contour tone restrictions.

4.1 Number of pitch levels

Despite the agreement among many phonologists that a pitched syllable in Chinese generally consists of two (surface) pitch targets (M. J. Yip, 1980; Duanmu, 2007; H. Zhang, 2016), which is also implicit in Chao (1930), there is a considerable dispute on how to represent the underlying forms of a syllable's tonal information.

M. J. Yip (1980) proposed the Register Theory (M. J. Yip, 2001), which have been modified by Pulleyblank (1986); Duanmu (1990); Chen (2000), among others. The basic idea is that the feature [+/- Upper] divides the pitch range into two registers, and the registers are fine-tuned by the feature [+/- high].

In Ganyu, only one feature [+/-H] (or [-/+L]) is enough in representing underlying tones. Even the creak can be treated as underlyingly low but surfaces as a creak in certain environment. A fuller discussion of creaky voice in Tone I is provided in Section 5.1.4.

4.2 Underspecified target

The Register Theory fails to explain the neutrality of Mid tones in many Chinese dialects. Therefore, M. J. Yip (2001) proposed the Target Theory, which analyzes contours tones as consisting of only one target, and level tones as consisting two targets. Yip provided some phonetic support for her analysis. The notion of **target** in target theory can be reformulated as underlying pitch information. The correspondence between targets and their surface representation can be partially captured by the mapping from lexicon to P-structure. While the remaining change of pitch information can be explained via my Optimality Theoretic account of tone sandhi, as shown in Section 5.

Myers (1998) demonstrated the distinction between H and Ø an absence of tonal target, in some Bantu languages. In Tommo So, the underspecified target Ø plays an important role in tone spreading, which is quite similar to the phenomenon that I have observed from Ganyu (McPherson, 2012). In Ganyu, tone sandhi in many cases can be described as the spreading of tonal information from neighboring syllables, and tonal spreading is also mostly triggered by Ø. For example, III-II and III-III sandhi is realized as HH.HH and HL.LH. The second part of tone III in these environments is identical to the first pitch target of the following syllable. Therefore it is reasonable to propose that there is an Ø at the end of Tone III's underlying form.

It is worth mentioning that in Ganyu, underspecified pitch targets do not have default realizations, which is different from many African languages, where underspecified pitch targets usually have a default realization of L or H (Hyman & Valinande, 1985; Hyman, 2009; M. Yip, 2002). The realization or deletion of underspecified pitch targets in Ganyu is determined by their environments and the interaction of various OT constraints.

4.3 Number of targets

Contour tones are resulted from the concatenation of more than one underlying pitch targets in a syllable (Duanmu, 1994, 2009; Hyman, 2009).

While representing underlying tones, there is a possibility that the underlying form consists of more pitch information than a syllable could realize. For example, in Shanghai Wu, syllables in multisyllabic prosodic words are all mono-moraic and can only accommodate one single pitch target. Consequently, multi-targeted/contour tones of the first syllable are realized as split over the whole prosodic words, with all underlying tonal information of the remaining syllables utterly eliminated. Phonetic support is also available for this possibility. In T. Zhang and Zhu (2012), it is discovered that Tone 1 has two citation realizations: $L^{C}M$, and LLH, with the latter being the more carefully produced one[11]. LLH differs from LH in that there is a stable period of low pitch at the beginning of LLH, which is absent in LH.

Phonetically, a realized underlyingly pitched syllable usually consists of two pitch targets, with the exception of Tone I syllables, even in their citations forms. In Section 5.1.4, I will treat the creak as a dependent target. Therefore, we can maintain that in the surface realization of citation tones, only two targets in a syllable are allowed.

In M. J. Yip (2001)'s formulation of Target Theory, level tones have only one target, and the target is associated with two segments, while contour tones have two targets, with each target associated with one segment. However, based on tone sandhi patterns of Ganyu, I propose that underlyingly each syllable is associated with three pitch targets. Targets can be underspecified. Consequently, the underlying tone and surface tonal realization of a syllable is never able to be the same, because an underlying tone either have three fully specified pitch targets, or contain an underspecified pitch target. The underlying tones of Tones I to IV are respectively: LLH, HHL, LHØ and HLØ, with Ø representing an underspecified pitch target.

4.4 Underlying forms

My analysis of Ganyu tone sandhi takes advantage of the two important possibilities: underlyingly a syllable could contain more than two pitch targets; and one or more pitch targets in the underlying form of a syllable could be underspecified, or **pitch-less**. My proposal of underlying tones in Ganyu proves both adequate and powerful in the explanation of various sandhi patterns, as will be shown in Chapter 5. In this section, I will present

[11] Lianyungang (Haizhou) Mandarin is a remote cousin of Ganyu Mandarin, although the two dialects share a considerably long border (most of it is a river).

some justification for the proposed underlying forms.

4.4.1 Tone I: LLH

There are two different surface realizations of Tone I: HH in non-phonological-phrase-final positions, and $\mathrm{L}^{C}\mathrm{H}$ in phonological-phrase final positions. If there are only two underlying targets, there are two possibilities: LL or HH. LL cannot explain why Tone I is mostly realized as HH. Changing two pitch targets would violate pitch faithfulness constraints and does not have clear phonetic or structural motivation. HH cannot explain why Tone I's phrase-final realization is LLH. Adding two L's and deleting an H is not desirable in any conceivable OT setting.

By treating its underlying form as LLH, Tone I's surface behaviors have clear phonetic motivations. Phonetic evidence shows that in citation and phrase-final forms, there are three distinct periods: L, creak and H. The surface form departs from the underlying form only in that the second period of pitch is realized as a creak. This is reminiscent of tonal downstepping (Hombert, 1974), although the phonological behavior is different. In Section 5.1.4 I will propose the notion of dependent pitch target, which is successful in explaining the change from L to creak. As for non-phrase final positions, the surface form is realized as HH because rising pitch contour is prohibited in non-final positions (c.f. Section 5.1.1's discussion on pitch saliency). In Shahe vernacular, Tone I is realized as LL in non-final positions and $\mathrm{L}^{C}\mathrm{H}$ in final positions. This can be explained by the difference in constraint ranking, and also serves as an evidence that the underlying form of Tone I is LLH.

4.4.2 Tone II: HHL

There are two different surface realizations of Tone I: HL or HH in non-phonological-phrase-final positions, and HH in phonological-phrase final positions. If we assume the underlying form of Tone II is simply HH, we cannot explain why in II-II sandhi the outcome is HL.HH, and in [II-II]-II sandhi the outcome is HL.HH.HH. If we assume the underlying form is simply HL, we cannot explain why in citation form and phrase-final form it is realized as HH. HL is actually allowed in these positions, as shown in the surface realization of Tone IV.

Instead, if we assume the underlying form is HHL, the realizations can be easily derived from OT tableaux. Important constraints include MAX-RIGHTMOSTPITCHTARGET(SD) and *R-BetweenFt. The latter constraint penalizes rising contour at foot boundaries. Because trisyllabic foot can be further analyzed as minor foot, so *R-BetweenFt is relevant in trisyllabic tone sandhi but not in disyllabic tone sandhi. Therefore, II-II is realized as HL.HH, while [II-II]-II is realized as HL.HH.HH.

4.4.3 Tones III and IV: LHØ and HLØ

As argued by McPherson (2012), the underspecified target Ø plays an important role in tone spreading in Tommo So. In Ganyu, tone spreading is also triggered by the existence of Ø. For Tone I and II, whose underlying pitch targets are fully specified, surface forms are mostly derived from deleting one of the targets because of a metrical integrity requirement which penalizes non-binary pitch targets in surface realizations, as will be discussed in Section 5.1.2.

However, as for Tone III's non-final realization, the second pitch target is consistently determined by the first underling pitch target of the following syllable. Therefore, it is useful to assume the existence of Ø which can trigger such tone spreading. In addition, including Ø will also make our prediction of target deletion correct. If we assume that the underlying form is simply LH, then there is no reason why we cannot realize it as HH in all non-final cases, after considering the fact that rising tone is prohibited in non-final positions.

Tone IV's behavior is similar to that of Tone III except that Tone IV is uniformly realized as LL in all non-final positions. This is due to the fact that rising is prohibited in these positions. However, in focus-stressed tone sandhi, if the following syllable is unstressed and underlying starts with a H pitch target, then Tone IV is realizes as $\mathrm{L}^{C}\mathrm{H}$. If the following syllable starts with L, Tone IV is realized as HL. This phenomenon shows that tone IV's realization is also sensitive to the following syllable. Therefore, it is reasonable to assume that there is also an underspecified pitch target in Tone IV.

