PHONOLOGICAL PROMINENCE AND ITS INTERACTION WITH TONE IN CHINESE DIALECTS

by

Ping Wang

M.A., Tianjin Normal University, 2001

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

Kenneth P. Dietrich School of Arts and Sciences

This dissertation was presented

by

Ping Wang

It was defended on

August 12, 2014

and approved by

Dr. Hiroshi Nara, Professor, Department of East Asian Languages and Literature

Dr. Claude E. Mauk, Lecturer and Director of Less Commonly Taught Languages,

Department of Linguistics

Dissertation Advisor: Dr. David Mortensen, Visiting Researcher

Language Technology Institute, Carnegie Mellon University

Dissertation Advisor: Dr. Shelome Gooden, Associate Professor, Department of Linguistics

PHONOLOGICAL PROMINENCE AND ITS INTERACTION WITH TONE IN CHINESE DIALECTS

Ping Wang, PhD

University of Pittsburgh, 2015

Earlier studies on Chinese have recognized that there are prominent positions, and there are interactions between tone and prominent positions. However, the earlier studies have not rigorously applied phonetic or phonological inspections for establishing prominent positions in Chinese. When more rigorous phonetic and phonological examinations have been applied in this study, a more constrained and principled set of prominence phenomena emerged. This set excludes the postulation of a generalized trochaic system in Standard Mandarin, accentual prominence in New Chongming and peripheral prominence in Zhenhai. On the other hand, this set includes metrical prominence in the Northern Wu dialects and Fengkai Cantonese, and the interaction between tone and metrically prominent positions.

In this study, two types of interaction between metrical prominence and tone are attested. First, metrically strong positions are characterized by the preservation of lexical tones, or the ability to determine the shape of the neighboring tones. Thus, the stressed position normally licenses a larger range of tonal contrast; whereas, unstressed syllables tend to go tonal modification, reduction, or loss. Second, tone can condition stress placement. Observations made in the Northern Wu dialects suggest that stress assignment is sensitive to tone properties. In the Northern Wu group, the distribution of stress tends to avoid syllables with a low tone, or a short tone.

To summarize, although Chinese is widely recognized as a canonical tone language, stress and tone as two independent phonological properties do co-exist in Chinese. The co-existence of tone and stress leads to some interesting interactions. However, tone-stress interaction in Chinese produces a limited set of phonological processes, which is only attested in a limited number of dialects.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This dissertation would not have been possible without the guidance and support of many people. Above All, I would like to thank my dissertation advisor, Dr. David Mortensen, for his support and encouragement as I worked through the Ph.D. program and completed this dissertation. Despite his busy schedule as a professor in the linguistics department, David always found time to meet with me individually, to review my work, to give technical help in Praat scripting, and to answer my questions. I also owe a great deal of thanks to my co-chair and committee members, Shelome Gooden, Claude Mauk, and Hiroshi Nara for their thoughtful and helpful comments, advice and feedback. I have learned a great amount from them and I am grateful to have been the beneficiary of their expertise and attention.

I want to express my appreciation to the Department of Linguistics, from which I have received important support, financial and otherwise, to help address the costs of my doctorate study in the United States. I am grateful to Professor Marta Ortega-Llebaria, Yasushiro Shirai, Na-Rae Han and other faculty members for their in-class guidance and after-class advice and discussion. The teaching and sharing of linguistic knowledge and experience from the faculty have deepened my understanding of linguistics in general.

Thank you to all of my friends for keeping me in good spirits. In particular, I would like to extend my sincere gratitude to Jing Wang for being my company, listener and providing transportation back and forth from grocery stores. I would also like to thank the past and present

linguistics graduate students for providing a supportive and joyful atmosphere: Jie Cui, Adrienne Washington, Kristopher Geda, Derek Ho-Leung Chan, Nozomi Tanaka, Zoe Pei-sui Luk, Sally Kim, Andrew Jeske, Draucker Fawn, Lauren Collister, Abdesalam Soudi, Carrie Bonilla, Nausica Marcos Miguel, William Price, Maria Laura Lenardon, Meghan Stehle, Katherine Martin, Benjamin Friedline, Holman Tse, Lorraine Denman, and Mike Olsen.

I have been enormously supported by my family. The understanding and continued love that my father Jiaji Wang and my mother Guiying Yin have provided me have incredibly encouraged me through tough times. I have been blessed with their love and support in all aspects of my life. I also thank my borther Bingwei Wang for widely contacting his friends in America in order to find support for me in the early days after my first arrival in Pittsburgh. I also heartily thank Zhanwu Liu and Song Nan for the willingness to provide temporary accommodation and find a long-term apartment for me. Thank you, Jim (I don't even know your last name), for the generous help and constant encouragement when they were most needed.

My advisor from my Master's degree, late Professor Jialing Wang, deserves mention as my first teacher in linguistics. He nurtured my interest in languages, learning and linguistics as he encouraged me to pursue further study. I dedicate this dissertation to the memory of Jialing Wang, my linguistic advisor for life.

1.0 INTRODUCTION

1.1 REVIEW OF PREVIOUS STUDIES ON PROMINENCE IN CHINESE

Word prosodic systems are prototypically either tonal or metrical. Chinese is widely accepted as a canonical tone language. Such being the case, a great number of phonological studies of Chinese have been devoted to tone.

1.1.1 Studies on tone sandhi in Chinese

The study of tone has a long tradition in Chinese linguistics, which can date back to the rime book *Qieyun* (AD 601) where four historical tone categories were firmly established. The earliest recorded attempt to investigate tone sandhi is found in the fourteenth century pronouncing dictionary *Zhongyuan Yinyun* (Chen, 2000). In modern times, systematic descriptive work on Chinese dialects dates back to the 1920s. Liu (1924) is a pioneering work of the cross-dialectal study of Chinese tones. It initiated the modern tradition, in Chinese linguistics, of paying close attention to phonetic details achieved by experimental methods and phonetic instruments. Liu's research inspired other early studies by Luo (1930) and Wang (1931) on tonal behavior of Chinese dialects. Since then, numerous works, including field investigations and doctoral dissertations, have been devoted to topics related to Chinese tone systems. Especially with the

onset of the publication of journal *Fangyan* in 1979, linguists have been exposed to the diversity of tone systems and tone sandhi phenomena in Chinese dialects.

Beginning in late 1970s and 1980s, tone became an active area in theoretical phonology. With the concepts and formalisms of autosegmental approach to tones, generative work during this period of time was centered on African tone systems (Goldsmith, 1976). A notable exception is Yip's work (Yip, 1980), which deals with the tonal dialectology of Chinese (Hyman, 2011). Later, Bao (1999) provides a model with a more complex geometry of tone to address the theoretical issue regarding the phonological representation of tones. The integration of contemporary phonological developments and the fruitful results of experimental studies, as well as field investigations greatly broadened and deepened the knowledge and understanding of tones in Chinese.

Over the past several decades, no other area of Chinese phonology and phonetics has attracted as much attention from phonologists and phoneticians as tones.

1.1.2 Studies on stress in Standard Mandarin

Compared to the efforts devoted to the study of tone, stress has been one of the least studied aspects of Chinese phonology. In some influential textbooks on modern Chinese which are still widely used today (Hu, 1981; Qian, 1990; Zhang, 1983), one can hardly find any mention of stress of Chinese. It is also not clear about how much consensus there is with regard to whether lexical stress exists in Chinese, or at least in some dialects. Gao & Shi (1963) clearly states that there is no stress at the lexical level in, for instance, Standard Mandarin. Hyman (1977) and Selkirk & Shen (1990) also hold the no-stress view for Chinese in general, or in a given dialect under their discussion. Linguists sometimes speak or write on the subject with great conviction,

but the views that are expressed often conflict on both general and specific issues. For instance, there is disagreement as to whether there is lexical stress in Chinese, or whether the Shanghai dialect has a strictly left-headed metrical system (Duanmu, 1995; Zhu, 1996).

Chao (1968: 147) asserts that in Mandarin "a syllable is either stressed, in which case it has a tone, or unstressed and in the neutral tone." In Mandarin, neutral tone is sometimes referred to as "light syllable" or "weak syllable" (Cao, 2007). The term light syllable in this context is not referring to the moraic content, but the impressionistic perception that a neutral tone is articulated with less energy. Neutral tone syllable occurs in the final position in disyllabic sequences or non-final positions in multisyllabic sequences, but it never occurs in word-initial position. In most cases, a neutral tone syllable closely follows a regularly toned syllable. Neutral tone syllables are underlyingly toneless, and short in duration. Since they are atonic underlyingly, their pitch is largely determined by the preceding regular tone.

Before Chao's study in 1968, linguists had already made attempts to sort out the complex nature of neutral tone and made pronouncements on how neutral tone should be defined phonologically. Early discussions of the neutral tone center on its tonal characteristics. Zhang (1947) claims that neutral tone is the "fifth tone" of Beijing Mandarin.[1] Other linguists proposed that neutral tone is a tone sandhi phenomenon (Luo & Wang, 1957), which has been a dominant view among Chinese linguists. Chao (1968) is a pioneering work in that it pronounces that the neutral tone in Mandarin has to do with metrical prominence.

Some recent studies still focus on the tonal aspect of the neutral tone. For instance, attentions have been devoted to the phonetic details on the pitch value, pitch range, and tonal co-articulation of neutral tone in different contexts (Shen, 1990; Y. Wang, 1993). Wang (2002)

[1] Beijing Mandarin is spoken in the urban area of Beijing, China. It is well accepted that there are four tones in Beijing Mandarin. Beijing Mandarin is classified as a Mandarin dialect. Mandarin covers a large and diverse group of dialects spoken across northern, central, and southwestern China.

applies regular tone sandhi rules to neutral tone alternations, achieving an integrated account of tone behavior in Tianjin dialect. However, there is a growing consensus among linguists that the nature of the neutral tone cannot be fully addressed without considering its prominence property. Li (1981) points out that the neutral tone is lexically unstressed. Li seems to be the first to argue that, similar to English and Russian, there is word stress in Standard Mandarin,[2] because there is a clear stress contrast between a regular syllable and a neutral tone syllable. For instance, the following two words are otherwise identical, if they are not differentiated by stress.

(1) da yi HL HL '*main idea*'

da yi HL ø[3] '*careless*'

In the first meaning, the word *da yi* consists of two regular syllables. In the second meaning, the second syllable of the word is realized with a neutral tone, which is produced with shorter duration and reduced rhyme. For native speakers, the contrast between a regular syllable and a neutral tone syllable, as shown in (1), is easy to perceive.

1.1.3 Is there stress in regular disyllabic words in Standard Mandarin?

Yin (1982) extends the discussion of stress to regular disyllabic words in Beijing Mandarin. A regular disyllabic word is a word consisting of two regular syllables, which is held in contrast to a disyllabic word with a word-final neutral tone. He proposes that there are three levels of prominence at word level: heavy stress, medium stress, and light stress.

[2] Standard Mandarin (also known as Standard Chinese, or Putonghu) is used as the standard dialect in China. The phonology of Standard Mandarin is based on Beijing dialect. In the literature of phonology, Beijing Mandarin, Standard Mandarin, Standard Chinese, and Putonghua are often used interchangeably.

[3] The symbol "ø" is used to indicate the toneless status of the neutral tone syllable.

The neutral tone syllable is believed to have light stress, in view of the fact that it is well accepted as the least prominent. When a disyllabic word is composed of a regular syllable and a following neutral tone syllable, the regular syllable is noticeably more prominent than the neutral tone syllable.

The purported contrast between "heavy stress" and "medium stress" is held between two regular syllables. A heavy stress + medium stress combination is trochaic, whereas a medium stress + heavy stress is iambic. Being "medium" is claimed to be intermediately prominent—less stressed than a heavy stress syllable but more stressed than a light stress (neutral tone) syllable. While a neutral tone syllable is always unstressed, a regular syllable can be either heavy-stressed or medium-stressed.

However, the three-way contrast has been subject to numerous objections. Although the prominence contrast is well established between a regular syllable and a neutral tone syllable, it is extremely difficult to judge the relative degree of prominence between two regular syllables. In other words, the contrast between the "heavy stress" and the "medium stress" is very hard to perceive.

Around fifty years of research on stress since Fry (1955, 1958) has shown that stress can be phonetically cued by one or more of the properties such as increased loudness, greater duration, peripheral vowel quality, etc. In addition, stress realization is context-sensitive, subject to language specific constraints. Cumulative evidence from previous research has shown that duration is a consistent correlate of word stress across languages (Ortega-Llebaria & Prieto, 2011). This observation receives support from Lin & Yan (1980) and Lin, Yan & Sun (1984). It has been reported that the neutral tone syllable is 30% to 50% shorter than a regular syllable. The result of experiment in Lin (1983) provides evidence that the perception of neutral tone syllable

is primarily cued by duration. Furthermore, the neutral tone syllable exhibits rhyme reduction (Lin & Yan, 1980). However, the intensity of a neutral tone syllable is not necessarily less than that of a regular syllable (Lin & Yan, 1980). Therefore, the lack of stress of neutral tone is well supported by phonetic correlates.

However, the contrast between a heavy-stressed syllable and a medium-stressed syllable is not convincingly supported by phonetic studies in both production and perception. Apart from the minimal pairs showing sharp contrast between [T. T] (regular disyllabic words) and [T. ø] (words with a neutral tone) as shown in (1), Yin (1982) also reports minimal pairs of the following type :

(2)	Words	Tones	Chinese	Gloss	Stress Pattern
a.	gong shi	HH HL	工事	'*fortification works*'	medium stress-heavy stress
	gong shi	HH HL	攻势	'*the offensive*'	heavy stress-medium stress
b.	gong ji	HH HH	公鸡	'*rooster*'	medium stress-heavy stress
	gong ji	HH HH	攻击	'*attack*'	heavy stress-medium stress
c.	sheng qi	HH HL	生气	'*to get angry*'	medium stress-heavy stress
	sheng qi	HH HL	生气	'*vitality*'	heavy stress-medium stress

Yin claims that the pairs of words in (2) are minimal pairs. The first word in each pair is said to have the medium stress-heavy stress pattern, whereas the second has the heavy stress-medium stress pattern.

However, Lin & Yan (1990) expresses doubt about this claim. Their production experiment does not confirm that a given regular disyllabic word is produced with a stable stress pattern. Similarly, the perception test in Lin, Yan & Sun (1984) does not support the stress

contrast demonstrated in (2) either. Cao (1989, 1994) finds that the duration of syllables in regular disyllabic expressions are roughly the same. A so-called heavy-stressed syllable is not necessarily longer than a medium-stressed one.

To date, it is widely agreed that the neutral tone syllable is unstressed. The postulation of the three-way stress contrast in Standard Mandarin remains controversial. This state of affairs leads to a continuous debate over whether there is stress in regular disyllabic words in Standard Mandarin.

1.1.4 Tone and phonological prominence

Neutral tone as a weak element is widely attested in dialects spoken throughout China. Cantonese (Yue) is the only dialect group where neutral tone is reportedly absent. Field linguists have published a great number of descriptive reports on the tonal behavior of the neutral tone. However, the descriptive studies are by and large based on impressionistic methodology. Studies on the acoustic properties of the neutral tone in dialects other than Standard Mandarin are scarce.

Studies on phonological prominence other than metrical stress in Chinese are even more limited in number. The inadequacy of descriptive and theoretical studies obscures the understanding of prominence phenomena of Chinese in general.

Some studies have started to discuss the role phonological prominence may play in the Chinese tone system. Chen (2000) argues that some Northern Wu dialects possess essentially accentual systems since they display accent-like properties such as culminativity and edgemostness. In New Chongming dialect, a member of the Northern Wu group, for instance, there can be at most one tone-bearing syllable in a trisyllabic or longer rhythmic unit. In addition, the tonic syllable tends to occur at the leftmost or the rightmost margin, which is recognized as

edgemostness, a function of demarcating accent/stress of head-terminal foot construction (Chen, 2000:226).

Li (2003) distinguishes peripheral prominence and metrical prominence in Zhenhai. Peripheral prominence is defined as an abstract prominent property associated with domain boundaries. The co-existence of the two types of prominence is argued to give rise to complex surface tonal patterns in Zhenhai.

There are also explicit claims that tone sandhi in Chinese is closely related to stress. Kennedy (1953), Yip (1980), and Chan (1995) all argue that the tone system and metrical system co-exist in Chinese, and stress plays a pivotal role in shaping tonal behavior. In particular, Duanmu (1995, 1999) proposes that the tonal sandhi domain is determined by metrical effects in Shanghai and Xiamen.

Bao (2003, 2004) argues against the metrical analysis of the tone system in the Northern Wu dialects. Bao points out some crucial problems in the relevant studies, particularly Chen (2000) and Duanmu (1995, 1997). For instance, Bao views it problematic that the metrical analysis can derive variable size of foot, which ranges from a monosyllabic form to a polysyllabic form without exhibiting the rhythmic effect typical of metrical systems. Secondly, the metrical analysis cannot correctly predict the domain where the tone sandhi rules operate, especially in polysyllabic compounds.

1.1.5 Summary

The no-stress view for Chinese is now a minority position among linguists. The neutral tone has been firmly identified as being lexically unstressed. However, there is no conclusive evidence to confirm that there is lexical stress in words consisting of two regular syllables. Although there

are studies discussing the role of phonological prominence in tone systems of Chinese, these attempts have not been wholly successful.

1.2 THIS STUDY

Earlier studies on Chinese phonology have recognized that there are prominent positions, and that there are interactions between tone and prominent positions in Chinese (Chen, 2000; Duanmu, 1997, 1999; Kennedy, 1953; Yip, 1980). However, the earlier studies have not rigorously examined phonetic, phonological, or general linguistic data for establishing prominent positions in Chinese. Therefore, the prominence-based analyses of phonological data have not been always carried out in principled ways. This dissertation will show that when more rigorous phonetic and phonological analysis is conducted, a more constrained and principled set of prominence phenomena will emerge. This set excludes the postulation of a generalized trochaic system in Standard Mandarin, [4] accentual prominence in New Chongming, or the dual prominences in Zhenhai. On the other hand, this set includes the existence of metrical stress in the Northern Wu dialects and Fengkai Cantonese, and the interaction between tone and stress.

Previous investigations have sought to apply metrical prominence to explain a wide range of tone sandhi phenomena (Chen, 2000; Shih, 1986; Yip, 1980; Wright, 1983). However, in this study, only two types of tone-stress interaction are attested in some representative dialects of Chinese. The first type of interaction is that lexical tones are preferably preserved in the metrically strong positions, while the lack of stress can often induce tone reduction, modification

[4] Although there is a consensus that the neutral tone syllable is lexically unstressed in Standard Mandarin, no conclusive evidence supports that there is lexical stress in words composed of two regular syllables. Considering the research findings up to date, the postulation of a trochaic system for Standard Mandarin as a whole is too strong. I therefore refer to such a postulation as the generalized stress view of Standard Mandarin.

and complete loss. As a result, the stressed position licenses a wider range of tonal contrasts whereas unstressed syllables can be modified to low tones, assigned a default tone, or acquire a tone through tone spreading. Thus, tonal neutralization can be triggered by the absence of metrical stress.

The second type of interaction is that tone can condition stress placement. Data from Zhenhai, a Northern Wu dialect, suggests that stress assignment depends on tone in the disyllabic domain. Stress in Zhenhai is not found in syllables which bear a low tone, or a short tone. In this dialect, stress is not only tone-sensitive, but also weight-sensitive.

To summarize, although Chinese is generally recognized as a tone language, there can be metrical prominence in Chinese dialects. The co-existence of tone and stress leads to some interesting, but limited set of interactions, which are attested in a limited number of dialects, such as Northern Wu dialects and Fengkai Cantonese.

1.3 OUTLINE OF THIS DISSERTATION

This dissertation is structured as follows. Chapter 1 provides the introductory and necessary background knowledge for our investigation of phonological prominence in Chinese. In addition, this part will briefly review the major works that have been done on the subject.

A close examination of the metrical analysis of Standard Mandarin proposed by Duanmu (2007) will be presented in Chapter 2. Duanmu argues strongly that metrical structure has to be incorporated into the phonology of Standard Mandarin. Otherwise, Duanmu contends, some linguistic observations such as word-length variation and word-order variation won't be

revealingly explained. However, scrutiny of Duanmu's generalized stress view shows that the suggested metrical analysis can be problematic and technically inadequate at times. This chapter will present an alternative analysis to Duanmu's from a grammatical perspective. It shows that we can make sense of, and explain, certain puzzling linguistic facts Duanmu was unable to explain if grammatical and semantic factors are taken into consideration. In addition, this chapter will show that previous acoustic studies do not support the proposition that there is lexical stress in regular disyllabic words in Standard Mandarin. This chapter concludes that there is no convincing evidence to extend trochaic stress to regular syllables in Standard Mandarin.

Pitch-accent is claimed to exist in some Northern Wu dialects. Chapter 3 will re-examine the pitch-accent reinterpretation of the New Chongming dialect and the Shanghai dialect (Bao, 2004; Chen, 2000). This chapter argues that accent prominence in New Chongming is not lexically specified, metrically determined, or rule-based as in other prototypical accentual systems. Instead, evidence shows that it is derived from tones. Therefore, I will argue that New Chongming exhibits a reduced tonal system rather than a pitch-accent system. For Shanghai, rich experimental data (Zhu, 2005) supports the existence of a metrical system, where both trochaic and iambic feet can occur.

Chapter 4 will argue against the recognition of multiple prominences in Chinese. With regard to the alleged peripheral prominence observed in Zhenhai, a close examination will show that it is not the kind of inherent/abstract prominence associated with peripheral positions suggested by Zoll (1997), but an emerging left-headed stress which co-exists with the more conservative iambic (right-headed) stress. It will be argued that the sole prominence attested in Chinese so far is metrical in nature.

Chapter 5 will review the previous metrical analyses of tonal operations in Chinese. In addition, this chapter will offer a case study of stress-driven tone sandhi found in Cantonese. It shows that the OT model of tone-stress interaction proposed by De Lacy (2002) works nicely with Fengkai Cantonese. But, the OT model needs to be further developed before it can cover a wide range of tone-stress interaction, for instance, tone deletion and tone spreading in Shanghai.

The final chapter summarizes the findings in this dissertation.

2.0 STUDIES ON METRICAL PROMINENCE IN STANDARD MANDARIN

2.1 DUANMU'S METRICAL STRESS VIEW

2.1.1 Introduction

Duanmu (1990, 1995, 1999, 2007) has conducted the most comprehensive and in-depth study of metrical stress in Standard Mandarin to date. Duanmu argues that a metrical system exists in Standard Mandarin and some other Chinese dialects, such as Shanghai and Southern Min (Taiwanese), despite the fact that stress is not always perceivable, or phonetically manifested.

Standard Mandarin has been confirmed as having lexical stress, as a result of the phonetic studies on the so-called neutral tone conducted in the past thirty years (Cao, 1995). A neutral tone syllable and a regular syllable show considerable contrast in their phonetic realizations. First of all, as the name suggests, a neutral tone syllable does not carry a lexical tone. Neutral tone syllables never occur word-initially. By and large, the pitch with which a neutral tone syllable is realized is determined by the preceding syllable with a regular tone. In contrast, regular syllables are underlyingly specified with a tone, and are louder as well as longer than the neutral tone counterparts. In addition, neutral tone syllables show great reduction in the rhyme. In this regard, a neutral tone syllable not only lacks lexical tone, but also lacks stress. Therefore, Duanmu (2000, 2007) refers to a regular syllable as full syllable, and a neutral tone syllable as weak syllable or light syllable. Full syllables can take stress whereas weak syllables cannot. For native

speakers, it is easy to perceive the stress contrast between full syllables and weak syllables (Duanmu, 2007; Lin & Yan, 1980).

The previous phonetic studies give us a better understanding on the nature of the neutral tone syllable. There is now consensus among linguists that the neutral tone syllable is not just neutralized in terms of tone, but also phonetically unstressed (Cao, 1995; Li, 1981; Lin & Yan, 1990), and that the tone neutralization is likely a result of the lack of stress. As far as Standard Mandarin is concerned, it would not be controversial if Duanmu's metrical stress analysis were to apply only to the contrast between regular syllables and neutral tone syllables. However, Duanmu (1990, 1995, 2000, 2007) argues for a general metrical stress system for Standard Mandarin, which holds that there is not only stress contrast between the regular syllable and the neutral tone syllable, but also between regular syllables. Hereafter, this view will be referred to as the generalized metrical stress view. In the coming sections of this chapter, I will argue against this view.

Chao (1968) points out that it is often difficult to obtain agreement from native speakers on the location of stress in a regular disyllabic word of Standard Mandarin. Phonetic studies do not support the existence of stress in disyllabic words either. For instance, phonetic correlates observed cross-linguistically for unstressed syllables such as vowel reduction, shorter duration are not typically found in the claimed unstressed regular syllable in Standard Mandarin.[1] Therefore, agreement has not been reached yet with regard to whether there is word level stress in Standard Mandarin, aside from the disyllabic words with a neutral tone syllable.

To account for the fact that the so-called stress in Standard Mandarin is often not possible to perceive because of the lack of phonetic cue, Duanmu (2000, 2007) suggests that stress is

[1] "Unstressed regular syllable" is the equal of the "medium-stressed" syllable in Yin's three-way contrast proposal. Duanmu (2007) does not explicitly suggest a three-way contrast as Yin does. For Duanmu, a neutral-tone syllable is always unstressed, and a regular syllable can either be stressed or unstressed. See §1.1.3 for Yin's proposal.

realized abstractly in Standard Mandarin. For the abstract realization of stress, Duanmu proposes a dual trochee metrical system. In this proposal, stress is sensitive to both moras and syllables, in that Standard Mandarin has both moraic feet and syllabic feet (Duanmu, 2007). With the dual trochee proposal, Duanmu argues that the existence of stress in Standard Mandarin is revealed by phrase/compound formation patterns, in particular the word-length patterns and word-order patterns.

2.1.2 Metrical system in Standard Mandarin

2.1.2.1 Dual trochee

Duanmu (1999, 2007) assumes that Standard Mandarin has a dual trochee system because a foot contains both a moraic trochee and a syllabic trochee. Duanmu posits that a foot must have (at least) two beats at both syllabic and moraic levels. Moreover, stressed syllables must be heavy. Duanmu's assumptions of the metrical system of Standard Mandarin can be summed up in (1).

(1) a. Trochee (at both moraic and syllabic levels)

The head of a foot is on the left.

b. Binarity (at both moraic and syllabic levels)

A foot must have (at least) two beats.

c. A stressed syllable must be heavy.

(1c) requires that a stressed syllable must be heavy, which reminds us of the Weight-to-Stress Principle. The basic idea of this principle is that syllables are measured on a weight scale: light, heavy, and superheavy; and a heavier syllable is more likely to attract stress than a lighter one.

(2) The Dual Trochee

x	x	
(s s)	(s s)	Syllabic foot
(m m). (m m)	(m m). m	Moraic foot
x x	x	
heavy- heavy	heavy- light	

(3) Bad foot structure

x	x	
(s s)	(s s)	Syllabic foot
m. m	m. (m m)	Moraic foot
	x	
light-light	light-heavy	(Duanmu, 2007:139)

Neutral tone syllables are phonetically short in duration, which can be rendered phonologically as a light syllable with one mora as shown in (2) and (3). A monomoraic syllable cannot form a monomoraic foot, which violates the Binarity rule.[2] A monomoraic syllable can form a foot with another syllable at syllabic level. However, it cannot be stressed according to the requirement of (1c), which explains the bad foot structures in (3).

Under Duanmu's dual trochee proposal, the light-light and light-heavy disyllabic words are ill formed, because they violate the rule in (1c) when stress falls on the light syllable in the syllabic foot. Therefore, the absence of word-initial neutral tone syllable can be easily explained

[2] Both monosyllabic and monomoraic feet have been claimed to exist. Being smaller than binary, monosyllabic and monomoraic feet are referred to as degenerate feet. Degenerate feet are often discriminated against by metrical rules and conditions. For instance, degenerate feet are repaired by various strategies such as lengthening and reparsing (Kager, 1995). In Duanmu's dual trochee proposal, binarity is required for foot formation. Therefore, monomoraic and monosyllabic feet are not allowed to occur in Standard Mandarin.

in the present analysis, in the sense that legitimate syllabic trochee cannot be possibly formed in light-heavy or light-light disyllabic words.

2.1.2.2 Empty beats

Although Duanmu assumes trochaic (left-headed) stress for Standard Mandarin, he accepts the conclusion from other works that the dominant stress pattern in disyllabic words is iambic. A survey in Xu (1982) receives wide recognition, which offers a rough count of the stress patterns of 20,000 commonly-used disyllabic words and compounds in Standard Mandarin. Xu (1982) does not explain the methodology and procedures in the survey. Considering the fact that the standard phonetic description of Mandarin in early 1980s was largely based on impressionistic data, it is most likely that the conclusion in the survey is dependent on Xu's impressionistic judgment as a native speaker of Beijing Mandarin.

(4)

	Pattern	Count	%
a.	Initial stress (heavy-light)	1,500	7.5%
b.	Initial stress (heavy-heavy)	4,500	22.5%
c.	Final stress (heavy-heavy)	14, 000	70.0%
	All	20, 000	100.0%

(Duanmu, 2007:141)

Xu's data in (4), which is quoted in Duanmu (2007), show that the dominant stress pattern in Standard Mandarin is the word-final stress in disyllabic words and compounds. However, this result is largely in contradiction to the trochaic stress argued strongly for Standard Mandarin by Duanmu (1999, 2000, 2007). Thus, Duanmu (2007) attributes the final stress dominance in Xu's data to the effect of final lengthening. Duanmu argues that the final stress of

a disyllabic word changes to initial (left-headed) stress when other word follows, since final lengthening is not possible for words in phrase-medial or compound-medial position.

In order to maintain his dual trochee proposal in light of Xu's data, Duanmu (2007) assumes an empty beat in the foot structure of words with final stress. An empty beat is a beat which does not correspond to an overt syllable. In Duanmu's scheme, the empty beat must be phonetically concrete. In particular, he assumes that when an empty beat occurs in word-final position, it is realized either as a pause or as the lengthening of the preceding syllable.

(5)

x	
s (s Ø)	Syllabic foot
(mm). (mm). Ø	Moraic foot
x x	
su she	
'*dorm*'	

The stress derivation in (5) demonstrates how an empty beat can be used to assign word-final stress in a heavy-heavy disyllabic word. As seen above, the word-final stress is assigned to the last syllable in (5), because the last syllable forms a trochaic foot with the empty beat. Consequently, the first syllable has to be left unfooted at syllable level, and therefore cannot bear a word level stress.

However, the empty beat solution is not entirely successful in that it cannot explain the co-occurrence of words with initial stress and words with final stress in phrase-final positions or in isolation. According to Duanmu, all final stress is derived from initial stress when final lengthening effect is available. But, not all words with initial stress are lengthened to derive final stress when they occur in phrase-final position or in isolation. For instance, *huo che* (heavy-

heavy), meaning “train” in English, is considered to have final stress in isolation; whereas, *guan li* (heavy-heavy), which means “management”, is listed as having initial stress in isolation in Xu (1982). In response to this situation, Duanmu acknowledges that “it is better to recognize final stress directly as a stress pattern” (Duanmu, 2007:142), which is, to some degree, admitting the weakness of his dual trochee theory.

Duanmu argues strongly for a generalized trochaic system in Standard Mandarin, even though such an argument is not always supported by acoustic evidence. His strongest evidence for this view is from linguistic patterns observed in the formation of phrases and compounds in Standard Mandarin, i.e. the word-length variation and word-order variation. Duanmu believes metrical stress is the key to the understanding of these variations. In the following section, we will see, just as the empty beat solution is not satisfactory, the metrical stress analysis is not theoretically satisfactory either in that it still leaves many problems on the word-length and word-order variations unanswered. Moreover, the problems can be better accounted for if grammatical and semantic factors in these variations are taken into consideration. Thus, this state of affairs leads to serious doubt on whether it is necessary to pursue a generalized trochaic stress system for Standard Mandarin, especially when such a stress system cannot successfully cover the language data and is not motivated by other independent linguistic phenomena observed in Standard Mandarin.

2.1.3 A metrical account of word-length variation

2.1.3.1 Introduction

The word-length variation in Chinese refers to the fact that most disyllabic words in Chinese have monosyllabic alternatives, but the use of the monosyllabic and disyllabic alternatives is restricted when words are combined together to form compounds or phrases.