Directional spreading without assuming the existence of underspecified tone sandhi is a common way of analysis in the literature (J. Zhang, 2007; Chen, 2000). In languages which can be explained by simple directional spreading, tone spreading works for all cases. In Ganyu, tone spreading is only observed in Tone III and Tone IV. If we do not assume Ø, there is no reason to forbid directional tone spreading for Tones I and II. Treating all syllables as having three pitch targets underlyingly observes metrical integrity at various prosodic levels, making the analysis more systematic.

5 Tone sandhi explained

At first sight, Ganyu tone sandhi cannot be simply explained by ease of production. The same is true for Xiamen, where no clear phonetic motivation can be attested effortlessly for tone sandhi (Hsieh, 2005). However, as will be shown later, an Optimality Theoretic explanation can unveil many hidden phonological factors which are responsible for tone sandhi in Ganyu. The superficially complex patterns are actually explainable by a constraint-based phonological reasoning. And relevant OT constraints all have either phonetic or structural motivations.

With the help of analyses of prosodic structure and underlying tones in previous chapters, I am now in a position to provide a simple answer to all sandhi phenomena.

My main task here is to unveil the mechanism of tone sandhi as listed in Tables 2 and 3, which are reproduced here as Table 7 and Table 8, respectively.

1st syl. \2nd syl	**I (L^CM)**	**II (HH)**	**III (LM)**	**IV (HL)**
I (L^CM)	HH.L^CM	HH.HH	HH.LM	HH.HL
II (HH)	HL.L^CM	HL.HH	HL.LM	HL.HL
III (LM)	HL.L^CM	HH.HH	HL.LM	HH.HL
IV (HL)	LL.L^CM	LL.HH	LL.LM	LL.HL

Cell colors group cases where the 1st syllables are of the same tone after sandhi.

Table 7: Disyllabic tone sandhi

5.1 Constraints

In order to explain Tables 7 and 8, we have to look for OT constraints from three major perspectives.

Firstly, there should be constraints which can capture the relationship between syllable status and pitch saliency. This would help us understand why some tonal outputs are only allowed in certain prosodic positions. Secondly, there should be constraints which can capture the integrity of metrical arrangements, because metrical arrangements are of vital importance in parsing linear sound signals into various prosodic levels from the perspective of hearers. Thirdly, there should be constraints which can capture the behavior of pitch targets. These third set of constraints are discussed in three subsections, which treat the realization of pitch, the dependent creak peculiar to Ganyu, and some normal faithfulness constraints for pitch.

1st syl.	2nd syl. \3rd syl	I (L^CM)	II (HH)	III (LM)	IV (HL)
I (L^CM)	I (L^CM)	HH.HH.L^CM	HH.HH.HH	HH.HH.LM	HH.HH.HL
	II (HH)	HH.HL.L^CM	HH.HH.HH	HH.HL.LM	HH.HH.HL
	III (LM)	HH.HL.L^CM	HH.HH.HH	HH.HL.LM	HH.HH.HL
	IV (HL)	HH.LL.L^CM	HH.LL.HH	HH.LL.LM	HH.LL.HL
II (HH)	I (L^CM)	HL.HH.L^CM	HL.HH.HH	HL.HH.LM	HL.HH.HL
	II (HH)	HL.HL.L^CM	HL.HH.HH	HL.HL.LM	HL.HH.HL
	III (LM)	HL.HL.L^CM	HL.HH.HH	HL.HL.LM	HL.HH.HL
	IV (HL)	HL.LL.L^CM	HL.LL.HH	HL.LL.LM	HL.LL.HL
III (LM)	I (L^CM)	HL.HH.L^CM	HL.HH.HH	HL.HH.LM	HL.HH.HL
	II (HH)	HH.HL.L^CM	HH.HH.HH	HH.HL.LM	HH.HH.HL
	III (LM)	HH.HL.L^CM	HL.HH.HH	HH.HL.LM	HL.HH.HL
	IV (HL)	HH.LL.L^CM	HH.LL.HH	HH.LL.LM	HH.LL.HL
IV (HL)	I (L^CM)	LL.HH.L^CM	LL.HH.HH	LL.HH.LM	LL.HH.HL
	II (HH)	LL.HL.L^CM	LL.HH.HH	LL.HL.HH	LL.HH.HL
	III (LM)	LL.HL.L^CM	LL.HH.HH	LL.HL.LH	LL.HH.HL
	IV (HL)	LL.LL.L^CM	LL.LL.HH	LL.LL.LM	LL.LL.HL

Yellow cell color indicates cases where the last two syllables are inconsistent with disyllabic sandhi patterns.

Table 8: Trisyllabic tone sandhi of pattern $[[\sigma\sigma]\sigma]$

5.1.1 Tonal saliency

M. Gordon (2007) argued that there is a close link between sonority and contour tone licensing. In many languages, contour tones are limited to the right edge of a word. In addition, vowels in word-final position are typically longer than word-medial vowels (Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992). Chen (2000) provided a detailed account of tonal saliency. In New Chongming, Tonal Saliency is shown in (18). In Chen's original formulation, level tones are called 'even tones' and contour tones are called 'oblique' tones.

(18) Tonal Saliency (Chen)

a. (At the syllabic level) level tones are more salient than contour tones

b. High register tones are more salient than low register tones.

c. Tone in prominent position is more salient than tone in weak position.

Since there is only one phonologically relevant tone register in my analysis, (18b) is omitted. In Chen's original formulation, tonal saliency includes both the saliency of pitch and the prominence of positions. These two notions are in effect very distinct from each other. Positional prominence provides accommodation for pitche contours of different saliency. Therefore, I will discuss them separately.

Pitch saliency is related both to perceptual and to production. Perceptually, it is the ranking of different pitch contours according to their prominence in a continuous utterance. Perceptual salience can be measured by frequency-following-response (FFR) in listeners' brains. As for production, more elaborate and more effortful pitch contours have greater saliency. Combined together, production and perceptual saliency has an overall ranking and is reflected phonologically. More salient pitch can be accommodated by less prominence syllable positions. Pitch saliency ranking in Ganyu is summarized in (19).

(19) Pitch saliency:

HH, LL > HL > LH

Phonologically, LH only occurs at the final position of a phonological phrase. HL cannot occur at a non-phonological-phrase-final foot end. While HH and LL can occur in syllables of any position within a phonological phrase.

First, Ganyu Mandarin disfavors contour tones. In many languages, contours tones are restricted to long vowels or stressed syllables (McDonough, 1999; Lanham, 1958). J. Zhang (2004) provides phonetic and typological evidence for the restricted distribution of contour tones. Compared with level tones, they require longer production duration.

In addition, Ganyu Mandarin strongly disfavors rising tones. Neuhoff (1998) showed that there is perceptual disparity between rising and falling tones at an experimental setting. Gandour and Harshman (1978); Gandour (1978) provided cross-linguistic evidence for the difference in perceptual saliency among different pitch contours. More recently, Krishnan, Xu, Gandour, and Cariani (2004); Wayland, Zhu, and Kaan (2015) provided experimental evidence which shows that in Mandarin rising tone is more perceptually salient than falling tones. The above-mentioned phonetic evidence is supportive of the phonological constraints in Ganyu that rising tones are restricted to more prominent positions as well as focused syllables.

In Ganyu, Rising tone is prohibited in non-final or non-focus-stressed positions, while falling tone is still acceptable, although not preferred in many cases, in non-final positions. Therefore, rising pitch contour is not only disfavored syllable-wise, but is generally disfavored at all prosodic categories/levels. Typologically speaking, falling tone is the most frequent tone that is attested in Chinese languages (1125 falling tones in 737 dialects), while level tone has been attested 1086 times and rising tone has only been attested 790 times (Cheng, 1973). The fact that falling tone is even more common than level tone can be explained by the fact that in many Chinese languages, there are more than two levels of pitch targets. A four-tier distinction (two levels multiplied by two registers) can accommodate at most four level tones, but six falling tones.

Positional prominence is the relative ranking of lower prosodic units according to their position in a higher prosodic unit. Here I focus on positional prominence of syllables. It serves as a cue for hearers to parse utterances more easily. Positional prominence of syllables can be realized as phonetic evidence such as phrase-final lengthening (J. Zhang, 2004; Edwards, Beckman, & Fletcher, 1991), or phonological evidence such as difference in the ability to accommodate different pitch contours.

Positional prominence across difference levels of the prosodic structure is usually uniform. For example, Ganyu Mandarin shows right-prominence in many other aspects. In Ganyu, syntactic heads usually occurs at the right side in a constituent. Within a foot, the right-most syllable is longer in duration. Smith (2004) showed that in many languages, the right-edge is the default stress position. A consistent right-edge prominence attribution makes phonological parsing easier. When foot interacts with phonological phrase, the right edge of a phonological phrase is also the right edge of a foot. Positional prominence of syllables can be identified and ranked with the help of two parameters, as summarized in (20).

(20) Positional prominence (of syllables):

[+PPhR, + FootR] > [-PPhR, +FootR] > [-PPhR, -FootR]

The interaction between (19) and (20) gives rise to the following optimality constraints:

(21) Tone saliency:

i. *R(ShortDur): No rising tone in syllables of shorter duration. Syllables of shorter duration are neither phonological-phrase-final nor focused.

ii. *R-BetweenFt: No rising pitch contour at foot boundaries in a phonological phrase.

iii. *Contour: No contour tones in any syllable.

iv. *PhraseFinal(σ,ShortDur): A phrase-final syllable cannot have short duration.

v. *NonPhraseFinal(σ,LongDur): A non-phrase-final syllable cannot have long duration.

The interaction of pitch saliency and positional prominence is important for *R(ShortDur). *R-BetweenFt is a more another constraint against rising pitch, but it works at foot boundaries. It is irrelevant in disyllabic tone sandhi because there is no internal foot boundary within a disyllabic foot. However, it is important for trisyllabic tone sandhi because a trisyllabic syllable is actually composed of a syllable and a minor foot, as discussed in Section 2.2. This constraint is irrelevant when the syllable at the left side of the minor foot boundary is Tone I, because the final pitch target of Tone I is H, and no cross-foot rising contour can occur. Nor is this constraint relevant when the syllable at the left side of the minor foot boundary is Tone III or Tone IV, because their final pitch target is underspecified and realized as identical with the first pitch target of the following syllable. In sum, in trisyllabic sandhi, *R-BetweenFt is particularly important when Tone II occurs at minor foot boundaries.

*Contour is an even more general constraint which penalizes any occurrence of contour tones in a syllable. *R(ShortDur) and *R-BetweenFt rank higher than *Contour.

Constraints *PhraseFinal(σ,ShortDur) and *NonPhraseFinal(σ,LongDur) are responsible for the relative timing of a syllable. They are very high-ranked. Applying them will results in a set of filtered candidates where all phrase-final syllables are long in duration, and all non-phrase-final syllables are short in duration. They interact with *R(ShortDur) to ensure that *R(ShortDur) only penalizes rising tones in non-phrase final positions.

I do not present *R(ShortDur) as *R(NonFinal) because focus-stressed non-final syllables can also be of longer duration. Although I do not discuss focus-stressed syllables in the thesis, they exhibit some similar behaviors with normal phrase-final syllables, and such similarity can be attributed to the fact that they are both relatively long in duration.

5.1.2 Metrical integrity

Within a given language, phonological manipulations on inputs can also be attributed to metrical integrity, or in Chen's (2000) term, melodic integrity, which is a requirement that there should be consistency in structural arrangements in order to make phonological parsing successful and easier.

Classical metrical theory provided us with the insights that there should be metrical integrity at many prosodic levels (Kager, 1995). Usually, a language's phonological system is either right-prominent or left-prominent. Within Chinese languages, Mandarin and Southern Min languages are usually right-prominent; while Shanghai Wu and many other Northern Wu languages are left-prominent (Yue-Hashimoto, 1986; Duanmu, 1993; Chen, 2000). As a Mandarin dialect, Ganyu shows exactly right-prominence at all prosodic levels.

Apart from right-prominence, Ganyu also has some constraints on how many child-level components a parent prosodic level can accommodate. The most important constraint is that in the output a pitched syllable can and must accommodate two pitch targets, as shown in (22v). Other metrical integrity constraints are realized as faithfulness constraints (22i-iv).

(22) Metrical integrity:

i. Max-RightmostPitchTarget(SD): The rightmost pitch target in the input of a shorter-duration syllable cannot be deleted.

ii. *Doub-Pitch(R): An input pitch target cannot be realized as two pitch targets. If a pitch target's pitch is also associated with the following pitch target in the output, this constraint is violated.

iii. *Doub-Pitch(L): An input pitch target cannot be realized as two pitch targets. If a pitch target's pitch is also associated with the preceding pitch target in the output, this constraint is violated.

iv. *Cross-Foot-T: The pitch target of one foot cannot be used in another foot.

v. 2T/σ: In the output a pitched syllable can and must accommodate two pitch targets.

Constraint (22i) Max-RightmostPitchTarget(SD) is a manifestation of the uniform right-prominence across different prosodic levels in Ganyu Mandarin. Constraints (22ii)

*Doub-Pitch(R) and (22iii) *Doub-Pitch(L) are crucial in our analysis. The underlying rationale is that the same piece of information in the input is better not used more than once in the output. The two constraints are faithful to the one-to-one pitch correspondence between the input and the output. The left part of Figure 2 shows a violation of *Doub-Pitch(L); and the right part of Figure 2 shows a violation of *Doub-Pitch(R). And both cases violate Max-Pitch. The motivation for constraint (22iv) *Cross-Foot-T, which is important for non-final foot sandhi, will be discussed in Section 5.2.4.

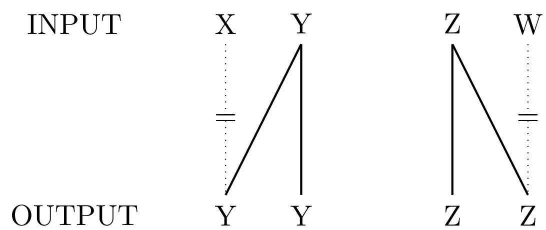


Figure 2: *Doub-Pitch constraints

Apart from these constraints, there are other metrical integrity constraints, which are responsible for foot parsing, as already been discussed in Chapter 3. They are separately treated as 'prosodic constraints' because they do not interact with pitch information.

5.1.3 Mandatory realization and realizability

In Chapter 4, I argued for the fact that it is possible to have underspecified pitch target in the input, which is supported by typological evidence in many other languages (Davison, 1992; West, 2016). When it comes to the output, the underspecified pitch target must either be realized or fully deleted, according to the requirements and ranking of various constraints.

Relevant constraints are shown in (23).

(23) Realization and realizability:

i. *UnspecifiedPitchTarget: In the output, all pitch targets must have a specific pitch value (L or H). Suppose an unspecified pitch targets could be realized as M, M is not allowed in Ganyu phonologically. H could be realized not so high, but is always phonologically treated as H. I assume all candidates have fulfilled the requirement of this constraint.

ii. Max-3Pitch: If the the input there are three consecutive targets of the same pitch, it must be realized at least once in the output.

Constraint (23ii) is an additional language-specific faithfulness constraint which requires that three consecutive targets of the same pitch must be realizes at least once in the output. For example, if we have a sequence of HHL.LLH in the input, the constraint penalizes any candidates where the middle L's are utterly deleted. This constraint is particularly related to sandhi involving LLH because in non-phrase-final cases it is always realized as HH, where two L's have been deleted. In simpler words, this constraint asks the candidates not to delete too many identical pitches, which would otherwise obliterating all cues for the existence of such pitches.

5.1.4 Dependent creak

In the citation tone of Tone I, the surface realization $L^C H$ has maintained the overall tone shape (Low level + Rising), and successfully avoided being too long. This is only possible for Tone I, whose underlying tone has two consecutive L's. In the output, the creak is treated as a dependent pitch to the previous L.

Javkin and Maddieson (1983), Andruski and Ratliff (2000), and Steed et al. (2006) reported that a creak can have phonological status. And they showed that in realization, creaky sounds/tones are usually much shorter than the corresponding modal sounds/tones.

In the literature, similar phenomena have been analyzed as downstepping (Hombert, 1974). Although the phonologically-relevant creak in Ganyu is a manifestation of downstepping, it deserves some additional treatment. Creak is the result of the interaction among several conflicting constraints: 2T/σ, Max-Pitch, *R(ShortDur) and I-Contig-Pitch.

In my analysis of Ganyu tone sandhi, I propose a new concept: dependent association, which is inspired both by Pierrehumbert Janet and Beckman (1988); Prieto, D'imperio, and Fivela (2005) and dependent type theories as advocated by Martin-Löf and Sambin (1984), where a type can depend on the value of another type[12]. Dependent association of a pitch target means that the pitch target is associated with another pitch target, instead of a prosodic unit. A dependent pitch target is not counted by the syllable as a primary

[12] Here is a more formal formulation of different kinds of pitch realization in type theory: primary pitch association is understood as the usual type judgement such as H:σ; secondary pitch association is understood as coercive subtyping (Luo, 1999); and dependent pitch association is understood as dependent type.

pitch target. Therefore, the realization L^CH has two instead of three primary pitch targets and does not violate constraint (22v) $2T/\sigma$.