(6)	Disyllabic	Monosyllabic	Gloss
a.	mei-tan	mei	
	coal-charcoal	*coal*	*'coal'*
b.	shang-dian	dian	
	business-store	*store*	*'store'*
c.	da-suan	suan	
	big-garlic	*garlic*	*'garlic'*
d.	lao-hu	hu	
	old-tiger	*tiger*	*'tiger'*
e.	zhong-zhi	zhong	
	plant-implant	*plant*	*'to plant'*
f.	yan-jing	yan	
	eye-eyeball	*eye*	*'eye'*

In the above examples, the extra syllable in the disyllabic form is semantically redundant. For instance, in (6a) *mei-tan* is semantically repetitive in that *mei-tan* does not mean "coal and charcoal" but "coal". Similarly, in (6d) *lao-hu* does not mean an "old tiger", it just means a "tiger". The presence of two synonymous forms, one monosyllabic and one disyllabic, creates a

dual vocabulary system in Chinese (Duanmu, 2007). The dual vocabulary phenomenon does not occur sporadically; most Chinese words have dual forms (Duanmu, 2007).

Although the dual forms of Chinese words can be used alternatively in most cases, there are restrictions on the use of dual forms in [N N] (Noun-Noun) compounds and [V O] (Verb-Object) phrases, as illustrated in (7), where 1 indicates a monosyllable, 2 a disyllable.

(7) a.

	[N	N]
[2 2]	ji-shu (*skill-technique*)	gong-ren (*labor-people*)
[2 1]	ji-shu (*skill-technique*)	gong (*laborer*)
*[1 2]	ji (*skill*)	gong-ren (*labor-people*)
[1 1]	ji (*skill*)	gong (*laborer*)
Gloss	'*mechanic*'	

b.

	[V	O]
[2 2]	zhong-zhi (*plant-implant*)	da-suan (*big-garlic*)
*[2 1]	zhong-zhi (*plant-implant*)	suan (*garlic*)
[1 2]	zhong (*plant*)	da-suan (*big-garlic*)
[1 1]	zhong (*plant*)	suan (*garlic*)
Gloss	'*to plant garlic*'	

When both constituent lexemes in the above compound/phrase have dual forms (a monosyllabic form and a disyllabic form), there are four possible word-length combinations: [2 2], [2 1], [1 2], and [1 1]. For a [N N] nominal compound, all other combinations are fine except [1 2]. Similarly, of the four combinations of a [V O] phrase, while three of them are acceptable, [2 1] is usually ill formed. As noted by Lü (1963), this contrast is quite general in Chinese: for

[N N] compounds, [1 2] is usually ill formed; for [V O] phrases, [2 1] is. The native intuition on such contrast is quite sharp.

2.1.3.2 The stress-length approach

Duanmu and Lu (1990) argues that word-length variation in Chinese is governed by stress. They propose that in a two-word construction, the word with more stress should not be shorter than the word with less stress. This proposal was further developed in Duanmu (2007). The main ideas of the proposal can be summed up in below.

(8) a. *Nonhead Stress*

In the syntactic structure [X XP] (or [XP X]), where X is the syntactic head and XP the syntactic nonhead, XP should be stressed. (Duanmu, 2007:146)

b. *Stress Length*

Phrasal stress should be carried by a syllabic foot. (Duanmu, 2007:174)

c. *Anti-Allomorphy*

A stressed word should keep the same phonological shape. (If a word has a disyllabic shape, it should be used when the word has phrasal stress.) (Duanmu, 2007:174)

d. *Foot Shelter*

A foot or potential foot can be treated as a word whose internal morphosyntactic structure can be ignored. (Duanmu, 2007:126)

The Nonhead Stress rule (8a) was first proposed in Duanmu (1990), which assigns stress to compounds and phrases. About the Nonhead Stress rule, Duanmu (2007:146) explicates that:

> In standard X-bar syntax, the head X is an element at the word (or affix) level, the nonhead XP, an element at the phrase level. Since there are more possible phrases than

possible words (or affixes), the occurrence of a nonhead (phrase) is less predictable than the occurrence of a head (word or affix). Therefore, the information load of XP is far greater than that of X. By the Information-Stress Principle, therefore, the default phrasal stress should go to XP.

The analysis of compound stress is similar. A compound [N1 N2] is parallel to a phrase [XP X], where N1 is the counterpart of XP and N2 that of X. Therefore, N1 should receive phrasal stress.

The Stress Length rule (8b) and Anti-Allomorphy (8c) are proposed in Duanmu (2007). The Stress Length requirement can be satisfied "either by two syllables (SS) or by one syllable (SØ), where Ø is an empty beat" (Duanmu, 2007: 174). Anti-Allomorphy requires that if a word has a disyllabic form, it should be used when the word carries phrasal stress. The Foot Shelter rule (8d) applies to any disyllabic unit. A potential foot is a disyllabic unit with final stress: S (SØ). It is considered a potential foot in the sense that, when it occurs in non-final positions, the empty beat will not be present, it will then become (S S).

The requirements in (8) can be applied to the variations of [N N] (Noun-Noun) compounds and [V O] (Verb-Object) phrases to account for the flexible word length. The analysis of [N N] compounds is shown in (9), where [2 2], [2 1], and [1 1] are acceptable and [1 2] is unacceptable. The "#" symbol in (9) indicates a word boundary.

(9) [N N] compounds

a. Second noun is (SS) or (SØ)

[2 2]	[2 1]	[1 1]	*[1 2]
x x	x x	x	x
(S S)#(S S)	(S S)#(S Ø)	(S # S)	S # (S S)

b. Second noun is S(SØ)

[2 2] *[1 2]

x x x

(S S)#S(S Ø) S#S(S Ø) (Duanmu, 2007:175)

According to the Nonhead Stress rule, the first noun in [N N] should have the phrase-level stress. The first noun therefore should form a syllabic foot according to the requirement of Stress Length rule. Thus, [2 2] and [2 1] are well formed in terms of word length, since the first disyllabic noun meets the requirement that a phrasal stress should be carried by a syllabic foot. In [1 1], the two syllables can be treated as a single word under the protection of Foot Shelter, so there is no problem either. One might wonder whether the ill-formed S#S(SØ) for [1 2] in (9b) can be restructured to (S#S)(SØ) so that the phrase stress can be realized on the nonhead. Duanmu (2007) argues that (S#S)(SØ) is not acceptable due to "syntactic or perceptual reasons, because of the effect of Foot Shelter, (S#S)(SØ) is likely to be perceived as [2 1]" (Duanmu, 2007:175). Thus, [1 2] is the only ill-formed word length combination for [N N] compounds.

Turning to the word length patterns of [V O] phrases, all but [2 1] are well formed. Among the four combinations, [1 2] and [2 2] are most common. [1 1] is comparatively marginal and [2 1] is generally unacceptable. The metrical structures of the four possible word length combinations of [V O] phrases are shown in (10).

(10) [V O] phrases

a. Object is (SS) or (SØ)

[2 2] *[2 1] [1 1] [1 2]

x x x x x x

(S S)#(S S) (S S)#(S Ø) S#(S Ø) S#(S S)

b. Object is S(SØ)

[2 2] [1 2]

x x x

(S S)#S(S Ø) S#S(S Ø) (Duanmu, 2007:175)

In line with the Nonhead Stress rule, the object in [V O] phrases should carry phrasal stress. The verb has no phrasal stress but can have word stress if it is disyllabic. All the forms with a disyllabic object satisfy the requirement of the Stress Length rule, since phrasal stress is always carried by a binary syllabic foot. Therefore, there is no problem for [2 2] and [1 2]. In the cases of [2 1] and [1 1], the Anti-Allomorphy rule is violated, because the object has a disyllabic form which should be used when the phrasal stress falls on the object. This explains why [2 1] is usually ill formed. Duanmu further argues that [1 1] is not as unacceptable, because Foot Shelter treats it as a word, a VO compound, and "this compound is unrelated to the disyllabic object so there is no violation of Anti-Allomorphy" (Duanmu, 2007:176).

Although [1 2] and [2 1] are generally unacceptable for [N N] and [V O] respectively, there are still some counterexamples which take the undesired word-length combination, as shown by the examples in (11).

(11)

[N	N]	[V	O]
1	2	2	1
xi	zhu-ren	sheng-chan	lü
department	head	produce	aluminum
'*department chair*'		'*to process aluminum*'	

Duanmu argues that these counterexamples contain fixed-length words; accordingly, the fixed-length words do not have word-length variation. In the above examples, *lü* "aluminum"

and *xi* “department” do not have disyllabic forms. Although they carry phrasal stress, they cannot be lengthened to a disyllabic form, and therefore they do not violate Anti-Allomorphy. Furthermore, in the exceptional [2 1] type of [V O] phrase, the Stress Length rule can be easily satisfied by assuming an empty beat in the foot structure: (S S)#(SØ). For the exceptional [1 2] type of [N N] compound, Duanmu suggests the foot structure (S#S)(SØ) where the nonhead can be parsed into a binary foot to realize the phrase level stress, so that the Nonhead Stress rule can be satisfied. In this case, Duanmu does not further explain why the structure of (S#S)(SØ) is syntactically all right even though the restructuring of S#S(SØ) into (S#S)(SØ) is banned for other [1 2] type of [N N] compounds due to syntactic and perceptual reasons as seen in (9b).

2.1.3.3 A crucial problem of the stress-length approach

A crucial problem of Duanmu’s proposal is that it fails to explain the word-length variation of [A N] nominal compounds, where A indicates an adjective. In this part, we will elaborate on this problem.

Wang (2001) insightfully points out that nominal compounds can be [A N] (Adjective-Noun) compounds, apart from the [N N] (Noun-Noun) compound discussed by Duanmu.

Like nouns and verbs in Chinese, adjectives can alternate between monosyllabic and disyllabic forms. Therefore, a [A N] compound also has four possible word-length combinations: [1 2], [2 2], [2 1] and [1 1]. According to Duanmu’s Nonhead Stress rule, the phrasal stress carrier in a [A N] nominal compound should be the A(djective), because it is the counterpart of the nonhead. Therefore, the A(djective) should be a disyllabic form in line with the Stress Length rule. Thus, [2 2] and [2 1] types of [A N] should be fine. [1 1] should be fine too, under the protection of Foot Shelter. The [1 2] type of [A N] should be banned. Nonetheless, Wang (2001) shows that this is not the case:

Table 1. Word length combinations for [V O] phrases, [N N] and [A N] compounds

[V O] Phrase	[N N] Compound	[A N] Compound
*[2 1]	*[1 2]	*[2 1]
[2 2]	[2 2]	[2 2]
[1 2]	[2 1]	[1 2]
[1 1]	[1 1]	[1 1]

Wang (2001) agrees that [2 1] and [1 2] are the most disfavored word-length combinations for [V O] and [N N] respectively. Notably, she argues that the most restricted word-length combination for [A N] is [2 1] rather than [1 2].

The ill-formedness of [2 1] type of [A N] is quite unexpected, if Duanmu's proposal is the explanation for the word-length variation. [2 1] satisfies the rules proposed by Duanmu perfectly: the adjective is a disyllabic form, which qualifies the adjective (nonhead) to be a phrasal stress carrier.

Some examples of the [A N] compound are shown in (12):

(12)

	[A	N]
[2 2]	rou-ruan (*tender-soft*)	tou-fa (*head-hair*)
*[2 1]	rou-ruan (*tender-soft*)	fa (*hair*)
[1 2]	ruan (*soft*)	tou-fa (*head-hair*)
[1 1]	ruan (*soft*)	fa (*hair*)
Gloss	'*soft hair*'	

Wang (2001) further points out the [2 2] type of [A N] is rather restricted too, though it is not as ill formed as [2 1]. [1 1] is less restricted; the least restricted is [1 2].

Wang's observations are important ones because they hold with native speakers' intuitive judgment. Native speakers can easily sense some ungrammaticality in [2 1] type of [A N] compounds. In addition, the examples of [2 1] type of [A N] are not as numerous as other length

combinations for [A N] in Standard Mandarin. For instance, the following instances of the [2 1] type of [A N] compounds never occur grammatically in Standard Mandarin:

(13)	[A	N]	Gloss
	*gao-da (*high-big*)	fang (*house*)	'*big house*'
	*qiang-zhuang (*powerful-strong*)	ren(*man*)	'*strong man*'
	*xing-fu (*happy-blessed*)	jia (*family*)	'*happy family*'
	*gan-jing(*dry-clean*)	yi (*clothes*)	'*clean clothes*'

In § 2.1.4, we will continue to work on the word-length variation of [A N], [N N] and [V O] in Standard Mandarin, and try to approach it from a grammatical perspective first proposed by Wang (2001).

2.1.4 A grammatical account of word-length variation

2.1.4.1 Syntactic functions: nouns, verbs and adjectives in Standard Mandarin

One may wonder what causes the apparent discrepancy between [A N] and [N N] compounds with regard to the word length well-formedness. Wang (2001) proposes a different solution to the problem. Wang believes the word-length restrictions on [A N], [N N] and [V O] are results of: 1) the uneven number of disyllables in nouns, adjectives and verbs, and 2) the different grammatical functions the three word categories developed in the disyllabification progression of Chinese.

In the historical development of Chinese, disyllabification first started with nouns (Chen, 1987). Nouns lead verbs and adjectives in terms of the number of disyllables. The statistics of *Putonghua San Qian Chang Yong Ci Biao* (3,000 Frequently Used Words in Standard Mandarin)

offers clear evidence that the number and percentage of disyllables vary considerably among the three word categories (Wang, 2001: 247).

Table 2. Number and Percentage of monosyllabic words and disyllabic words by word category

	Total	Monosyllable	Monosyllable (%)	Disyllable	Disyllable (%)
Noun	1753	312	18%	1287	73%
Verb	1028	422	41%	604	59%
Adjective	479	143	30%	329	69%

3,000 Frequently Used Words in Standard Mandarin was completed by the Chinese Language Reform Committee Research and Popularization Office in 1959. The study involved about forty scholars for nearly three years. The frequency of the words in the list was based on a study of a selection of modern written texts of around 130,000 characters (Duanmu 2007:160). According to the data in Table 2, for the three word categories, disyllables comprise 73% of nouns, 59% of verbs and 69% of adjectives.

Adjectives and verbs have more monosyllabic forms than nouns, whereas the majority of nouns are disyllables. These facts partly explain why [1 2] is much more common for [A N] and [V O] than [N N]. But, the statistics cannot adequately answer why certain word-length combinations are avoided.

Wang (2001) points out that the key to the understanding of the avoidance of certain word-length combinations is relevant to the different syntactic functions that nouns, verbs, and adjectives developed in the history of vocabulary disyllabification. This view can be represented by the statements and examples in (14), (15), and (16).

(14) Disyllabic verbs show considerable word class flexibility. Lexical flexibility is the relative fluidity of a word being used for different grammatical functions. It is common in Chinese that a disyllabic verb can shift between the syntactic functions of verbs and nouns. Monosyllabic verbs do not usually show such fluidity.

a. wo *huai-yi* ta shi zei.
I suspect he is thief
'I suspect he is a thief.'

b. zhe zhi-shi wo-de *huai-yi.*
this only is my suspicion
'This is only my suspicion.'

Change of lexical class in Chinese is not often associated with morphological marking. Thus, the same word can have more than one lexical class without overt morphological changes. As shown by the above sentences in (14a) and (14b), the same disyllabic form *huai-yi* can function either as a verb or as a noun without being marked by any derivational affixes.

However, the fuzzy verb-noun boundary is not commonly observed in monosyllabic verbs. The vast majority of monosyllabic verbs cannot be used for syntactic functions other than what prototypical verbs can perform. Li (1990) and Liu (1992) are also aware of such a distinction. Below are some examples adapted from Li (1990) and Liu (1992).

c. *zhong-zhi/zhong* shu-cai
plant vegetable
'to plant vegetables'

d. *zhong-zhi/*zhong* fang-fa
plant method
'*the planting method'*

e. huai-ren *qi-pian/pian* le wo-men.
bad person cheat *asp* us
'*the bad person cheated us.'*

f. huai-ren de *qi-pian/*pian*
bad person 's cheat
'*the bad person's cheating*'

In (14c), for instance, both *zhong-zhi* and *zhong* mean "to plant", and both of them can make (14c) grammatical. However, only *zhong-zhi* can be used in (14d). This is because the modifier in (14d) requires a nominal, and only a disyllabic verb can function as a nominal. Similarly, *pian*, the monosyllabic form for "cheat" can be used as a verb in (14e), but not in the nominal position in (14f).