The difference among simple association, double association and dependent association is shown in Figure 3.

Black line: primary association;
Blue line: dependent association

Figure 3: Different kinds of pitch associations

Creak is a realization of dependent low pitch; while a dependent high pitch can be realized as falsetto, which was attested by T. Zhang and Zhu (2012). T. Zhang and Zhu (2012) analyzed Lianyungang Mandarin, a neighbor dialect of Ganyu Mandarin, as having a three-register tone system, where the dependent falsetto pitch target behaves exactly like creaky voice, with the only difference being its underlying pitch is H. Falsetto is not attested in Ganyu.

In Ganyu, each syllable can only accommodate two pitch targets. Usually speaking, each pitch target can only be realized as a L or H. However, this does not mean that we can have multiple fillers for each target. For example, they can be LL or HH underlyingly. In order to indicate such multiplicity, the mode of pronunciation should be changed. Otherwise, given limited time, it is not possible to distinguish L from LL within one pitch target. Hence the second pitch will naturally become a 'dependent' pitch. In Ganyu, LLH becomes L^CH, where the creak is dependent on the previous L. In Lianyungang, MHH becomes MH^F, where the falsetto pitch is dependent on the previous H. Lianyungang boasts both phonologically important creaky and falsetto pitches T. Zhang and Zhu (2012). Pictures of falsetto pitch realization can be found in T. Zhang and Zhu (2012).

There are two constraints related to the mysterious creak, as shown in (24).

(24)
i. Markedness

*Creak-Independent: In Ganyu tonal phonology, a creak pitch is a dependent pitch target, which must follow an L.

ii. Faithfulness

*H→C: An underlying H cannot surface as creak.

5.1.5 Pitch faithfulness constraints

Based on McCarthy and Prince (1995, 1999) and tone sandhi patterns in Ganyu, I have identified some MAX, IDENT, DEP and CONTIG constraints related to pitch, as shown in (25).

(25) Faithfulness:

i. IDENT-SHAPE(LONGDUR): The contour shape of the output should be identical with that of the input in longer syllables. Longer syllables are either phonological-phrase final or focused.

ii. IDENT-PITCH: The pitch of the output pitch targets should be identical with the input pitch targets.

iii. MAX-PITCH: All underlying pitch targets must be realized in the output. This constraint has very low ranking. It is almost always violated in each and every case.

iv. DEP-PITCH: Do not use a pitch which does not occur in the input.

v. I-CONTIG-PITCH: Contiguity of pitch targets in the input should be preserved in the output. The result is that there should be no internal deletion of a pitch target in a syllable.

IDENT-SHAPE(LONGDUR) acts in a strict way. For example, the shape of LLH is regarded as Level + Rising, instead of simply being Rising. Actually, it is simply another IDENT-PITCH constraint, but at the syllable level. It can be alternatively written as IDENT-PITCH(SYLLABLE). Deleting a pitch target often violates IDENT-SHAPE(LONGDUR), but does not violate IDENT-PITCH, because IDENT constraints only compare materials that are present both in the input and the output. A single pitch target in the input can be simply deleted in the output, while the overall pitch shape of the syllable can never be fully deleted in Ganyu, even pitchless syllables have to be realized either as a short L or a short H based on their environments. On the other hand, IDENT-SHAPE(LONGDUR) collects all pitch information, regardless of whether the pitch target is a primary pitch target or a dependent target. Therefore, $L^{C}H$ retains the shape of LLH although it has been shortened as having only two primary pitch targets. The first pitch target is actually a pair of pitch ⟨L,C⟩, with the creak depending on the L.

I-Contig-Pitch is a stricter requirement than the standard Linearity constraint against metathesis: the precedence relation of input elements must be maintained in the output (McCarthy & Prince, 1995, 1999). In I-Contig-Pitch, the immediate precedence relation of input elements must be maintained in the output (Horwood, 2002; Landman, 2002; Lamontagne, 1996; Fleischhacker, 2005). In other words, when we have three targets in the input, and the output can only make licit two of them, the middle target must be retained . How I-Contig-Pitch works is shown in (26)

(26) Penalized: X~~Y~~Z

Not penalized: X~~YZ~~, XY~~Z~~, ~~X~~YZ, ~~XY~~Z

In sum, constraints in (25) require that a pitch target should not be deleted, appearing from nowhere in the input, or changed in value. In addition, there should be no internal deletion of pitch targets.

5.2 Analysis

Although assuming parallelism, I am presenting my analysis in two steps for ease of discussion. In the first step, I used constraints which are relevant for prosodic domains delimitation, as shown in Section 3.2. In the second step, I listed those candidates that have not violated constraints used in step one and look primarily at the pitch change from the input to the output, as shown in Section 5.1.

Pitch-related constraints, as discussed in Section 5.1, are summarized in (27).

(27) Markedness Constraints:

a. *R(ShortDur): No rising tone in syllables of shorter duration. Syllables of shorter duration are neither phonological-phrase-final nor focused.

b. *R-BetweenFt: No rising pitch contour at foot boundaries in a phonological phrase.

c. *Contour: No contour tones in any syllable.

d. 2T/σ: In the output a pitched syllable can and must accommodate two pitch targets.

e. *UnspecifiedPitchTarget: In the output, all pitch targets must have a specific pitch value (L or H).

f. *PhraseFinal(σ,ShortDur): A phrase-final syllable cannot have short duration. I will assume all candidates have fulfilled the requirement of this constraint.

g. *NonPhraseFinal(σ,LongDur): A non-phrase-final syllable cannot have long duration. Again, I will assume that all candidates have fulfilled the requirement in the following discussion. This constraint and the previous one are based on the phonetic implementation phenomenon 'phrase-final lengthening'. They interact with constraint which are sensitive to the duration of the syllable such as *R(ShortDur). J. Zhang (2004) provides a detailed phonetic explanation for the ability to accommodate contour tones of syllables with different duration.

h *Creak-Independent. In Ganyu tonal phonology, a creak pitch is a dependent pitch target, which must follow an L.

Faithfulness Constraints:

i. Max-RightmostPitchTarget(SD): The rightmost pitch target in the input of a shorter-duration syllable cannot be deleted. Ganyu shows right-prominence at many levels, such as syllable, and phonological phrase. This constraint shows right-prominence of suprasegmental information at the syllable level.

j. Ident-Shape(LongDur): The contour shape of the output should be identical with that of the input in longer syllables. Longer syllables are either phonological-phrase final or focused.

k. Ident-Pitch: The pitch of the output pitch targets should be identical with the input pitch targets.

l. Max-Pitch: All underlying pitch targets must be realized in the output. This constraint has very low ranking. It is almost always violated in each and every case.

m. *Doub-Pitch(R): An input pitch target cannot be realized as two pitch targets. If a pitch target's pitch is also associated with the following pitch target in the output, this constraint is violated.

n. *Doub-Pitch(L): An input pitch target cannot be realized as two pitch targets. If a pitch target's pitch is also associated with the preceding pitch target in the output, this constraint is violated.

o. Dep-Pitch: Do not use a pitch which does not occur in the input.

p. Max-3Pitch: If the the input there are three consecutive targets of the same pitch, it must be realized in the output.

q. I-Contig-Pitch: Contiguity of pitch targets in the input should be preserved in the output.

r. *Cross-Foot-T: The pitch target of one foot cannot be used in another foot.

s. *H→C: An underlying H cannot surface as creak.

The ranking of the above constraints is shown as follows:

(28)

*PhFinal(σ,SD) *Non-PhFinal(σ,LD) I-Contig-Pitch *Creak-Independent *H->C	≫	Id-Shape(LD) 2T/σ	≫	Max-RPitch(SD) Max-3Pitch *Cross-Foot-T	≫	*R(ShortDur)	≫
*R-BetweenFt	≫	Id-Pitch	≫	*Contour Max-Pitch Dep-Pitch	≫	*Doub-Pitch(R)	≫
*Doub-Pitch(L)							

In the following discussion, I will assume that none of the candidates have violated the constraints *PhFinal(σ,SD), *Non-PhFinal(σ,LD) and I-Contig-Pitch because they are high-ranked.

5.2.1 Citation tones

LLH	2T/σ	Ident-Shape(LD)	*Contour
☞ a. $L^C H$			*
b. LL		*!	
c. HH		*!	
d. LH		*!	*
e. LLH	*!		*

Table 9: Tone I

Candidate (b) also violates Max-RightmostPitch, which is not violated by any other candidate. Constraints which are ranked lower than Ident-Pitch and Ident-Time, or do not have effect on a single syllable, are omitted from the above tableau.