Lexical flexibility is not just confined to verbs. Adjectives in Mandarin also show this fluidity. As previously mentioned, the change of word class is not normally manifested by morphological changes in Chinese. The lexical flexibility of adjectives is not marked by morphological changes.

(15) Adjectives (both disyllabic and monosyllabic) can function as adjectives or verbs in Chinese. Furthermore, when used as verbs, disyllabic forms can be used in causative environment and followed by a NP object, whereas monosyllabic forms cannot.

As a group, adjectives in Mandarin are, in many ways, like stative verbs and because of this, they have been treated as a special type of verb (Chao, 1968). Like verbs, adjectives in Chinese can function as predicate. When they do, they do not need to be preceded by a linking verb equivalent to such verbs as *is, become* and *grow* in English.

a. ni zhen *yu-chun/chun*.
you really stupid
'*you are really stupid.'*

b. ta-de shen-ti hen *qiang-zhuang/zhuang.*
his body very strong
'*he is very strong physically.*'

c. ta-de zhang-xiang *chou-lou/chou.*
his looking ugly
'*He looks ugly.'*

The adjectives above, both the monosyllabic and the disyllabic forms, behave like stative verbs and can be modified by degree adverbs. However, monosyllabic and disyllabic adjectives differ in whether or not they can express causative events when functioning as verbs.

Compare the following pairs of adjectives. It is easy for native speakers to see that the disyllabic forms can stand alone in the causative environment, while the monosyllabic ones cannot:

d. yun-dong neng-gou *qiang-jian/*qiang* shen-ti.
excercise can strong body
'*Exercise can make people strong*.'

e. zhe-zhong xiang-bo ke-yi *nong-mi/*mi* tou-fa.
this kind shampoo can thick hair
'*This kind of Shampoo can make one's hair thick.*'

f. rou-shun-ji ke-yi *rou-ruan/*ruan* yi-wu.
fabric conditioner can soft clothes
'*fabric conditioner can make clothes soft*.'

In the above examples (15d~15f), the adjective is functioning as a verb indicating that the subject in the sentence causes a change of state of the NP object. As shown, the disyllabic adjective can be grammatically followed by an NP object in the causative environment. In contrast, the monosyllabic adjective cannot be used as verbs in the same environment.

(16) Monosyllabic nouns in Mandarin are more often bound morphemes than the disyllabic nouns. Therefore, bound monosyllabic nouns often cannot stand alone to form phrases or compounds with other words. In contrast, disyllabic nouns are usually free to combine with other words to form phrases or compounds.

Wang (2001) does not provide statistic data to prove monosyllabic nouns are more likely to be bound forms than disyllabic nouns. However, if we examine the monosyllabic and disyllabic nouns in Mandarin, it is often the case that disyllabic nouns are the compounded or affixed free forms developed out of a bound monosyllabic morpheme.

For instance, *zan*, which means "hairpin", cannot be used by itself. In order to be used freely, it has to be attached with a suffix *–zi*. The derived form z*an-zi*, still meaning "hairpin", is a free form. Compounding is another way to develop free words from bound forms. *fa-zan,* which literarily means "hair hairpin", is a compounded word which adjoins two bound forms, *fa* "hair", and *zan* "hairpin".

In the following sections §2.1.4.2, §2.1.4.3 and §2.1.4.4, the observations presented in (14), (15) and (16) will be applied to the word-length variation of the relevant compounds and phrases in Standard Mandarin, on the basis of a grammatical analysis presented in Wang (2001).

2.1.4.2 A grammatical analysis of word-length variation of [A N] compounds

In line with the fact that [2 1] is the least common word-length combination for [A N], Wang (2001) claims that [2 1] is the least preferred [A N] form. Wang argues that the [2 1] type of [A N] tends to be perceived as [V O] phrases, because disyllabic adjectives can be followed by NP objectives when they are used to indicate causative events. This fact excludes quite a number of [2 1] from the pool of [A N] nominal compounds. On the other hand, many [2 1] type of [A N] compounds cannot be taken as grammatical [V O] phrases either, even when the disyllabic

adjective is interpreted as expressing a causative meaning. For instance, forms like **rou-ruan fa* ([2 1]) "to make hair soft", **qiang-zhuang shen* ([2 1]) "to make body strong" are never considered grammatical by native speakers. Wang (2001) explains that the noun in the [2 1] combination is a monosyllabic form which is more likely to be a bound form, and therefore cannot be freely used in [V O] phrases.

Thus, only a small number of prototypical disyllabic adjectives can be used in the [2 1] type of [A N]. The disyllabic adjectives are prototypical in the sense that they can only function as adjectives. Furthermore, the monosyllabic noun in the [2 1] type of [A N] is required to be a free form. These two requirements considerably narrow down the number of qualified [2 1] type of [A N] nominal compounds.

(17) a. gan-jing *ku-zi/*ku*
dry-clean trousers-affix/ trousers
'*clean trousers*'

b. gan-jing *xie*
dry-clean shoes
'clean shoes'

c. hu-tu *nao-dai/*nao*
bewildered brain-bag/ brain
'*muddlehead'*

d. hu-tu zhang
chaotic account
'chaotic account'

Examples in (17) demonstrate how the two requirements are regulating the word-length combination choices for the [A N] nominal compounds. The disyllabic adjectives *gan-jing* "clean", and *hu-tu* "bewildered" are prototypical adjectives. Therefore, they won't be

misinterpreted as verbs. However, only the disyllabic nouns are chosen to collocate with them in (17a) and (17c), because the monosyllabic forms in the pairs are bound forms. The [2 1] combinations in (17b) and (17d) are grammatically well formed, because the monosyllabic nouns *xie* "shoes" and *zhang* "account" are free morphemes already.

Apart from the [2 1] type of [A N], the [2 2] type of [A N] is also rather restricted in its use, although it is not entirely ungrammatical. The [2 2] type of [A N] can be ambiguous, shifting between the reading of a nominal compound and a [V O] phrase.[3] The ambiguity leads to the avoidance of [2 2] at times. When an attributive particle *de* proceeds [2 2], it will be interpreted as a [A N] compound only and therefore won't cause ambiguity. In Chinese, particle *de* is used to mark possession or modification. The presence of particle *de* indicates a post-particle nominal position. See the following examples for illustration:

(18) a. si-chou ban de rou-ruan tou-fa
silk like *attributive particle* soft hair
'*silk like soft hair'*

b. shan-ding de xi-bo kong-qi
mountain-top *attributive particle* thin air
'*rarified air at mountain top'*

The most common word-length combination for [A N] is [1 2], which is the perfect combination for [A N] in that the monosyllabic adjectives won't be mistaken as a verb followed by a NP object in the causative environment, and the disyllabic nouns are free forms.

The [1 1] type of [A N] compounds is quite common as well. Bound monosyllabic nouns can normally occur in [1 1]. The combination of a monosyllabic adjective and a monosyllabic noun is one of the basic ways to create free noun forms in Standard Mandarin. For instance,

[3] Recall that disyllabic adjectives can be used as verbs to express causative meaning, where they are followed by NP objects. See statement (15) in §2.1.4.1 for details.

chang fa “long hair” and *mei nü* “beautiful woman” are grammatical, though the second nouns are bound forms. In addition, the [1 1] type of [A N] is usually regarded as a single lexical item rather than a compound. Notably, the [1 1] type of [V O] phrases is quite common as well. Similarly, the [1 1] type of [V O] is often lexicalized. Follow-up discussion of the lexicalization of [V O] phrases will be given later in this chapter.

It can be seen from the above analysis that the grammatical approach can straightforwardly explain the preference and avoidance of certain [A N] word-length combinations. It also catches the fact that [2 2] type of [A N] is not as common as [1 2], though it is not entirely unacceptable.

In §2.14.3, we will extend the grammatical analysis to [V O] word-length variation.

2.1.4.3 A grammatical analysis of word-length variation of [V O] phrases

For [V O] phrases, [1 2] is the most well-formed combination. In this combination, the monosyllabic verb is a prototypical verb. And, the disyllabic object is usually a free noun form. Consequently, [1 2] is unambiguously a qualified [V O] phrase.

The vast majority of [1 1] are well formed too. Interestingly, many of the [1 1] type of [V O] phrases are referred to as VO *verbs* (as opposed to [V O] *phrases*), since they, as a matter of fact, are used as disyllabic intransitive verbs (Li & Thompson, 1981).[4] The VO verb is often associated with a lexicalized meaning. For instance, *da jia* which literally means “to beat price” has been lexicalized with the meaning “to bargain”. More discussion about the lexicalization of VO verbs will be offered in the section on word-order variation.

[4] As mentioned in §2.1.4.2, the [1 1] type of [A N]/ [V O] is often a single lexical item rather than a compound or phrase.

The [2 2] type of [V O] is also well formed. It is worth pointing out that a [2 2] type of [V N] (Verb-Noun) structure is not necessarily a [V O] phrase; alternatively, it can be a nominal compound. A [2 2] type of [V N] is interpreted as a [V O] phrase, only when the N(oun) is the logical O(bject) undergoing the action indicated by the preceding V(erb). Otherwise, it is often interpreted as a nominal compound by native speakers. Compare the following pair in (19):

(19) a. zhong-zhi ji-shu
plant technique
'*planting technique*'

b. zhong-zhi da-shuan
plant garlic
'*to plant garlic*'

As seen in (19a), when the noun is not syntactically playing the role of a logical object of the preceding verb, a [2 2] type of [V N] structure can only be interpreted as a nominal compound. In (19a), the verb *zhong-zhi* (to plant) is nominalized and functions as a nominal modifier of the following noun *ji-shu* (technique). When the noun is playing the thematic role of theme/patient with respect to the action indicated by the preceding verb, the [2 2] type of [V N] is perceived as a [V O] phrase, as shown in (19b).

The [2 1] type of [V O] phrase is the ill-formed combination. [2 1] is the word-length combination to form [V N] nominals. As we already know, disyllabic verbs in Chinese can directly function as nominal modifiers because of the lexical flexibility. Zhang (1989) argues that [2 1] is the standard word-length template to form [V N] nominals. Such a combination is preferred, because monosyllabic nouns are frequently bound morphemes and bound nominal morphemes cannot stand alone in a [V O] phrase. Therefore, the monosyllabic noun in a [V N] is a clear indication of a verb-modified nominal construction. Moreover, the [2 1] type of [V N] is often considered a single lexical item because of the structural solidness caused by the adhesiveness of the bound head noun, although it is morphologically derived by compounding.

Consider the following [2 1] type of [V N] in (20), where all the disyllabic verbs function as modifiers of the following noun. Notably, the [2 1] type of [V N] is always interpreted as a nominal regardless of the syntactic relation between the V(erb) and the following N(oun). As seen in (20a), [2 1] is interpreted as a nominal even when the monosyllabic noun is syntactically the logic object of the preceding verb.

(20) a. yang-zhi xia
breed shrimp
'farmed shrimp' (as opposed to wild caught shrimp)

b. chu-zhu che
rent car
'*taxi*'

c. hong-gan ji
dry machine
'*dryer*'

d. jiang-luo san
descend umbrella
'*parachute*'

e. ju-zhu tiao-jian
live condition
'*living condition'*

The example in (20e) is a [2 2] type of [V N]. (20e) is a nominal compound rather than a [V O] phrase, where the noun is modified by the preceding nominalized verb. Note that the disyllabic noun in (20e) is syntactically not the logical object of the preceding verb, otherwise (20e) will be interpreted as a [2 2] type of [V O] phrase.

The rest examples listed above are [2 1] type of [V N] nominals. Interestingly, (20a) and (20b) present a case where the monosyllabic noun is the logic object of the preceding disyllabic

verb. Nevertheless, these two examples are still nominals rather than [V O] phrases, even though the noun undergoes the action indicated by the verb according to the thematic relation. The examples shown in (20c) and (20d) are nominals too. Noteworthily, the monosyllabic noun in (20c) and (20d) is syntactically the logic *subject* of the preceding verb. For instance, *san* which means "umbrella" in English, is the logical subject of the preceding verb *jiang-luo* "descend".

The examples in (20) illustrate that [2 1] type of [V N] is generally interpreted as nominals regardless of the syntactic relation between the disyllabic V(erb) and the monosyllabic N(oun). By contrast, for the [2 2] type of [V N], the syntactic relation between the V(erb) and the N(oun) is playing an important role in its grammatical interpretation.

2.1.4.4 A grammatical analysis of word-length variation of [N N] compounds

As to [N N] nominal compounds, it seems apparent that [2 2] is a good combination of word length. Nouns can be modifiers of nouns, and disyllabic nouns are free to be modifiers or to be modified.

The [1 1] and [2 1] combinations are acceptable too. As noted before, the [1 1] combination is the common way to develop free disyllabic nouns out of two bound forms. Thus, the [1 1] type of [N N] compounds are often lexicalized. Likewise, [2 1] is often lexicalized too, because of the adhesiveness of the second monosyllabic noun. Compared to the lexicalized status of [2 1] and [1 1], the [2 2] type of [N N] is structurally loose, and is grammatically somewhere between a phrase and a lexeme (Wang, 2001: 249).

The [1 2] word length combination is generally unacceptable for [N N] compounds. Wang (2001) argues that, similar to the [2 2] type of [N N] compounds, [1 2] is grammatically somewhere between a phrase and a lexeme. But, [1 2] is not as well formed as [2 2], because the monosyllabic modifier (the first noun) cannot offer as much information as a disyllabic modifier.

Therefore, [1 2] is an unpreferred redundant form. Secondly, monosyllabic nouns are often bound forms, and therefore cannot be used freely to modify another noun. Note that this is different from [1 1]. The [1 1] type of [N N] is often lexicalized. However, the [1 2] type of [N N] is not normally interpreted as a lexical item, due to the fact that the second disyllabic nouns are often free forms.

2.1.4.5 Summary

While Duanmu's proposal advocates a metrical approach to the word-length variation, Wang (2001) argues for a grammatical approach. The grammatical and semantic considerations discussed so far can account for the word-length preferences of compounds and phrases in Standard Mandarin. Moreover, the grammatical considerations Wang presents are sound ones in that they do not seem unnatural or counterintuitive.

Wang's study provides an alternative solution to the word-length variation. It empirically catches more facts of the word-length preference and avoidance, and therefore explains more data. For instance, Wang's grammatical approach can explain why [2 1] is usually the disfavored word length combination for [A N] nominal compounds.

Apart from the word-length variation, Duanmu also argues that the metrical factors are essential to the analysis of word-order variation in Standard Mandarin. In the next section, we will look into the word-order issue from Duanmu's metrical perspective, and then explore an alternative solution from the grammatical perspective.

2.1.5 A metrical account of the word-order variation

2.1.5.1 Introduction

Word-order variation is found in two types of Chinese compounds. The first is [V O N] compound, the second, [X Y N]. The word-order variation is illustrated with the following examples:

(21) a. [V O N] [O V N]

sui	zhi	ji	zhi-zhang	fen-sui	ji
break	paper	machine	paper	break	machine
'*shredder*'			'*shredder*'		

b. [X Y N] [Y X N]

da-xing	han-yu	ci-dian	han-yu	da	ci-dian
large-scale	Chinese	dictionary	Chinese	big	dictionary

Two [V O N] compounds are presented in (21a). Depending on the word length of the constituents, the compound can be either [V O N] or [O V N]. When V and O are both monosyllabic, [V O N] is the order realized. When V and O are both disyllabic, [O V N] is the only possible order.

In a [X Y N] compound, X and Y are modifiers of N. In (21b), the modifier *large* precedes another modifier *Chinese* in the [X Y N] compound; in the [Y X N] compound, the order is reversed. Again, it seems that the order variation is relevant to word length: *large* is disyllabic in [X Y N], but monosyllabic in [Y X N].

It is not immediately apparent why word order varies in Chinese [X Y N] and [V O N] compounds. Duanmu (1997, 2000, 2007) has made attempts to answer the question. Duanmu believes that the order variation in [V O N] and [X Y N] compounds is driven by metrical factors.