The input-output correspondence of the Tone I's citation realization is shown in Figure 4.

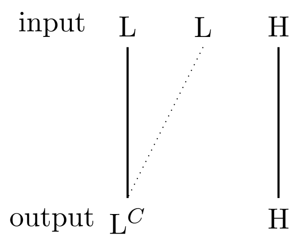


Figure 4: Citational Tone I

HHL	2T/σ	Ident-Shape(LD)	*Contour	*Max-Pitch
a. HHL	*!		*	
☞ b. HH		*		*
c. HL		*	*	*!

Table 10: Tone II

Constraints which are never violated, such as Max-3Pitch, *Cross-Foot-T, *R(SD), two Ident constraints, Dep-Pitch, and two Doub constraints, are not shown in the above tableau.

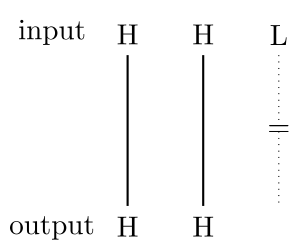


Figure 5: Citational Tone II

LHØ	Id-Shape(LD)	Id-Pitch	*Contour	Max-Pitch
☞ a. LH			*	*
b. HH	*!			
c. LL	*!	*		*

Table 11: Tone III

Constraints which are never violated, such as $2\mu/\sigma$, Max-3Pitch, *Cross-Foot-T, two constraints on Rising tone, two Ident constraints, Dep-Pitch, and two Doub constraints, are not shown in the above tableau. In addition, low-ranked constraints such as Max-RPitch are also omitted.

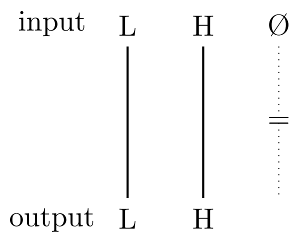


Figure 6: Citational Tone III

HLØ	Ident-Shape(LD)	Id-Pitch	*Contour	Max-Pitch
☞ a. HL			*	*
b. HH	*!			
c. LL	*!	*		*

Table 12: Tone IV

The realization of Tone IV is mostly conditioned by the same set of constraints which have shaped Tone III.

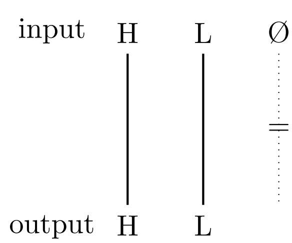


Figure 7: Citational Tone IV

5.2.2 Disyllabic tone sandhi

In this section, I discuss cases where disyllabic feet are phonological-phrase final. Tone sandhi for non-phonological-phrase final foot will be discussed later in Section 5.2.4.

First syllable is of Tone I:

LLH.σ	Max-RP(SD)	*R(SD)	Id-Pitch	*Contour
a. LL␣.XY␣	∗!			
b. ␣LH.XY␣		∗!		∗
☞ c. ␣HH.XY␣				
d. ␣LL.XY␣			∗!	

Table 13: First syllable is of Tone I

In Table 13, σ represents an underlying tone which is realized as XY. Constraints that have been omitted from the tableaux are (1) constraints that have never been violated: Max-3P, *Cross-Foot-T, Dep-Pitch; (2) constraints that is violated the same number of times by all candidates: Max-Pitch, and (3) constraints that are low-ranked: the two Doub constraints.

It is also worth noting that Max-RP(SD) is outranked by some other constraints in many southern Ganyu dialects, where Tone I uniformly changes into LL in non-phonological-phrase-final positions.

First syllable is of Tone II:

HHL.σ	Max-RP(SD)	Id-Pitch	Max-Pitch	*Contour
a. HH␣.XY	∗!		∗	
☞ b. ␣HL.XY			∗	∗
c. ␣HH.XY		∗	∗!	

Table 14: First syllable is of Tone II

In Table 14, the constraint Id-Shape(LD) has no effect on the first syllable and is therefore omitted. In addition, constraints without any violation by the candidates include $2\mu/\sigma$, Max-3Pitch, *Cross-Foot-T, Dep-Pitch, *R-BetweenFt, and two Doub constraints.

First syllable is of Tone III:

LHØ.σ_h	Max-RP(SD)	*R	Dep-Pitch	*Contour	Doub-P(R)
☞ a. ␣HH.HX					*
b. LH␣.HX	*!	*		*	
c. ␣HL.HX		*	*!	*	

LHØ.σ_l	Max-RP(SD)	*R	Dep-Pitch	*Contour	Doub-P(R)
☞ a. ␣HL.LX		*		*	*
b. LH␣.LX	*!	**		*	
c. ␣HH.LX		*	*!		

Table 15: First syllable is of Tone III

In the two tableaux in Table 15, the constraint *R is a shorthand for two rising constraints (a) and (b). The constraint Id-Shape(LD) is irrelevant for the first syllable. High-ranked constraints are assumed to have been obeyed by all spelt-out candidates.

In many tableaux, I represented syllable-combinations abstractly. I do not assume cyclic operations. σ_h represents a syllable input which will have an optimal output which starts with a high pitch; and σ_l represents a syllable input which will have an optimal output which starts with a low pitch.

There are two sources of L in the second target of LL. One is from the following syllable, one is from the first pitch target of the same syllable. So the output LL.LX could only violate one of the two non-Double Association constraints. Figure 8 is a concrete example.

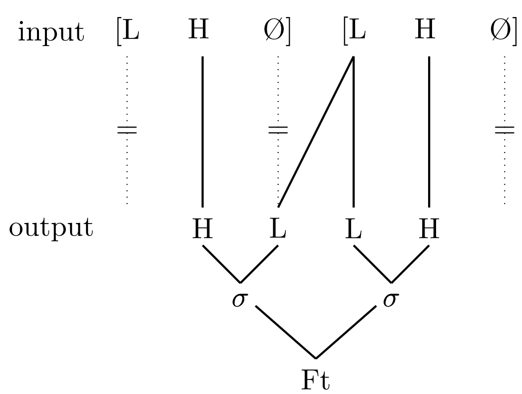


Figure 8: III-III sandhi

First syllable is of Tone IV:

Constraints that are always obeyed are: $2T/\sigma$, Id-Pitch, Dep-Pitch. Irrelevant constraints for our discussion include Max-3Pitch, *Cross-foot-T. The constraint Id-Shape(LD) is irrelevant for the first syllable. High-ranked constraints are assumed to have been obeyed by all spelt-out candidates.

HLØ.σ_l	MAX-RP(SD)	*R(SD)	*CONTOUR	*DOUB-P(R)	*DOUB-P(L)
☞ a. ␣LL.LX		*	*	(*)	(*)
b. ␣LH.LX		**!	**		
c. HL␣.LX	*!	*	**		

HLØ.σ_h	MAX-RP(SD)	*R(SD	*CONTOUR	*DOUB-P(R)	*DOUB-P(L)
☞ a. ␣LL.HX		*	(*)	(*)	(*)
b. ␣LH.HX		**!	*(*)		
c. HL␣.HX	*!	*	*(*)		

Table 16: First syllable is of Tone IV

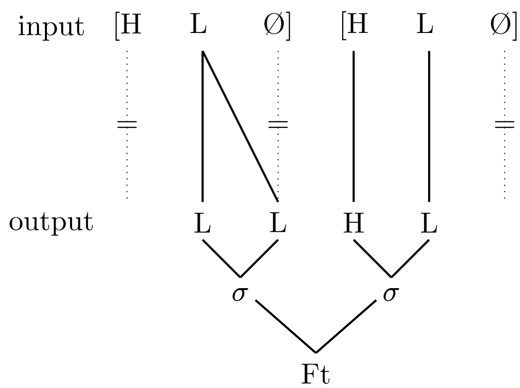


Figure 9: IV-IV sandhi

5.2.3 Trisyllabic tone sandhi

In a $\sigma_1\sigma_2\sigma_3$ final-foot trisyllabic tone sandhi, $\sigma_2\sigma_3$ usually follow disyllabic sandhi rules. However, there are a few exceptions, which could be mostly explained by the interaction of *R-BETWEENFT, which penalizes any rising contour at foot boundaries in the whole phonological phrase, ALIGN-R, and IDENT-PITCH constraints. In these cases, the major-minor foot structure is important.

In this section, I will show five representational cases of trisyllabic tone sandhi which deviates from disyllabic sandhi. All remaining cases of trisyllabic sandhi could be explained with similar tableaux.