2.1.5.2 [X Y N] word-order variation: metrical structure improvement

Let's consider the [X Y N] compounds first. It is widely agreed that the sequence of the modifiers in languages is restricted by their meaning. For example, a partial hierarchy of modifier ordering is Size > Shape > Color > Provenance… (Quirk *et al*. 1972; Duanmu, 2007). This word order hierarchy is supposed to be universal.

With regard to a [X Y N] compound, there are eight word-length combinations when each constituent in a [X Y N] compound can be monosyllabic or disyllabic. Six of them do not show alternative word order, two have alternatives:

(22)	a. [X [Y N]]	da mu chuan	'*big wooden boat*'
	b. [X [Y NN]]	xiao hong deng-long	'*small red lantern*'
	c. [XX [Y N]]	xue-sheng shi tang	'*student dining hall'*
	d. [XX [Y NN]]	zhong-guo mian da-yi	'*Chinese cotton coat*'
	e. [XX [YY N]]	Shang-hai Huo-che Zhan	'*Shanghai Train Station*'
	f. [XX [YY NN]]	da-xing han-yu ci-dian	'*large-scale Chinese dictionary*'

(23)	a. [X [YY N]]	Dong Chang-An Jie
		East Chang-An Road
		'*East Chang'an Road*'
	Alternative: [YY [X N]]	Si-Chuan Bei Lu
		Si-Chuan North Road
		'*North Sichuan Road*'

b. [X [YY NN]] Xin Ying-Han Ci-dian

new English-Chinese dictionary

'*New English-Chinese Dictionary*'

Alternative: [YY [X NN]] Han-Yu Da Ci-dian

Chinese large dictionary

'*Large Chinese Dictionary*'

The above examples are quoted from Duanmu (2007: 193). The duplicated forms, such as YY, NN, or XX, indicate a disyllabic form. Singleton forms are monosyllabic. The bracketing in (22) and (23) is taken from Duanmu (2007), which is based on constituent analysis and semantics. For example, *da mu chuan*, which means "big wood boat" is not [[X Y] N] but [X [Y N]], because the former means "boat made of big wood" which is not the intended meaning. Similarly, *Han-yu Da Ci-dian* "Chinese Large Dictionary" is not [[YY X] NN] but [[YY [X NN]], because "Chinese large" is not a possible constituent, but "large dictionary" is.

Examples in (22) have a fixed word order, which exactly follows the word-order hierarchy listed earlier. Examples in (23) have an alternative word order. [X [YY N]] can be alternatively realized as [YY [X N]]. For instance, *Si-Chuan Bei Lu* "Si-Chuan North Road" is actually the north part of Si-Chuan Road, so this is a case where the name of the road is fronted. See the detailed glosses given in (23) for the [X [YY N]] and [YY [X N]] compounds.

Duanmu (2007) argues that the Y-fronting in the [X Y N] compound is motivated by the need to improve metrical structure. The foot boundaries of the [X Y N] compounds listed in (22) and (23) are shown in (24).

(24)	[X [Y N]]	(X Y) (N Ø)
	[X [Y NN]]	(X Y) (N N), (X Y) N (N Ø)
	[XX [Y N]]	(X X) (Y N), (X X) Y (N Ø)
	[XX [Y NN]]	(X X) (Y N) (N Ø)
	[XX [YY N]]	(X X) (Y Y) (N Ø)
	[XX [YY NN]]	(X X) (Y Y) (N N)
	[X [YY N]]	(X Y) Y (N Ø)
Alternative order	**([YY [X N]])**	(Y Y) (X N), (Y Y) X (N Ø)
	[X [YY NN]]	(X Y) Y (N N), (X Y) Y N (N Ø)
Alternative order	**([YY [X NN]])**	(Y Y) (X N) (N Ø)

The foot boundaries shown above are based on Duanmu (2007:193). According to Duanmu (2007), in the first six cases, there is no need to move X or XX close to the head noun, because they get footed in an appropriate way. In the first two cases, XY forms a foot, in the next four cases, XX does. The last two cases have an alternative word order, which both start with XYY. In [X [YYN]], YY should be stressed in the [YY N] compound according to the Nonhead Stress rule, therefore YY forms a foot; however, X cannot, unless the stress on YY is deleted to make it possible for X to form a binary foot with the following Y. If YY is fronted, it can form a foot and XN can form another, which is metrically well formed. Likewise, in [X [YY NN]], YY forms a foot; X cannot, unless the stress on YY is deleted. If YY is fronted, X and the following NN can be footed as (X N) (N Ø).

However, YY fronting is purportedly used alternatively to improve the metrical structure rather than being an obligatory requirement. This is evidenced by the co-existence of [X [YY N]] and [YY [X N]] (see examples in (23)). Duanmu suggests that there is a cost for both stress

deletion and YY fronting. “Stress deletion seems to have a cost” (Duanmu, 2007: 194), which is avoided sometimes so that YY fronting can happen. YY fronting may have a cost too, because it violates the default word order. So, if the original foot structure is already well formed, the YY fronting strategy is not used (e.g. [YY XX NN] is always avoided, because it already has a perfect metrical structure).

2.1.5.3 Problems of the metrical structure improvement proposal

Duanmu (2007) claims that when the original metrical structure is already good, YY fronting will not occur. However, Mandarin has [XX YY NN] compounds which can be alternatively realized with the word order [YY XX NN], examples are shown in below:

(25) a. [XX [YY NN]] [YY [XX NN]]
lü-se you-ji shi-pin you-ji lü-se shi-pin
green organic food organic green food
Gloss: ‘*organic food’*

b. [XX [YY NN]] [YY [XX NN]]
yu-le jian-shen huo-dong jian-shen yu-le huo-dong
entertainment fitness activity fitness entertainment activity
Gloss: ‘*entertainment fitness exercise*’

The compounds in (25) can be realized either as [XX YY NN] or [YY XX NN], there is no difference in meaning. As shown in (26), the foot structure of the compounds is already well formed in that every syllable is footed properly:

(26) [XX YY NN] (XX) (YY) (NN)
[YY XX NN] (YY) (XX) (NN)

The examples in (25) raise a question with regard to Danmu’s analysis: since every word in (25) is disyllabic and can form a foot, then what causes the word-order variation to happen?

It has been mentioned previously that the placement of multiple modifiers is semantically restricted. The sequence of multiple modifiers is believed to be arranged in accordance with the Semantic Proximity principle, which claims that the linguistic distance between expressions corresponds to the conceptual distance between them, i.e. the modifier which is semantically more coherent to the head noun is expected to stay closer to the head noun (Lu, 1998).

Therefore, on the basis of the inherence/relevance of the modifiers to the head noun, Lu (1998: 64) proposes the following order of in-compound modifiers in a NP for Chinese:

(27) age > size > color > origin > material > manner > function > N

Turning back to the order variation of the modifiers in (25a) and (25b), we will see semantics is at work. Consider the two modifiers in (25a) first. *lü-se* "green" and *you-ji* "organic" are actually synonymous. Both of them are used to specify food produced by organic farming. Semantically, they are equally relevant to the head noun, and consequently they are not internally ordered with regard to their distance to the head noun in the way shown in (27). Likewise, the modifiers in (25b) are not internally ordered either, because both of them denote the same thing: the function of the head noun. That being said, the word-order variations seen in (25) are semantically fine, because they do not violate the hierarchy stated in (27).

In the following example, we see a case where the placement of XX and YY is realized with a fixed word order.

(28)	[XX	[YY	NN]]	*[YY	[XX	NN]]
	da-xing	fa-dian	she-bei	fa-dian	da-xing	she-bei
	large-scale	power-generating	equipment	power-generating	large-scale	equipment

Gloss: '*large power-generating equipment'*

In (28), the XX modifier *da-xing* "large-scale" specifies the size of the head noun *she-bei* "equipment". The YY modifier, *Fa-dian* "power-generating", specifies the function of the head

noun. According to the word-order hierarchy in (27), YY has to be placed closest to the head noun. Thus, only the word order in [XX [YY NN]] is allowed.

The examples in (25) and (28) present a situation where metrical structure is irrelevant to the word-order variation, or the prohibition of word-order variation in the [X Y N] compounds. Instead, semantics is playing a role.

In addition, Lu (1989) proposes a different analysis from the perspective of the "chunk size" of modifiers. He suggests that Chinese has a preference to avoid "big-belly" structures, in which a long word occurs in the middle of short words. Instead, Chinese favors "small-belly" structures, in which long words occur at the beginning or the end of an expression. [X [YY N]] and [X [YY NN]] are big-belly structures; [YY [X N]] and [YY [X NN]] are small-belly structures. In line with Lu's proposal, the alternative word order is a remedy to the big-belly problem.

Duanmu (2007:194) cites Lu's big-belly examples which show that word-order variation is not necessarily triggered by the need to improve metrical structure:

(29) a. [gao-ceng [[gang-jin shui-ni] jian-zhu]]
multi-story steel-grid concrete building

b. [[gang-jin shui-ni] [gao-ceng jian-zhu]]
steel-grid concrete multi-story building

The structure in (29a) has a big belly: *gang-jin shui-ni* "steel-grid concrete (reinforced concrete)", a long word occurs between two short words *gao-ceng* "multi-story" and *jian-zhu* "building". Therefore, the middle long word can be alternatively fronted to derive the order in (29b). As seen above, there is no need to improve the metrical structures in (29), since every word in (29) is disyllabic and can be footed properly.

Hu & Lu (1988) argues that the big-belly structure is not favorable, because it places a short modifier at a distance from the head noun. When multiple modifiers occur on the same side of a head noun, it is better to place the short modifier close to the head noun so that modifiers can be more evenly distanced from the head noun. The avoidance of big-belly structure may cause violation of the word-order hierarchy in (27), which shows that the Semantic Proximity principle can be overridden by the structure configuration needs in Standard Mandarin.

As a response to the big-belly avoidance proposal, Duanmu (2007:195) suggests that (29b) is not derived from (29a), but from (30).

(30) [[gang-jin shui-ni] de [gao-ceng jian-zhu]]
steel-grid concrete *de* multi-story building
'*multi-story building of steel-grid concrete*'

Duanmu argues that the structure in (30) contains an underlying *de* and it can be changed to (29b) via *de*-omission. *De* is an attributive particle in Chinese, which conjoins a modifier and a noun. All [M *de* N] structures are phrases,[5] and *de* can be omitted occasionally. If Duanmu's suggestion is correct, then (29b) is derived from a phrase, and it therefore won't be a compelling example against Duanmu's analysis.

However, there is no valid method to test whether (29b) is derived from a phrase. The *de-omission* proposal seems a remedial suggestion. Duanmu (2007) argues strongly that [M *de* N] is a phrase but [M N] is a compound. Similarly, [M [M N]], [[M N] N], [[M N] [M N]], etc. are compounds. However, there is no clear evidence in the vast majority of cases that a compound is related to a phrase in that the compound is derived from a phrase by *de*-omission.

[5] "M" in the [M *de* N] structure indicates a modifier, and "N" a noun.

Instead, as a regular compounding process, a modifier can be added to a head noun to form a compound. Theoretically, it is not necessary to assume *de*-omission.

The challenge raised by examples in (25) and (29) is not convincingly addressed by Duanmu's metrical structure improvement proposal, leaving his analysis of the [X Y N] word-order variation theoretically inadequate.

In §2.1.5.4, the [V O N] word-order variation will be introduced, and details of Duanmu's analysis on the issue will be discussed.

2.1.5.4 [V O N] word-order variation: Duanmu's metrical analysis

The following data in (31) and (32) are taken from Duanmu (2007:159-198), which show all the possible word orders for the [V O N] or [V MO N] type of compounds. N represents a noun, V a verb, M a modifier and O an object. So, in a [V MO N] compound, M is the modifier of the object. The duplicated forms, such as VV and NN, mean a verb or a noun of disyllabic form. In [M O VV NN], the object and its modifier are monosyllabic, the verb and the head noun, disyllabic.

(31) Compounds with a monosyllabic verb

	Phrase Order	Compound word-order variation
a.	[[V O] *de* N]	[[V O] N]
b.	[[V O] *de* NN]	[[V O] NN]
c.	[[V OO] *de* N]	?[[V OO] N] ?[OO [V N]]
d.	[[V OO] *de* NN]	?[[V OO] NN] ?[OO [V NN]]
e.	[[V [M O]] *de* N]	?[[V [M O]] N] ?[[M O] [V N]]

f.	[[V [M O]] *de* NN]	?[[V [M O]] NN] ?[[MO][V NN]]
g.	[[V [M OO]] *de* N]	?[[V [M OO]] N] ?[[MOO][V N]]
h.	[[V [M OO]] *de* NN]	?[[V [M OO]] NN] ?[[M OO][V NN]]
i.	[[V [MM O]] *de* N]	[MM [[V O] N]] ?[[V [MM O]] N] ?[[MM O] [V NN]]
j.	[[V[MMO]] *de* NN]	[MM [[V O] NN]] ?[[V[MMO]] NN] ?[[MM O][V NN]]
k.	[[V [MM OO]] *de* N]	?[[V [MM OO]] N] ?[[MM OO][V N]]
l.	[[V [MM OO]] *de* NN]	?[[V [MM OO]] NN] ?[[MM OO][V NN]]

(32) Compounds with a disyllabic verb

	Phrase order	Compound word-order variation
a.	[[VV O] *de* N]	?[O[VV N]])
b.	[[VV O] *de* NN]	?[O[VV NN]])
c.	[[VV OO] *de* N]]	[OO [VV N]]
d.	[[VV OO] *de* NN]]	[OO [VV NN]]
e.	[[VV [MO]] *de* N]	[[M O][VV N]]
f.	[[VV [MO]] *de* NN]]	[[MO][VV NN]]
g.	[[VV [M OO]] *de* N]	[[M OO][VV N]]
h.	[[VV [M OO]] *de* NN]	[[M OO][VV NN]]
i.	[[VV [MM O]] *de* N]	[[MM O][VV N]]
j.	[[VV [MM O]] *de* NN]	[[MM O][VV NN]]

k. [[VV [MM OO]] *de* N] [[MM OO] [VV N]]

l. [[VV [MM OO]] *de* NN] [MM OO][VV NN]]

Listed in (31) and (32) are all the possible word orders for the [V (M)O N] type of compounds, given that the V, M, O, and N can be either a monosyllabic form or a disyllabic form. The corresponding [V(M)O *de* N] phrases are listed too, because they show clear syntactic relations between V, M, O and N: M modifies O; O is the object of V. The compound-internal [VMO], on the other hand, modifies N which is the logical subject of V. In the [V(M)O *de* N] phrases, the attributive particle *de* conjoins the preceding [V(M)O] phrase and the following head noun.

Following the convention of Duanmu (2007), a form with question mark can be either acceptable or unacceptable depending on a given expression. For instance, he states, "[[V OO] NN] is marginal for *qie luo-bo gong-ju* 'cut turnip tool (turnip-cutting tool)', but there is an existing expression *dong nao-jin ye-ye* 'move brain grandpa (the Wise Grandpa)', a columnist in a children's newspaper who answers science questions" (Duanmu, 2007: 198). A number of compounds listed above are always marginal no matter what the word order is, such as (31c~31h), (31k~31l) and (31a~31b). In such cases, the compounds are avoided and the corresponding phrases are used by native speakers. The compounds in (31a) and (31b) have only one word order. (31i) and (31j) have two marginal alternative word orders. All forms in (32) have no alternative word order. The bracketing in (31) and (32) is based on constituent analysis, and are not controversial.

Duanmu has made different attempts to account for the word-order variation of the [V O N] compounds. In his earlier work (Duanmu, 1997; Duanmu, 2000), Duanmu assumed phonologically triggered syntactic movement and multiple levels of stress assignment. Duanmu

(2007) acknowledges the earlier analyses are too complicated. He therefore has tried to offer a technically simpler analysis. In the most recent analysis, Duanmu continues to assume that foot must be binary and phrasal stress goes to the syntactic nonhead. Other assumptions of the proposal are shown below.

(33) [Verb Object] ([V O]) is a phrase

(34) a. *Internal Phrase: a compound should not contain an internal phrase.

b. Foot Shelter

c. *[V N]/[V NN]: a foot cannot cross a clause boundary.

The *Internal Phrase rule in (34a) prevents a [V O (M) N] compound from containing a [V O] phrase, such as [[VV OO] N] or [[VV O] N], unless the [V O] phrase is protected by Foot Shelter. For instance, a [[VO] N] compound does not violate (34a), because the compound-internal phrase [V O] is protected by Foot Shelter. *[V N]/[V NN] in (34c) rules out any compound that contain [V N] or [V NN]. Duanmu (2007) claims that, [V N] and [V NN] are not acceptable because their foot structures are (V N) and (V N) (N Ø) respectively, where the foot (V N) crosses a clause boundary. Rule (34c) needs more elaboration, we will be exposed with more details about the rule (34c) in §2.1.5.5.

The following data in (35) show how the rules listed in (34) can eliminate the disfavored word orders of [V (M)O N] compounds with a monosyllabic verb (Duanmu, 2007:200):

(35) Compounds with a monosyllabic verb

	Compounds	Comments
a.	[[V O] N]	[V O] protected by foot shelter
b.	[[V O] NN]	[V O] protected by foot shelter

c.	?[[V OO] N]	*Internal Phrase [V OO]
	?[OO [V N]]	*[V N]
d.	?[[V OO] NN]	*Internal Phrase [V OO]
	?[OO [V NN]]	*[V NN]
e.	?[[V [M O]] N]	*Internal Phrase [V [M O]]
	?[[M O][V N]]	*[V N]
f.	?[[V [M O]] NN]	*Internal Phrase [V [M O]]
	?[[M O][V NN]]	*[V NN]
g.	?[[V [M OO]] N]	*Internal Phrase [V [M OO]]
	?[[[M OO]][V N]]	*[V N]
h.	?[[V[M OO]] NN]	*Internal Phrase [V [M OO]]
	?[[M OO] [V NN]]	*[V NN]
i.	?[[V [MM O]] N]	*Internal Phrase [V [MM O]]
	[MM [[V O] N]]	[V O] protected by Foot Shelter
	?[[MM O] [V N]]	*[V N]
j.	?[[V [MM O] NN]	*Internal Phrase[V [MM O]]
	[MM [[VO] NN]]	[V O] protected by Foot Shelter
	?[[MM OO][V NN]]	*[V NN]
k.	?[[V [MM OO]] N]	*Internal Phrase [V [MM OO]]
	?[[MM OO][V N]]	*[V N]
l.	?[[V[MM OO]]NN]	*Internal Phrase [V [MM OO]]
	?[[MM OO] [V NN]]	*[V NN]

It can be seen above that the *[V N]/*[V NN], and the *Internal Phrase rule are robust in ruling out the marginal forms. The four acceptable word orders in (35) are all protected by Foot Shelter. Duanmu (2007: 201) shows the metrical structure of the four acceptable forms as follows:

(36)

Compounds with acceptable word order	Foot Structure
[[V O] N]	(SS) (SØ)
[[V O] NN]	(SS) (SS), (SS)S(SØ)
[MM [[V O] N]]	(SS) (SS) (SØ)
[MM [[V O] NN]]	(SS) (SS) (SS), (SS)(SS)S(SØ)

According to Duanmu, all the above foot structures are acceptable. Therefore, the word orders shown in (36) not only fulfill the requirements of the rules listed in (34), but also can be structured into acceptable metrical forms.

Duanmu's analysis of the word order of compounds containing a disyllabic verb is shown in (37):

(37) Compounds with a disyllabic verb

	Compound	Comments
a.	*[[VV O] N]	*Internal Phrase [VV O]
	?[O [VV N]]	Foot Binarity
b.	*[[VV O] NN]	*Internal Phrase [VV O]
	?[O [VV NN]]	Foot Binarity
c.	?[[VV OO] N]	*Internal Phrase [VV OO]
	[OO[VV N]]	
d.	?[[VV OO]NN]	*Internal Phrase [VV OO]
	[OO [VV NN]]	

e. ?[[VV [M O]]N] *Internal Phrase [VV [MO]]
[[MO][VV N]]

f. ?[[VV [M O]] NN] *Internal Phrase [VV [MO]]
[[MO][VV NN]]

g. ?[[VV [M OO]] N] *Internal Phrase [VV [M OO]]
[[M OO] [VV N]

h. ?[[VV [M OO]] NN] *Internal Phrase [VV [M OO]]
[[M OO][VV NN]]

i. ?[[VV [MM O]] N] *Internal Phrase [VV [MM O]]
[[MM O][VV N]]

j. ?[[VV [MM O]]NN] *Internal Phrase [VV [MM O]]
[[MM O][VV NN]]

k. ?[[VV [MM OO]]N] *Internal Phrase [VV [MM OO]]
[[MM OO][VV N]]

l. ?[[VV [MM OO]]NN] *Internal Phrase [VV[MM OO]]
[[MM OO][VV NN]]

Duanmu argues that the compounds containing a [VV O] phrase are completely ungrammatical, not only because they violate the *Internal Phrase rule, but also because O cannot form a binary foot: "In particular, the O cannot form a foot with the following N, because such a foot would cross a clause boundary" (Duanmu, 2007:202). Duanmu argues that the compound-internal [V O] phrase is like a relative clause. Therefore, there is a clause boundary between [V O] and N, details of which will be elaborated on in §2.1.5.5.

[O [VV N]] and [O [VV NN]] are unacceptable word orders because of the similar reason: O cannot be footed to realize the phrasal stress. O cannot form a foot with the following syllable, unless a structural change is made to delete the stress of VV. The rest of the problematic forms violates the *Internal Phrase rule.

Duanmu (2007) argues that the acceptable word orders shown in (37) have to be metrically well formed too, apart from being in line with the rules in (34):

(38)

Compounds with acceptable word order	Foot structure
[OO [VV N]]	(SS)(SS)(SØ)
[OO [VV NN]]	(SS)(SS)(SS), (SS)(SS)S(SØ)
[[M O] [VV N]]	(SS)(SS)(SØ)
[[M O] [VV NN]]	(SS)(SS)(SS), (SS)(SS)S(SØ)
[[M OO] [VV N]]	(SS)S(SS)(SØ)
[[M OO] [VV NN]]	(SS)S(SS)(SS), (SS)S(SS)S(SØ)
[[MM O] [VV N]]	(SS)S(SS)(SØ)
[[MM O] [VV NN]]	(SS)S(SS)(SS), (SS)S(SS)S(SØ)
[[MM OO][VV N]]	(SS)(SS)(SS)(SØ)
[[MM OO][VV NN]]	(SS)(SS)(SS)(SS), (SS)(SS)(SS)S(SØ)

In the compound-internal [M OO] constituent listed above, O is the syntactic head and head is not stressed at phrasal level. Hence, (M O) can form a foot where M get its stress, leaving the second syllable of the O(bject) extrametrical. In [M O] and [MM O], M can always properly footed. The compound-internal [VV N] and [VV NN] are not problematic either, because VV can easily form a foot to realize the phrasal stress. According to Duanmu's criteria, all the data above are not only syntactically acceptable, but also metrically well formed.

Duanmu's analysis of the word-order variation of [V (M)O N] is subject to a couple of problems. In particular, the *[V N] /[V NN] is an ad hoc rule. In the part to follow, we will discuss the problems in the foregoing analysis of Duanmu.

2.1.5.5 Problems of Duanmu's metrical analysis for [V O N] word-order variation

In Duanmu's analysis, compound-internal [V N] and [V NN] are considered ill formed. In the compound-internal [V N] or [V NN], N/ NN is the logical subject of the V(erb). The example below helps to illustrate the syntactic relations between V, O and N:

(39) [O O [V V N]]
xue - ya ce-liang yi
blood pressure measure machine
Gloss: '*blood-pressure monitor*'

In the above [V (M)O N] type of compound, OO (blood pressure) is the object of the VV (measure), and N (machine) is the logic subject of VV. Duanmu is curious about the lack of [V N] and [V NN] in the [V (M)O N] type of compound. He cannot simply rule out all compound internal verb-noun constituents, because [VV NN] and [VV N] are acceptable (see the well-formed word orders listed in (38)). Metrical considerations are not sufficient to provide an explanation in this regard. For instance, it is not clear why [V N] is not acceptable, considering that [V N] should be protected by Foot Shelter.

Duanmu proposes that the compound-internal [V N] and [V NN] are prohibited, because the foot structure derived there crosses a clause boundary:

(40) (SS) (SS)
[{$Ø_i$ [VV $Ø_j$]} NN_i]

(SS) (SØ)
[{$Ø_i$ [VV $Ø_j$]} N_i]

(S S)(SØ)
*[{$Ø_i$ [V $Ø_j$]} N N_i]

(S S)
*[{$Ø_i$ [V $Ø_j$]} N_i]

Duanmu assumes that a transitive verb always projects a subject position and an object position. In (40), V/VV projects an empty object $Ø_j$, and an empty subject $Ø_i$ which is co-indexed with N. The curly brackets indicate an inner clause. The structures shown in (40) are similar to a nominal with a relative clause. In the two ill-formed forms above, there is a foot formed across the clause boundary, which Duanmu "suspects is the reason why they are bad" (Duanmu, 2007: 206).

However, it is difficult to accept Duanmu's account as it stands. His account can be maintained only at the cost of adding ad hoc solutions as need arises. For instance, the *Internal Phrase rule has to be supplemented by Foot Shelter, so that the attested form [[V O] N] will not be eliminated. Similarly, the lack of compound-internal [V N] and [V NN] has to resort to an ad hoc rule *[V N]/*[V NN]: a foot cannot cross a clause boundary. Worse still, the rules can be in conflict with each other at times. For instance, Foot Shelter should protect [V N] whereas [V N] is not acceptable because it crosses word boundary.

The following is another set of examples where a special rule has to be added to fix a problem:

(41) a. Compounds with internal [VV OO+]

1. [[VV OO] N] zhi-zhao yao-yan zhe
   create rumor person
   '*people who create rumors*'

2. [[VV OO] N] guai-pian er-tong fan
   swindle child criminal
   '*criminals who swindle children*'

3. [[VV OO] N] zou-si du-pin zui
   smuggle drug crime
   '*crime of smuggling drugs*'

4. [[VV [OO OO]] N] xie-lou guo-jia ji-mi zui
   leak state secret crime
   '*crime of leaking state secrets*'

b. Compounds with internal [V OO+]

1. [[V OO] N] chou you-yan ji
   suck smoke machine
   '*kitchen exhaust fan*'

2. [[V OO] N] qu si-pi gao
   remove dead-skin cream
   '*exfoliating cream*'

3. [[V OO] N] kang yi-yu yao
   oppose depression medicine
   '*anti-depression medicine*'

4. [[V OOO] NN] fan fa-xi-si xuan-yan
   oppose Fascist declaration
   '*anti-Fascist declaration*'

Compound-internal [V O] phrases are banned from occurring in [V (M)O N] compounds if they are not protected by Foot Shelter. However, as shown in (41), there are many counter-examples to the *Internal Phrase rule in Standard Mandarin. Duanmu concludes that "such compounds seem to be limited to certain head nouns, in particular *zhe* 'the person who…', *zui* 'crime', and *fan* 'criminal'" (Duanmu, 2007: 209).

Duanmu (2007) argues that, cross-linguistically, compound-internal phrases are not uncommon. In English, for example, there are compounds like *nobody-cares attitude*, *take-it-or-leave-it offer*. Weise (1996) refers to such compounds as phrasal compounds. According to Weise, the internal phrase is not a true phrase but a quotation—an extra-linguistic symbol inserted into a compound. Weise argues that the quotation can be a phrase, a non-language sound, or a piece of another language.

Duanmu borrows Weise's idea, and suggests that the compound-internal [VV OO+] and [V OO+] phrases are quotations. However, the quotation analysis is not entirely convincing. A critical problem is that the quotation proposal fails to supply criteria to identify compound-internal quotations and compound-internal phrases. Such being the case, the truthfulness of the quotation proposal is difficult to be verified or disproved. It seems that, the quotation analysis is just prompted to fix the explanatory inadequacy of the rules in (34). Doing so, however, does affect the consistency and validity of Duanmu's proposal.

A more general theoretical problem of Duanmu's account for [V (M)O N] compounds is that it fails to bring out a further parallel between [V (M)O N] word-order variation and the characteristics of grammatical verb premodifiers in Chinese. In §2.1.6, we will present a grammatical account for the word-order issue.

2.1.6 A grammatical account of [V O N] word-order variation

2.1.6.1 Verb premodifiers in Standard Mandarin

It has long been noted that Chinese allows verbs to directly function as attributives (Shao, 1995; G. Wang, 1993; Zhang, 1989; Zhang, 1998). But, the verbs which can be used as attributives are confined to verbs with reduced verb properties (G. Wang, 1993; Wang, 2001; Zhang, 1989). Prototypical verbs cannot be directly used as attributives; when they are used attributively, they have to be followed by the attributive particle *de*.

As we already know, disyllabic verbs comprise a flexible class of verbs in that they can shift between the syntactic functions of verbs and nouns. Thus, disyllabic verbs often exhibit reduced verb properties and are not associated with an action meaning. For instance, it is sometimes preferable for disyllabic verbs to avoid being followed by aspect markers such as *zhe, le, guo*, as illustrated in (42):

(42) a. ? wo-men zai gong-zuo shang *lai-wang* le.
we at work on come-go *asp*
'*We have seen each other for business purpose*'.

b. wo-men zai gong-zuo shang *jing-xing* le *lai-wang*.
we at work on carry out *asp* come-go
'*We have seen each other for business purpose*'.

In (42a), when the disyllabic verb *lai-wang* "see each other" is made into a perfective by the aspect marker *le*, the sentence does not sound natural. One way to make the sentence more acceptable is to transfer the aspect marking to a dummy verb, such as *jing-xing* "carry out" as shown in (42b). It is apparent that *lai-wang* in (42b) is not predicating as a verb, but functioning as a verbal object.

Chen (1987) is an important and pioneering work on noun-verb class shift. Chen's study demonstrates that monosyllabic verbs show prototypical verb properties and, consequently, are prototypical members of verb class. For verbs, disyllabification is the necessary condition for noun-verb class shift to happen.

In Chinese, monosyllabic verbs are usually older. The majority of them entered the lexicon in the Old Chinese period (Wang, 2001).[6] Disyllabic ones were normally created later, along with the progression of disyllabification in Chinese. In this historical development, different syntactic functions were developed and came into being between mono- and disyllabic verbs. Monosyllabic verbs are frequently prototypical verbs. On the contrary, disyllabic verbs can generally be nominalized, and behave syntactically as nominals.

Thus, bare forms of monosyllabic verbs cannot be used attributively in front of nouns. Empirical evidence also proves that disyllabic verbs can, while monosyllabic verbs cannot, be used as attributive modifiers. For instance, Zhang (1989) tested 254 verbs randomly sampled from a medium-length novel. Zhang's test shows that all the 91 disyllabic verbs can be directly used before nouns as attributives. However, none of the 154 monosyllabic verbs can be used in the same way.

G. Wang (1993) has discussed the order of multiple modifiers before the head noun. The tests conducted in G. Wang (1993) reveal that when disyllabic verbs are used as attributive modifiers, they have to be placed closest to the head noun, and hence are nicknamed as the "close-to-skin attributive". Nothing may intervene between a verb premodifier and its head noun. Consider the following examples in (43):

[6] Old Chinese period dates from around 1200 BC to 3rd century BC.

(43) a. ling-dao fang-fa　　* ling-dao yi zhong fang-fa　　*ling-dao xin fang-fa
lead method　　lead one *cl* method　　lead new method
Gloss: '*leadership strategy*'

b. zhu-she qi-cai
inject instrument
'*injecting instrument*'
bo-li qi-cai
glass instrument
'*glass instrument*'

bo-li zhu-she qi-cai　　*zhu-she bo-li qi-cai
glass inject instrument　　inject glass instrument
Golss: '*glass injecting instrument*'

The examples in (43) are quoted from G. Wang (1993:71). The example (43a) demonstrates that a disyllabic verb cannot be separated from its head noun. The intervention of a classifier phrase or an adjective makes the V+N sequence ungrammatical. Example (43b) demonstrates that, when a disyllabic modifier co-occurs with another modifier, the disyllabic verb is the one that remains closer to the head noun.