[I-II]-II sandhi:

[LLH.HHL.]HHL	Id-Shape	Max-RP(SD)	*R-BetweenFt	Ident-Pitch	Contour
☞ a. ␣HH.␣HH.HH␣	*			**	
b. ␣HH.␣HL.HH␣	*		*!	*	*
c. ␣HH.HH␣.HH␣	*	*!		*	

Table 17: [I-II]-II sandhi

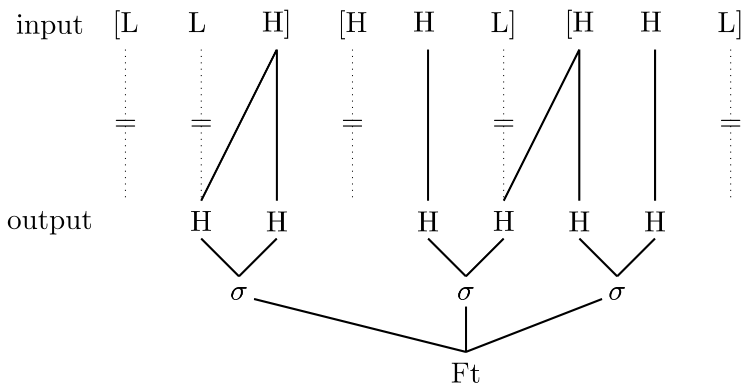


Figure 10: [I-II]-II sandhi's optimal output

As with previous tableaux, irrelevant, highest-ranked and lowest-ranked constraints are omitted (c.f (27)). From the above tableau, we can see that a rising portion at the phrase level, "L.H", is penalized. So the optimal output would rather change the L into an H.

If the major foot is parsed as [σ[$\sigma\sigma$]], the optimal output would be (b).

I-II-III sandhi:

LLH.HHL.LHØ	Max-RP(SD)	*R-BetweenFt	Ident-Pitch	*Contour
a. ␣HH.␣HH.LH␣	*		**!	*
☞ b. ␣HH.␣HL.LH␣	*		*	**
c. ␣HH.HH␣.LH␣	**!		*	*

Table 18: I-II-III sandhi

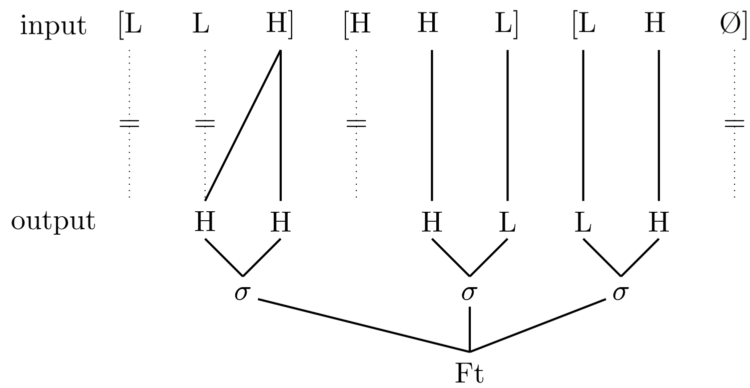


Figure 11: I-II-III sandhi's optimal output

In I-II-III sandhi, the internal foot structure is irrelevant because there is no cross-syllabic contour tones in the candidates filtered by higher constraints.

[II-I]-I sandhi:

[HHL.LLH.]LLH	MAX-RP(SD)	MAX-3P	*R-BETWEENFT	IDENT-PITCH
☞ a. ˽HL.˽HH.L^{C}H			*	**
b. ˽HH.˽HH.L^{C}H		*!		***
c. HH˽.˽HH.L^{C}H	*!	*		**

Table 19: [II-I]-I sandhi

In the input, there is a sequence of three consecutive L's. An output where none of them is realized is penalized by the constraint MAX-3P. As with previous tableaux, irrelevant, highest-ranked and lowest-ranked constraints are omitted (c.f (27))

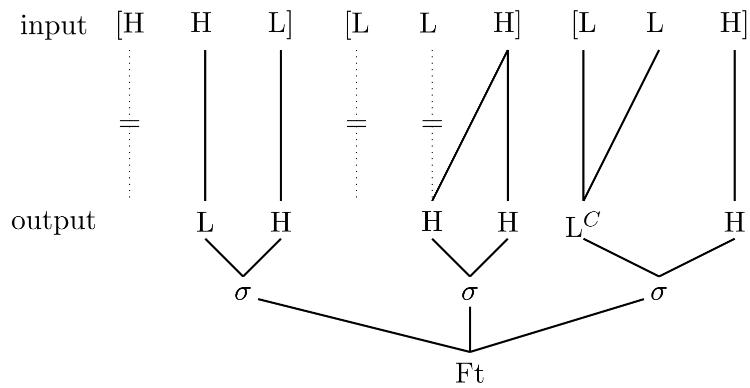


Figure 12: [II-I]-I sandhi's optimal output

If the major foot is parsed as [σ[$\sigma\sigma$]], the optimal output would be (b).

II-[II-I] sandhi:

HHL.[HHL.LLH]	Max-RP(SD)	*R-BetweenFt	Ident-Pitch
a. ␣HL.␣HL.L^CH		∗!	∗
b. ␣HL.HH␣.L^CH	∗!	∗	∗
c. ␣HL.␣HH.L^CH		∗!	∗∗
☞ d. ␣HH.␣HL.L^CH			∗∗
e. ␣HH.␣HH.L^CH			∗∗∗!

Table 20: II-[II-I] sandhi

As with previous tableaux, irrelevant, highest-ranked and lowest-ranked constraints are omitted (c.f (27)). And If the major foot is parsed as [[$\sigma\sigma$]σ], the optimal output would be (b).

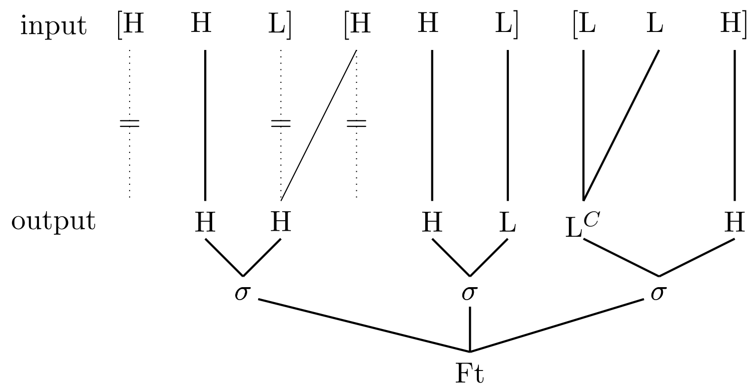


Figure 13: II-[II-I] sandhi's optimal output

II-[II-II] sandhi:

HHL.[HHL.HHL]	Max-RP(SD)	*R-BetweenFt	Ident-Pitch
a. ␣HH.␣HH.HH␣	*		**!
b. ␣HL.HH␣.HH␣	**!	*	
c. ␣HL.␣HH.HH␣	*	*!	*
☞ d. ␣HH.␣HL.HH␣	*		*
e. ␣HL.␣HL.HH␣	*	*!	

Table 21: II-[II-II] sandhi

As with previous tableaux, irrelevant, highest-ranked and lowest-ranked constraints are omitted (c.f (27)).

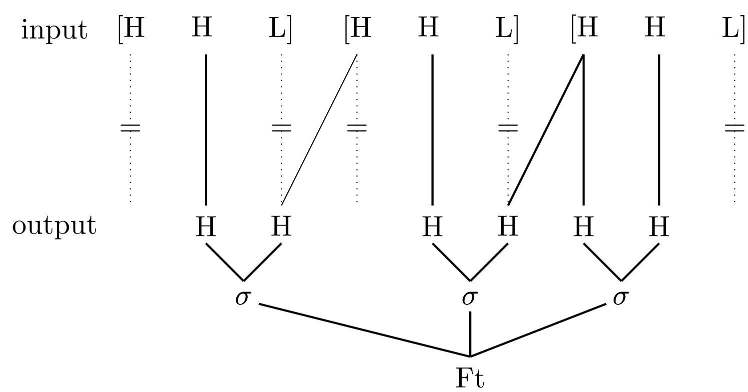


Figure 14: II-[II-II] sandhi's optimal output

5.2.4 Non-final foot sandhi

Non-final foot sandhi could be derived from the interaction of the same group of constraints which have been used in the explanation of final foot tone sandhi.

The most crucial difference between non-final feet and final feet is that in non-final feet, there is no syllable of longer duration. Therefore, the optimal output of the last syllable should always obey MAX-RIGHTMOSTPITCH because ID-SHAPE(LD) has no effect on it. Table 22 lists disyllabic non-final-foot sandhi.