It is not surprising that disyllabic verb premodifiers are placed closest to the head noun. Verb modifiers, by and large, signify an inherent property of the head noun, namely its function:

(44) a. jiu huang-se quan-mian y*un-dong* yi
old yellow cotton exercise suit　　(age> color> material > function)
'*old yellow cotton sport suit*'

b. ri-ben su-gang zi-dong *fu-yin* ji
Japanese plastic-steel automatic copy machine　　(origin>material>manner>function)
'*Japanese plastic and steel automatic photocopier*'

The attributive verb modifiers in (44) are italicized for easy identification. The relative order of the multiple premodifiers in (44) is consistent with the principle of Semantic Proximity

which contends that a semantic relation governs the placement of multiple premodifiers. According to this principle, modifiers signaling the function ascribed to the head noun should be placed in the innermost position.

As a summary of the foregoing discussion on verb premodifiers, there are two important characteristics that are worthy of attention:

a) Verb premodifiers exhibit reduced verb property, i.e. they are not prototypical verbs. Verbs that can be used as premodifiers are usually disyllabic, showing a fuzzy verb-noun word boundary. Monosyllabic verbs are usually prototypical verbs; they generally cannot be used as verb modifiers.

b) Verb premodifiers are placed closest to the head noun when there are multiple modifiers, because verb premodifiers semantically denote the function of the head noun they modify. According to the Semantic Proximity principle, the modifier signifying the function of the head noun should stay closest to the head noun.

2.1.6.2 [V O] phrases vs. VO verbs

Before a grammatical account can be presented for [V (M)O N] word-order variation, it is important to make a distinction between [V O] *phrases* and VO *verbs*. And, it is important to note that VO verbs grammatically behave as intransitive verbs. In addition, they can syntactically function as verb premodifiers.

The vast majority of the [1 1] type of VO expressions (both V and O are monosyllabic) are VO *verbs* rather than [V O] *phrases* (Li & Thompson, 1981). The VO verbs in general are used intransitively. They are not expected to take a direct object. Moreover, they are often lexicalized, associated with an idiomatic meaning. Therefore, the V and O constituents in a VO verb are more or less structurally inseparable:

(45) a. Inseparable [V O] verbs:

V	O		* V	*asp*	O
ge	ming		ge	*le*	ming
remove	mandate life		remove	*le*	mandate life
'*to revolt*'					

b. [V O] verbs with limited separability:

V	O		V	*asp*	O		* O,	V	asp
shang	feng		shang	*le*	feng		feng, shang	*le*	
hurt	wind		hurt	*le*	wind		wind, hurt	*le*	
'*to catch a cold*'			'*caught a cold*'						

The example in (45a), *ge-ming* "to revolt", is a lexicalized VO verb. The V and O constituents are inseparable in that nothing may intervene between the two constituents. (45b), *shang-feng* "to catch a cold", is also a lexicalized VO verb. The V and O in (45b) exhibit a limited degree of separability, which is evidenced by the perfective aspect marker *le* placed between the verb *shang* "hurt" and the noun *feng* "wind". However, (45b) does not allow the O(bject) to be placed in front of the V(erb) as a sentence-initial topic. The examples in (45) illustrate that the VO verbs have varied degree of structural solidness as a result of lexicalization.

In comparison to VO verbs, [V O] phrases are more loosely structured. Compare the examples in (45) with the [V O] phrase *wei-zao zheng-jian* "to forge a certificate" in (46):

(46) Separability tests for the constituents in [VV OO] phrases

VV	OO
wei-zhao	zheng-jian
forge	certificate
'*to forge a certificate*'	

a. VV and OO separated by an aspect marker:
Zhangsan *wei-zhao* le *zheng-jian*.
Zhangsan forge *asp* certificate
'*Zhangsan forged certificates.*'

b. VV and OO separated by a classifier phrase:
Zhangsan *wei-zhao* le san ben *zheng-jian*.
Zhangsan forge *asp* three *cl* certificate
'*Zhangsan forged three certificates.*'

c. VV and OO separated by a modifier:
Zhangsan *wei-zhao* le xueli *zheng-jian*.
Zhangsan forge *asp* educational certificate
'*Zhangsan forged educational certificate*.'

d. Placing OO in front of VV:
Xueli *zheng-jian* Zhangsan bu neng *wei-zhao*.
Educational certificate Zhangsan not can forge
'*The educational certificate, Zhangsan cannot forge.*'

The separability of the constituents in the [VV OO] phrase is tested in (46). The above syntactic operations show that not only regular interventions can occur in a [VV OO] phrase, but also the most drastic type of separation, where the object constituent is topicalized and placed in front of the verb. The contrast between the examples in (45) and (46) reveals that VO verbs are more of a monolithic unit than [V O] phrases. The former is grammatically an intransitive verb, whereas the latter is unarguably a phrase.

2.1.6.3 A grammatical analysis of [V O N] word-order variation

With the fundamental differences addressed for monosyllabic verbs vs. disyllabic verbs, and [VO] phrases vs. VO verbs, we are now at the point to give a grammatical explanation to the word-order variation of [V (M)O N] compounds.

A couple of comments are in order. First of all, Duanmu has been puzzled by the lack of [V N] and [V NN] in the [V (M)O N] compounds. It is not clear why Foot Shelter cannot protect compound-internal [V N], considering that Foot Shelter is a quite robust rule. It is not plausible to simply rule out all internal verb-noun compounds, because compound-internal [VV NN] and [VV N] are acceptable. Thus, the ad hoc *[V N]/[V NN] rule has to be postulated to address the problem.

However, this problem can be accounted for quite straightforwardly, if the grammatical characteristics of verb premodifiers are taken into consideration. As we already know, disyllabic verbs are generally not prototypical verbs. They can grammatically behave as nominals without overt morphological change, and directly functions as attributive modifiers. On the contrary, bare form of the monosyllabic verb cannot be used in the same way. Not surprisingly, [V N] or [V NN] are bad forms in contrast to the legitimate [VV N] and [VV NN]. In the latter case, the disyllabic verb can grammatically function as the modifier of the head noun. Furthermore, when used as verb premodifiers, disyllabic verbs have to be placed closest to the head noun in accordance with the Semantic Proximity principle, because verb premodifiers usually specify the function of the head noun. These two characteristics of verb premodifiers can explain the word order of the examples listed in (32). The examples in (32) are reproduced in (47):

(47) Compounds with a disyllabic verb

a. ?[O[VV N]]

b. ?[O[VV NN]]

c. [OO [VV N]]

d. [OO [VV NN]]

e. [[M O][VV N]]

f. [[MO][VV NN]]

g. [[M OO][VV N]]

h. [[M OO][VV NN]]

i. [[MM O][VV N]]

j. [[MM O][VV NN]]

k. [[MM OO] [VV N]]

l. [MM OO][VV NN]]

According to the judgment of Duanmu (2007), ?[O[VV N]] in (47a) and ?[O[VV NN]] in (47b) are the only two marginal forms in (47).[7] The marginal acceptability of (47a) and (47b) has to do with the word length well-formedness of [V O] phrases. Morphosyntactically, the O and VV constituents in ?[O[VV N]] and ?[O[VV NN]] form a [VV O] phrase. As we already know from the previous discussion on word-length variation, [2 1] is the avoided word-length combination for [V O] phrases. Rather, [2 1] is the standard template to form [V N] nominals. In (47a) and (47b), the VV and O constituents form a disfavored word length combination, which makes the two compounds marginally acceptable. As seen above, when O is not monosyllabic,[8] the corresponding compound is always well formed.

Next, let's turn to the word-order variation for compounds with monosyllabic verbs:

(48) Compounds with monosyllabic verb

a. [[V O] N]

b. [[V O] NN]

c. ?[[V OO] N]
?[OO [V N]]

[7] As a native speaker of Standard Mandarin, I share Duanmu's grammaticality judgment.
[8] Note that [MO] and [MM O] are not monosyllabic O(bject)s, since the modifier provides extra syllable(s).

d. ?[[V OO] NN]
?[OO [V NN]]

e. ?[[V [M O]] N]
?[[M O][V N]]

f. ?[[V [M O]] NN]
?[[M O][V NN]]

g. ?[[V [M OO]] N]
?[[[M OO]][V N]]

h. ?[[V[M OO]] NN]
?[[M OO] [V NN]]

i. ?[[V [MM O]] N]
[MM [[V O] N]]
?[[MM O] [V N]]

j. ?[[V [MM O] NN]
[MM [[VO] NN]]
?[[MM OO][V NN]]

k. ?[[V [MM OO]] N]
?[[MM OO][V N]]

l. ?[[V[MM OO]]NN]
?[[MM OO] [V NN]]

An inspection of the above compounds shows that all forms that contain [V N] or [V NN] are ruled out. This is readily comprehensible: the bare form of monosyllabic verbs cannot be modifiers. Secondly, all forms that contain a phrase, such as [V OO], [V [MO]], or [V [M OO]] are ruled out too, because [V O] phrases cannot be used to modify nouns without an attributive particle *de.*

Paradoxically, all forms with compound-internal [V O] ([1 1]) are acceptable. One may wonder why. A crucial difference between [V O] ([1 1]) and [V OO], [V [MO]], [V [M OO]] is, as already mentioned, [V O] ([1 1]) are actually VO *verbs*. As noted in the previous discussion, the [1 1] type of [V O] grammatically functions as intransitive verbs rather than phrases. Thus, a VO *verb* resembles a disyllabic verb, and can be used attributively before a nominal. Like typical verb premodifiers, they stay closest to the head noun.

Finally, let's consider the exceptional compounds Duanmu has talked about in (41). The examples presented in (41) all contain a [V O] phrase. To account for the violation of the *Internal Phrase rule, Duanmu suggests that the compound-internal [V O] phrases are quotations. As argued earlier, this suggestion is ad hoc. We will give an alternative analysis based on grammatical considerations. The examples presented in (41) are reproduced in (49):

(49) a. Compounds with internal [VV OO+]

1. [[VV OO] N] zhi-zhao yao-yan zhe
   create rumor person
   '*people who create rumors*'
2. [[VV OO] N] guai-pian er-tong fan
   swindle child criminal
   '*criminals who swindle children*'
3. [[VV OO] N] zou-si du-pin zui
   smuggle drug crime
   '*crime of smuggling drugs*'
4. [[VV [OO OO]] N] xie-lou guo-jia ji-mi zui
   leak state secret crime
   '*crime of leaking state secrets*'

b. Compounds with internal [V OO+]

1. [[V OO] N] chou you-yan ji
   suck smoke machine
   '*kitchen exhaust fan*'

2. [[V OO] N] qu si-pi gao
   remove dead-skin cream
   '*exfoliating cream*'

3. [[V OO] N] kang yi-yu yao
   oppose depression medicine
   '*anti-depression medicine*'

4. [[V OOO] NN] fan fa-xi-si xuan-yan
   oppose Fascist declaration
   '*anti-Fascist declaration*'

In line with the *Internal Phrase rule, the above compounds shouldn't have been allowed because [V O] phrases should not occur inside a compound. In order to account for these exceptional cases, Duanmu suggests that the compound-internal phrases are not real phrases but quotations. The problem of such a suggestion is that Duanmu does not provide a plausible criterion to differentiate compound-internal phrases from compound-internal quotations. It seems that this treatment allow for too much freedom in analytic practice.

The counter-examples listed in (49a) seem to be limited to certain head nouns, in particular *zhe* "the person who…", *zui* "crime", and *fan* "criminal". If we pursue a grammatical analysis, the forms in (49a) can be readily explained. *Zhe* in contemporary Chinese is not an independent word, but a suffix (Lü, 2001). *Zhe* is often attached to a preceding [V O] phrase to derive nominals with the meaning of "the person who VO", which is normally referred to as *zhe* construction (Wang, 2003). Similarly, *fan* and *zui* are used in the same way. If viewed from a

grammatical perspective, there is no need to give rise to the quotation proposal, and the examples in (49a) are not really exceptions but compounds derived from regular suffixation.

Similarly, the examples in (49b) are acceptable too, although they contain compound-internal [V O] phrases. All these compounds in (49b) are quite new; they came into being at around late 80s or early 90s of last century. The [[V OO] N] form for *chou you-yan ji* "kitchen exhaust fan" and *qu si-pi gao* "exfoliating cream" seems to be a compromise of the well-formed [[V O] N] form. *chou you-yan* ([V OO]), literally meaning "suck cooking smoke", cannot possibly be realized as *chou-yan* ([V O], "suck smoke") because *chou-yan* is already an intransitive VO verb, which means "to smoke (a cigarette)". Likewise, *qu si-pi* ([V OO], "remove dead skin cell") cannot be *qu-pi* ([V O], "remove skin"), since the latter is not the intended meaning.

The compound-internal [V OO] constituent in the rest of the examples of (49b) are not real phrases, because *kang*, and *fan* are actually used as affixes in the same way as "anti-" is used in English. The affix usage of *kang* and *fan* is introduced into contemporary Chinese through loan words from English. As a matter of fact, the above exceptions are limited in number. They cannot overshadow the fact that the standard and mainstream [V O N] compounds do not contain compound-internal [V O] phrases.

To summarize, the so called word-order issue and word-length issue can be accounted for by grammatical and semantic considerations. Thus, the word-order and word-length variations do not seem to serve as strong evidence for the claim that Mandarin has a generalized metrical system which covers both regular disyllabic words and neutral tone words. This view seriously challenges the dual trochee and binary foot suggestions Duanmu put forward, especially when stress of regular disyllabic words is supported neither perceptually nor phonetically.

2.1.7 Summary

The discussion offered so far has described Duanmu's metrical analysis for the word-length variation and word-order variation in Standard Mandarin. It is undeniable that Duanmu's metrical analysis is an innovative and noteworthy step forward for the explanation of the two long-noticed issues. Moreover, Duanmu's analysis is the most in-depth study from the stress-based perspective in this regard. However, it is easy to see the inadequacy of Duanmu's analysis. For instance, the analysis does not correctly predict the favored word-length combination for [A N] compounds, and it is technically inadequate to deal with exceptional cases without additional ad hoc rules. In contrast, the grammatical analysis naturally explains the various length patterns and word orders of nominal compounds and [V O] phrases in Mandarin, and therefore offers a more competitive alternative to the metrical analysis. This situation undermines Duanmu's argument for a generalized metrical system in Standard Mandarin.

2.2 PERCEPTION AND ACOUSTICS OF THE STRESS PATTERN IN REGULAR DISYLLABIC WORDS OF STANDARD MANDARIN

The preceding part in this chapter has presented the metrical view of Standard Mandarin and re-analyzed some stress-related issues form a grammatical perspective. The following section will present the results of previous acoustic studies centered on stress in regular disyllabic words of Standard Mandarin. Discussion on the experimental studies will follow, so that the acoustic findings can help to clarify some important facts about the stress in Standard Mandarin.

2.2.1 Acoustic studies

There are only a few studies that examine whether there is perceived contrastive stress in regular disyllabic words in Standard Mandarin. Lin, Yan & Sun (1984) provides an acoustic study in this regard. The test items in their study are produced by two native speakers who are asked to articulate 103 disyllabic words in citation form. The study finds that the sixteen listeners involved in the experiment tend to identify the test items as having the final stress. Listeners identify 91.2% of the 103 words articulated by the male speaker as having the final stress; 88.3% of the words articulated by the female speaker are identified as such. In the same study, the perception of stress is proved to be positively correlated to syllable duration, $r = 0.80$.

For the test items articulated by the male speaker, the average duration for the first syllable and the second syllable are 265 ms and 301 ms respectively. The average length of the first syllable produced by the female speaker is 317ms, and that of the second syllables is 346 ms.

Moreover, Lin, Yan & Sun (1984) noticed that the tone hosted by the second syllable is more ideally realized, in that the tone contour is close to the counterpart articulated in isolation. The authors believe the longer duration of the second syllable contributes to the idealized realization of the tone, which also enhances the perception of the final stress. Correlation between stress perception and articulation energy, or loudness is not found.

Lin *et al*.'s study has some merit. First, the test items are designed to cover a variety of morphosyntactic structures. Second, the data set covers all tonal combinations (neutral tone not included) for disyllabic words. However, the study has a noticeable limitation: all the test items are articulated in isolation. Although this study proves the dominance of final stress in regular

disyllabic words, it cannot clarify whether the final stress is an inherent property, or is caused by final lengthening effect commonly observed in words in citation form.