1st syl. \2nd Syl	**I (L^CH)**	**II (HH)**	**III (LH)**	**IV (HL)**
I (L^CH)	HH.HH	HH.HH	HH.HH	HH.LL
II (HH)	HL.HH	HL.HH	HL.HH	HL.LL
III (LH)	HL.HH	HH.HH	HL.HH	HH.LL
IV (HL)	LL.HH	LL.HH	LL.HH	LL.LL

Table 22: Disyllabic non-final-foot sandhi

The most important constraint here is that the pitch target of one foot cannot be used in another foot because the strict hierarchical delimitation of phonological resources. I will refer to it as *CROSS-FOOT-T. It is a high ranking constraint. Based on current evidence, its ranking is the same as MAX-RIGHTMOST(PITCH). Therefore, in non-final-foot tone sandhi, Tone II and Tone III are always realized as HH when they are foot-final. The following example shows the derivation of [III-III] [III-III] tone sandhi. As with previous tableaux, irrelevant, highest-ranked and lowest-ranked constraints are omitted (c.f (27)).

[LHØ.LHØ][LHØ.LHØ]	CROSS-FOOT-T	MAX-R(PITCH)	*R(SD)	*R-BETWEENFT
a. [⌞HL.⌞HL][⌞HL.LH⌟]	*	*!		*
☞ b. [⌞HL.⌞HH][⌞HL.LH⌟]		*		
c. [⌞HL.LH⌟][⌞HL.LH⌟]		**!	*	

Table 23: [III-III] [III-III] sandhi

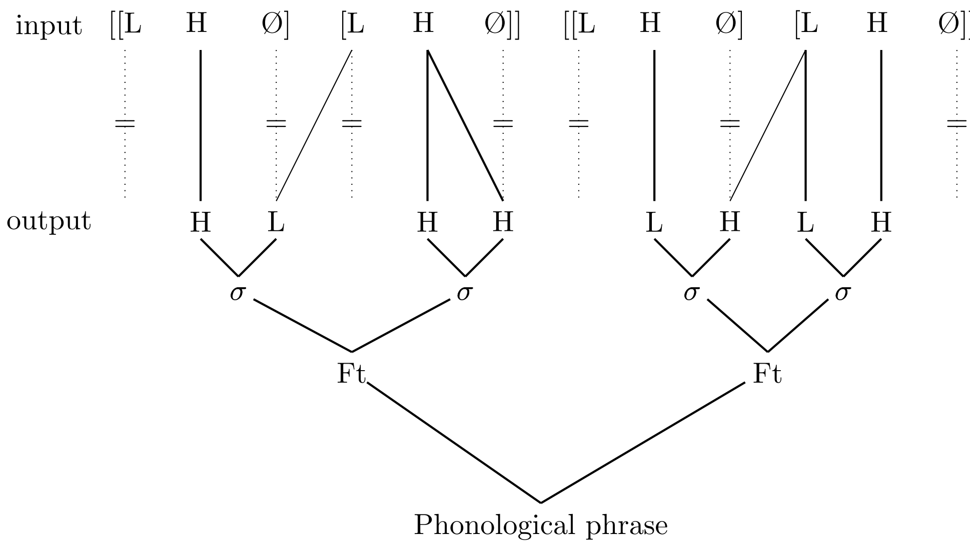


Figure 15: [III-III] [III-III] sandhi's optimal output

6 Conclusion

The thesis is the first attempt to analyze Ganyu tone sandhi from an explanatory perspective with the help of prosodic structure and Optimality Theory (OT).

In the thesis, I have answered the three questions as mentioned in Section 1.3. That is, (i) the proposed underlying tone forms are suitable for my specific analysis, as shown in Chapter 4; (ii) the proposed hierarchical prosodic structure, at least part of the whole structure which is useful for tone sandhi, as shown in Chapter 2; and (iii) I have provided explanations for prosodic arrangements and for specific sandhi processes, both from the perspective of OT, as shown in Chapters 3 and 5.

In addition, three innovative aspects in my analyses are worth mentioning.

First, I made an argument for the possibility that a syllable can have three pitch targets underlyingly, although at the surface form at most two pitch targets could be realized because of the various markedness constraints as well as the interrelation between different prosodic levels at the P-structure. Therefore, the whole underlying form is never actually realized. My proposal of underlying forms is based on investigation of the phonological behaviors of the various surface forms of a syllable's tone. My analysis of Ganyu tone sandhi provides a vivid example of how citation tones could be different from underlying tones and how they could be derived from underlying tones.

Second, I analyzed the ability to accommodate different pitch contours as the interaction between pitch saliency and positional prominence (of syllables). In syllables of longer duration, contour tones are more likely to appear in order to maintain maximal tonal contrast. But in syllables of shorter duration, i.e. non-phrase-final syllables, rising tone is the least favored one, while falling tone is less favorable than level tones. The ranking of pitch saliency in Ganyu is shown in (29).

(29) Tonal saliency: HH, LL $>$ HL $>$ LH

Third, non-final foot sandhi has rarely been touched upon in the literature. After examining the overall prosodic structure, I was able to provide a succinct and powerful explanation to non-final foot tone sandhi. In the future, I will extend my analysis to tone sandhi which involves pitchless syllables and focus-stressed syllables.

References

Andruski, J. E., & Ratliff, M. (2000). Phonation types in production of phonological tone: The case of Green Mong. *Journal of the International Phonetic Association*, *30*(1-2), 37–61.

Baxter, W. H., & Sagart, L. (2014). *Old Chinese: a new reconstruction.* Oxford University Press.

Chao, Y. R. (1930). ə sɪstəm əv "toun-lɛtəz". *Le Maître Phonetique*, *30*, 24–27.

Chen, M. Y. (2000). *Tone sandhi: Patterns across Chinese dialects* (Vol. 92). Cambridge University Press.

Cheng, C.-c. (1973). A quantitative study of Chinese tones. *Journal of Chinese linguistics*, 93–110.

Chomsky, N. (1986). *Barriers* (Vol. 13). MIT press.

Chomsky, N. (1998). *Minimalist inquiries: The framework* (No. 15). MIT Working Papers in Linguistics, MIT, Department of Linguistics.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* New York: Harper and Row.

Crowhurst, M., & Hewitt, M. S. (1995). *Directional footing, degeneracy, and alignment.* GLSA (Graduate Linguistic Student Association), Dept. of Linguistics, University of Massachusetts.

Davison, D. S. (1992). Parametric variation in pitch realization of 'neutral tone'syllables in mandarin. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 18, pp. 67–79).

Duanmu, S. (1990). *A formal study of syllable, tone, stress and domain in Chinese languages* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Duanmu, S. (1993). Rime length, stress, and association domains. *Journal of East Asian Linguistics*, *2*(1), 1–44.

Duanmu, S. (1994). Against contour tone units. *Linguistic inquiry*, *25*(4), 555–608.

Duanmu, S. (2007). *The phonology of standard Chinese.* Oxford University Press.

Duanmu, S. (2009). *Syllable structure: The limits of variation.* Oxford University Press.

Edwards, J., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *the Journal of the Acoustical Society of America*, *89*(1), 369–382.

Feng, S. (2009). *Interactions between morphology, syntax and prosody in Chinese (revised edition).* Perking University Press.

Fleischhacker, H. A. (2005). *Similarity in phonology: Evidence from reduplication and loan adaptation* (Unpublished doctoral dissertation). Citeseer.

Gandour, J. T. (1978). The perception of tone. In *Tone* (pp. 41–76). Elsevier.

Gandour, J. T., & Harshman, R. A. (1978). Crosslanguage differences in tone perception: A multidimensional scaling investigation. *Language and speech*, *21*(1), 1–33.

Gordon, M. (2007). *Syllable weight: phonetics, phonology, typology.* Routledge.

Gordon, M. J. (2002). A phonetically driven account of syllable weight. *Language*, *78*(1), 51–80.

Green, A. D. (1995). The prosodic structure of burmese: A constraint-based approach. *Working Papers of the Cornell Phonetics Laboratory*, *10*, 67–96.

Haegeman, L. (1994). *Introduction to government and binding theory.* Wiley-Blackwell.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies.* University of Chicago Press.

Hombert, J.-M. (1974). Universals of downdrift: their phonetic basis and significance for a theory of tone. *Studies in African Linguistics*, *5*, 169–183.

Horwood, G. (2002). Precedence faithfulness governs morpheme position. In *In proceedings of the twenty first west coast conference on formal linguistics, line.*

Hsieh, F.-f. (2005). Tonal chain-shifts as anti-neutralization-induced tone sandhi. *University of Pennsylvania Working Papers in Linguistics*, *11*(1), 9.

Hyman, L. M. (2009). The representation of tone. *UC Berkeley Phonology Lab Annual Reports, 5(5).*

Hyman, L. M., & Valinande, N. (1985). Globality in the kinande tone system. *African linguistics*, *6*, 239–260.

Iwata, R. (1989). Tone sandhi in the Qing-kou (Jiang-su, Gan-yu Prefecture) dialect (a new variety). *Journal of Asian and African Studies*, *37*, 221–236.

Javkin, H. R., & Maddieson, I. (1983). An inverse filtering study of burmese creaky tone. *UCLA Working Papers in Phonetics*, *57*, 115–125.

Jiǎng, X. (2005). *Hànyǔ yīnyùn fāngyán lùnwén jí (Papers on Chinese phonology and dialects).* Guizhou People's Publishing House.

Kager, R. W. J. (1995). The metrical theory of word stress. *Blackwell handbooks in* *linguistics*, *1*, 367–402.

Kornai, A., & Pullum, G. K. (1990). The X-bar theory of phrase structure. *Language*, 24–50.

Krishnan, A., Xu, Y., Gandour, J. T., & Cariani, P. A. (2004). Human frequency-following response: representation of pitch contours in chinese tones. *Hearing research*, *189*(1-2), 1–12.

Lamontagne, G. (1996). Relativized contiguity. *Ms., University of British Columbia*.

Landman, M. (2002). Morphological contiguity. *Carpenter et al.(2002)*, *139169*.

Lanham, L. W. (1958). The tonemes of Xhosa. *African Studies*, *17*(2), 65–81.

Li, X. (2004). Levels and classes of tone sandhi in the Chinese dialects [j]. *Fangyan (Dialect)*, *1*, 003.

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic inquiry*, *8*(2), 249–336.

Liberman, M. Y. (1975). *The intonational system of English.* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Lin, J.-w. (1994). Lexical government and tone group formation in Xiamen Chinese. *Phonology*, *11*(2), 237–275.

Liú, C. (2001). *Gànyú fāngyán zhì.* Beijing: Zhong Hua Book Company.

Luo, Z. (1999). Coercive subtyping. *Journal of Logic and Computation*, *9*(1), 105–130.

Martínez-Paricio, V., & Kager, R. (2015). The binary-to-ternary rhythmic continuum in stress typology: layered feet and non-intervention constraints. *Phonology*, *32*(3), 459–504.

Martin-Löf, P., & Sambin, G. (1984). *Intuitionistic type theory* (Vol. 9). Bibliopolis Napoli.

Matthews, S., & Yip, V. (2013). *Cantonese: A comprehensive grammar.* Routledge.

McCarthy, J. J. (2010). An introduction to harmonic serialism. *Language and Linguistics Compass*, *4*(10), 1001–1018.

McCarthy, J. J., & Prince, A. (1993). Generalized alignment. In *Yearbook of morphology 1993* (pp. 79–153). Springer.

McCarthy, J. J., & Prince, A. (1995). Faithfulness and reduplicative identity. *Linguistics Department Faculty Publication Series*, 10.

McCarthy, J. J., & Prince, A. (1999). Faithfulness and identity in prosodic morphology. *The prosody-morphology interface*, *79*, 218.

McDonough, J. M. (1999). Tone in Navajo. *Anthropological Linguistics*, 503–540.

McPherson, L. (2012). Underspecified tone in Tommo So (dogon, mali). In *Proceedings of the 29th west coast conference on formal linguistics. somerville: Cascadilla proceedings project* (pp. 169–177).

Molinier, G. (1920). *Les leys d'amours* (Vol. 17). E. Privat.

Myers, S. (1998). Surface underspecification of tone in Chichewa. *Phonology*, *15*(3), 367–391.

Napoli, D. J., & Nespor, M. (1979). The syntax of word-initial consonant gemination in Italian. *Language*, 812–841.

Nespor, M., & Vogel, I. (2012). Prosodic domains of external sandhi rules. *Sezione di Lettere*, *2*(1), 14–53.

Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, *395*(6698), 123.

Pan, W. (1982). Guanyu Hanyu shengdiao fazhan de jige wenti [several problems in the development of Chinese tones]. *Journal of Chinese Linguistics*, *10*, 359–385.

Pierrehumbert Janet, B., & Beckman, M. E. (1988). *Japanese tone structure.* Cambridge, Mass.: MIT Press.

Poser, W. J. (1989). The metrical foot in Diyari. *Phonology*, *6*(1), 117–148.

Prieto, P., D'imperio, M., & Fivela, B. G. (2005). Pitch accent alignment in Romance: primary and secondary associations with metrical structure. *Language and Speech*, *48*(4), 359–396.

Prince, A., & Smolensky, P. (1993). Optimality theory: Constraint interaction in Generative Grammar. *Optimality Theory in phonology: A reader*, 3–71.

Prince, A. S. (1980). A metrical theory for Estonian quantity. *Linguistic Inquiry*, 511–562.

Pulleyblank, D. (1986). *Tone in lexical phonology* (Vol. 4). Springer Science & Business Media.

Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook*, *3*, 371–405.

Selkirk, E. (1995). Sentence prosody: Intonation, stress, and phrasing. *The handbook of phonological theory*, *1*, 550–569.

Selkirk, E., & Shen, T. (1990). Prosodic domains in Shanghai Chinese. *The phonology-syntax connection*, *313*, 37.

Selkirk, E. O. (1980). The role of prosodic categories in English word stress. *Linguistic* *inquiry*, *11*(3), 563–605.

Shih, C. (1997). Mandarin third tone sandhi and prosodic structure. *Linguistic Models*, *20*, 81–124.

Shih, S.-h. (2017a). Binarity and focus in prosodic phrasing: New evidence from Taiwan Mandarin. In *Proceedings of the annual meetings on phonology* (Vol. 4).

Shih, S.-h. (2017b). Major phrases are binary: Evidence from Taiwan Mandarin flat structure. In *34th west coast conference on formal linguistics* (pp. 454–461).

Smith, J. L. (2004). *Phonological augmentation in prominent positions.* Routledge.

Steed, W., et al. (2006). Phonation type and tone sandhi as evidence of Chinese stress in Longquan Wu. In *Proceedings of the 11th Australian International Conference on Speech Science & Technology, Australian Speech Science & Technology Association inc.*

Sū, X., & Wàn, L. (2011). *Gànyú fāngyán yánjiū (Research in Ganyu dialect).* Zhonghua Book Company.

Tsay, J., & Myers, J. (1996). Taiwanese tone sandhi as allomorph selection. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 22, pp. 395–405).

Wayland, R., Zhu, Y., & Kaan, E. (2015). Perception of pitch contours by native and non-native tone listeners. *The Scottish Consortium for ICPhS*.

West, L. (2016). Pitch specification of minor syllables in Sgaw Karen. *The Journal of the Acoustical Society of America*, *139*(4), 2218–2218.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America*, *91*(3), 1707–1717.

Woo, N. H. (1969). *Prosody and phonology.* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Wurm, S. A., Li, R., & Baumann, T. (1987). *Language atlas of China.* Australian Acad. of the Humanities; Longman Group (Far East).

Yip, M. (2002). *Tone.* Cambridge University Press.

Yip, M. J. (1980). *The tonal phonology of Chinese* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Yip, M. J. (2001). Tonal features, tonal inventories and phonetic targets. *UCL Working Papers in Linguistics*, *13*, 161–188.

Yue-Hashimoto, A. O. (1986). Tonal flip-flop in Chinese dialects. *Journal of Chinese Linguistics*, *14*(2), 161–82.

Zhang, H. (2016). *Syntax-phonology interface: Argumentation from tone sandhi in Chinese dialects*. Taylor & Francis.

Zhang, J. (2004). The role of contrast-specific and language-specific phonetics in contour tone distribution. *Phonetically based phonology*, 157–190.

Zhang, J. (2007). A directional asymmetry in Chinese tone sandhi systems. *Journal of East Asian Linguistics*, *16*(4), 259–302.

Zhang, T., & Zhu, X. (2012). A triple-register tonal system: The case of Lianyungang [j]. *Dialect*, *3*, 003.

Zhu, X. (2006). A grammar of Shanghai Wu (vol. 66). *Munich, Germany: LINCOM Europa*.

Zhu, X. (2010). Yǔyīnxué (phonetics). *Beijing: Commercial Press*.

Zwicky, A. M., Kaisse, E. M., & Chen, M. Y. (1987). The syntax of Xiamen tone sandhi. *Phonology*, *4*(1), 109–149.

Zwicky, A. M., Kaisse, E. M., Hale, K., & Selkirk, E. (1987). Government and tonal phrasing in Papago. *Phonology*, *4*(1), 151–183.