J. Wang & L. J. Wang (1993) investigates the internal duration distribution of disyllabic words and multisyllabic words embedded in frame sentence. Their findings are notable in that the opposite duration pattern is found in regular disyllabic words. The experiment tests 83 disyllabic words embedded in the same carrier sentence articulated by two native speakers. During the recording, the speakers are asked to utter the sentence in a way that no perceivable pause can be detected before and after the nested item. The data show, for both speakers, the first syllable is longer than the second syllable. For the male speaker, the averages of the normalized duration of the first and second syllables are 0.81 and 0.66 respectively.[9] For the female speaker, the normalized duration of the first syllables is 0.77, and that of the second syllable is 0.68. Trisyllabic words and quadrisyllabic words exhibit the same pattern: the initial syllable is longer than the following ones.

Cao (1992) investigates the temporal distribution in disyllabic words with a more comprehensive experiment design. In this study, when the regular disyllabic words are articulated in isolation, the first syllable is shorter than the second one. The result is reversed when they are articulated in connected speech, where the first syllable is longer than the second one. Notably, Cao's data covers disyllabic words consisting of a word-final neutral tone syllable. For these test items, regardless of whether they occur in isolation, or in a carrier sentence, the first regular syllable is always longer than the neutral tone syllable (recall that a neutral tone syllable never occurs word-initially).

[9] The normalized duration was calculated by dividing the average duration (ms) of an embedded syllable by the duration (ms) of the same syllable in citation form. Therefore the normalized duration represents a ratio between the duration of a syllable in connected speech and the duration of the same syllable in isolation.

Wang *et al*. (2003) is the most in-depth study of stress perception in Standard Mandarin to date. Their experiments tests 1,898 disyllabic words chosen from 300 utterances in Chinese Speech Corpus of Microsoft Research Asia. Therefore, the test items are originally in the context of continuous speech, although they are presented as isolated forms in the perception test. The items are further divided into four groups according to their position in the hierarchy of prosodic boundary: B1 is the prosodic word level, where no boundary (pause) can be perceived; B2, B3, and B4 are domains of prosodic compound, prosodic phrase and intonation phrase respectively, which are followed by an increasing length of prosodic boundary pause. A one-way ANOVA shows a highly significant main effect of the prosodic boundary (F (3, 1762) =73.84, p=0.000) for stress perception. Post hoc tests show that the stress pattern for words before B1 are significantly different from those before B2, B3, and B4 (p=0.000). No significant difference of stress pattern is found among words before B2, B3, and B4.

Data analysis shows that nearly 50% of the disyllabic words in B1 group are perceived by the 21 listeners as having the initial stress; however, only 10% of the words are perceived as having the final stress in the same environment. In the combined group of B2, B3, and B4, around 25% of the disyllabic words are rated as having the initial stress, whereas around 33% are rated as having the final stress.

Another finding in the study shows that the relative prominence of the first syllable in a disyllabic word has a significant positive relation with the F0 difference (r=0.583, p=0.000), and the duration difference (r=0.426, p=0.000) between the two syllables. This means that, when the first syllable is longer in duration, or higher in F0 than the second syllable, it is more likely to be perceived as having stress.

Notably, their analysis also reveals that the inherent tonal contour of a syllable has significant impact on its stress perception. A two-way ANOVA of the data shows a highly significant main effect of tone category on stress perception of the first syllable (F (3, 1879) =23.549, p=0.000). Similarly, main effect of tone category is also detected for the stress perception of the second syllable (F (4, 1879) = 224.937, p=0.000). For the first syllable, when it bears T1 (HH) and T4 (HL), it is more likely to be rated as being stressed. The second syllable receives more stress rating when it bears T1 (HH). This neatly illustrates that when a syllable is articulated with a high tone, it is more likely to be perceived as having stress by listeners. T3 (LLH) has a high tone feature; however, it is often realized as LL in natural speech regardless of its position in a word (Duanmu, 2007). T2 (LH) has a high feature too, but the high feature is not normally realized as high as that in T1 and T4 (Wang *et al.*, 2003).

Li *et al.* (2011) is the most recent study on stress perception in continuous speech in Standard Mandarin. The stimuli are disyllabic words extracted from corpus sentences, presented in isolation to annotators. Li *et al.* (2011) labels the test items as b0, b1, b2 and b3 according to their levels in the prosodic hierarchy. b0 is the syllable level, which is marked at the place between the syllables in a disyllabic word; b1, b2 and b3 are the prosodic word level, prosodic phrase level and intonation phrase level respectively. The subjects involved in the experiments are three experienced annotators. They are asked to rate the prominence level of each syllable in the test items.

Instead of examining the influence of each tone category on stress perception, Li *et al.* (2011) examines the impact from tone combinations. A pattern is said to be "compatible" when the end point of the first tone and the onset of the following tone share the same tone feature. When they do not have the same tone feature, the corresponding disyllabic words are classified

as a “conflicting” pattern. For instance, the combination of T1+T4 (HH + HL) is a compatible pattern; whereas, T4+T1 (HL + HH) is considered conflicting. Correlation coefficient analysis indicates that, in the case of the conflicting tone combination, normalized duration contributes a lot more to perceived stress than it does in the case of the compatible tone combination, regardless of the prosodic hierarchy of the word boundary. Li *et al.* (2011) explains that, in the conflicting pattern, tonal co-articulation effect lowers the average pitch of the word, which hinders the role of F0 in stress perception. For both types of tone combination, the correlation coefficient between the normalized F0 of a syllable and the stress rating of the syllable decreased, b1> b2> b3. According to Li *et al.* (2011), this result can be accounted for by the increasing final lengthening effect of the prosodic hierarchy, b1< b2< b3. As a result, syllable duration gradually overrules the contribution of F0 to perceived stress.

Notably, this study investigates, for the first time, the consistency of perceived stress of disyllabic words. In all the disyllabic test items, only 16.8% (6,894) receives identical stress rating from the three annotators. The perceived stress pattern for the rest of the test items (34,144, 83.2%) varies greatly. There are 492 words that exhibit consistent stress rating and occur three times as stimuli. Analysis of the 492 test items shows that: (a) the initial syllable is more likely to be perceived as having stress when it bears T1 and T4, a result which is compatible with Wang *et al.* (2003), and (b) 76% of the words occurring before b1 are rated as having the initial stress. For words occurring before b2 and b3, only 53% and 46% are rated as such, which means there is no stable stress pattern in the b2 and b3 prosodic contexts.

It is important to point out that even the same word can be rated as having different *stress patterns*, although it exhibits consistent *stress rating* at the same time. For instance, a word is considered to be with consistent stress rating in the following situation: a given word before b1

can be rated as having the initial stress by all the three annotators, which, however, won't prevent the same word from being rated as having the final stress by the same annotators when it occurs before b2 or b3. Therefore, in the study of Li *et al.* (2011), consistency of stress perception means the agreement of *stress rating* by the subjects in the same context, which, however, doesn't mean the word is realized with the same *stress pattern* in all the contexts. It seems the standard for being consistent in stress rating is not strict enough in this study, but it should be noted that a given word can be stress-rated contrarily even in the same prosodic environment by the annotators.[10]

2.2.2 Discussion on the previous experimental results

The results of the previous acoustic studies show that, stress perception in Mandarin disyllabic words varies considerably in different prosodic environments. When the final lengthening effect is not available, words are more likely to be perceived as having the initial stress. Otherwise, the second syllable receives a higher likelihood of being rated as stressed by listeners. Furthermore, since Chinese is a tone language, the inherent F0 contour plays a role in stress perception: a syllable with a high tone feature is more likely to be perceived as having stress. However, there is a trading effect between duration and F0: in environment where lengthening effect is more prominent, the influence from F0 on stress perception is reduced, and vice versa.

In this sense, the results of Lin, Yan & Sun (1984), J. Wang & L. J. Wang (1993) are compatible with Wang *et al.* (2003) and Li *et al.* (2011). The study of Lin, Yan & Sun (1984) examines the stress pattern of disyllabic words in citation form— a case where longer boundary

[10] As mentioned earlier, 83.2% (34144) of the test items do not have consistent rating from the annotators.

pause can be detected. Not surprisingly, their results indicate a dominant final stress pattern for disyllabic words, because final lengthening effect is playing an important role. The study of J. Wang & L. J. Wang (1993) is concerned with a case where boundary pause is small or undetectable. In this study, the first syllable is found to be realized with longer duration than the second one.

Interestingly, the trochaic system Duanmu suggested for Standard Mandarin embodies the findings of these studies. Recall that Duanmu argues that perceived final stress in Standard Mandarin is attributed to the final lengthening effect. He proposes that a disyllabic word can have final stress in a trochaic system when the word occurs in a prepausal position. In Duanmu's scheme, the final stress of a disyllabic word can be assigned by a trochee with an empty beat:

x

S (S Ø)

Overall, the previous experimental findings do not seem to support a general metrical system for Standard Mandarin. This is more clearly revealed when a comparison is made between Standard Mandarin and more typical stress languages. We are now turning to a comparison between Standard Mandarin and English.

2.2.3 A comparison of stress between Standard Mandarin and other stress languages

Based on these findings, an essential, yet critical question arises as to how to theoretically classify the so-called lexical stress in Standard Mandarin. More specifically, it is doubtful that there is such a thing that stress is an inherent property of the regular disyllabic words in Standard Mandarin. If stress is a lexical property, then it is not expected that the stress pattern of a

disyllabic word vary considerably in different prosodic environments. Similarly, the lack of consistent stress rating is not expected either.

In typical stress languages, such as English, there is also interplay between prominence and prosodic structure. For instance, different types of phonological prominence are identified, with each type corresponding to a different level in the prosodic hierarchy (Ladd, 2008). On the lowest level, there is the unstressed syllable, which is held to differ from a stressed syllable, the second level of prominence. Furthermore, the perceptual prominence in an utterance is linked to its intonation structure: each stressed syllable is a potential point of association for an accent. The latter is the third and the highest level of prominence. An accented syllable is perceived as more prominent than the unaccented ones. Therefore, it is important to differentiate "stress" from "accent". The stressed-unstressed distinction is at lexical level whereas the accented-unaccented distinction is at intonation level (Kunter, 2011). A stressed syllable can be accented or unaccented in its intonation structure. In the former case, the stressed syllable gets additional prominent enhancement. In the latter, the lexical stress won't be erased from the syllable, though it is not prominent at the intonation level. This observation appears to hold for most languages that prominence is constructed bottom-up in such a way that higher level prominence is overlaid on top of word-level stress (Ladd, 1996). Regardless of focus, higher level of prominence selects the primary stress in a word for promotion, except in cases of stress retraction to a secondary stressed syllable, or an unstressed syllable to avoid stress clash (Gordon, 2011). Hence, in a stress language, lexical stress is an inherent property, which won't be easily altered according to its position in longer utterances.

In comparison to English, the so-called stress in Mandarin is very different. For instance, as noted earlier, the stress pattern of a regular disyllabic word can vary considerably according to

its level in the prosodic hierarchy. Judging from what has been observed in the previous acoustic studies, the claimed stress of Mandarin is not the prominence at the word level in the sense that the prominence of a regular syllable is largely determined by its prosodic contexts. This conclusion is supported by evidence from speech synthesis. Based on the study of the naturalness of synthesized speech, Cao (2007) argues that the stress contrast between the neutral tone syllable and the regular syllable is unconditional, i.e. unconstrained by the prosodic environment. Phonologically, the contrast can distinguish lexical meaning. However, the stress contrast between regular syllables is not phonologically significant. Cao (2007:287) concludes that "regular disyllabic words exhibit comparative stress at sentential level rather than the contrastive stress at lexical level".

When compared to a typical stress system, such as English, the so-called stress in Standard Mandarin lacks perceptual and articulatory stability. Therefore, it is questionable to claim that such a "stress" is a lexical property. Just as stress and accent are distinguished as prominence at different levels, it is important to discern whether the stress in Mandarin is an intrinsic word-level prominence, or actually is the prominence at higher accentual structures.

2.3 SUMMARY

There is a clear consensus among Chinese linguists that the neutral tone syllable is lexically unstressed. It still remains controversial as to whether there is contrastive stress between regular syllables. Duanmu proposes a generalized stress view for Standard Mandarin which argues for the existence of lexical stress in both neutral tone words and regular disyllabic words. Based on

this view, Duanmu (2007) argues that the long-noticed word-length variation and word-order variation in Standard Mandarin are motivated by metrical effects.

This chapter has presented the stress-based analysis Duanmu suggested for these variations. Re-examination of Duanmu's analysis showed that: (a) the variations can be accounted for by grammatical and semantic considerations, and (b) Duanmu's metrical analysis is not convincing in many aspects. For instance, the proposed dual trochee, empty beat, and the rules posited for word-order variation contain ad hoc properties; the metrical analysis makes wrong prediction for word-length combinations; the metrical analysis cannot offer a convincing explanation to the existing counter-examples.

This chapter also presented the results of previous acoustic studies focused on stress in Standard Mandarin. The acoustic findings showed that both syllable duration and F0 can contribute to the perception of stress. However, stress perception in regular disyllabic words is heavily influence by prosodic environments, a feature which is not found in prototypical stress systems. Overall, when we reexamine the previous metical analyses and experimental studies, no conclusive evidence is found in support of the generalized stress view for Standard Mandarin.

3.0 TONE AND ACCENT IN NORTHERN WU DIALECTS

Beginning with this chapter we turn our attention to the discussion on how tone can interact with stress. With regard to word prosody, some languages, like Chinese and Thai, are tone languages; others, like English and Russian, have word stress. In reality, not all languages can neatly fit into either the tone or the stress category, prompting many researchers to introduce a hybrid type of system based on pitch-accent (Van der Hulst, 2011). This chapter is going to discuss whether some of the Northern Wu dialects can be classified into the pitch-accent system. Chapter 4 and Chapter 5 will focus on the discussion of tone-stress interaction, i.e. tone-conditioned stress placement and stress-triggered tone sandhi processes.

3.1 INTRODUCTION TO THE PITCH-ACCENT SYSTEM

3.1.1 Introduction to word prosodic typology

Over the past decades, phonologists have attempted to set up a principled characterization of differences in word-prosodic systems. A common tripartite word-prosodic typology recognizes three kinds of prosodic system, as in (1):

(1) a. Stress system: a language with word-level metrical structure, e.g. English

b. Tone system: a language with word-level pitch features, e.g. Mandarin

c. Pitch accent : a language in which pitch is the primary correlate of prominence and there are significant constraints on the pitch patterns at word-level, e.g. Japanese.

(Bybee *et al.*, 1998; Hyman, 2009)

The definitions provided in (1a) and (1b) are widely accepted for stress systems and tone systems. According to Hyman (2009: 215), "stress is a structural property in which syllables are metrically hierarchized as relatively strong vs. weak (however this contrast is realized phonetically), while tone is a featural property referring to contrastive relative pitch." The definition of pitch accent in (1c) is adopted from Bybee *et al*. (1998: 277), where the "pitch" in question is most commonly /H/. The "significant constraints" stated in (1c) are the properties most commonly associated with the distribution of tone in the pitch-accent system, which are enumerated in (2).

(2) In a pitch-accent system, tone may have the following properties:

a. *obligatory*: at least one must occur per domain (e.g. word)

b. *culminative*: at most one can occur per domain

c. *privative*: the underlying contrast is between presence vs. absence of the tone

d. *predictable*: assigned to positions by rule

e. *restricted*: occurring only in certain positions (e.g. stressed syllable)

f. *reducible*: subject to reduction, subordination (e.g. in defocusing)

(Hyman, 2011:230)

The above properties listed in (2a~2f) are typically associated with pitch-accent languages, which are sometimes labeled as “defective” tone systems. They are defective in the sense that the tone specification of such systems is sparse: the number of contrasts of tone is restricted; tone feature is not required for every TBU, and the presence of tone is restricted to specific positions.

However, in early works of pitch-accent theory, accent is not an intrinsic phonetic property, but an analytic notion that governs the tone derivations in some African languages. In cases where tone-vowel linking is not predictable, the unpredictable linkage is encoded by accenting the relevant vowel with a star in the underlying representation; the accented syllable in turn attracts a tone, typically H (Goldsmith, 1984). So, the presence of an accent usually entails the presence of a tone, as exemplified below:

(3) Underlying representation *nebo ‘*sky*’

High tone association *nebo
H

Default L tone assignment *nebo
H L

As seen above, the derivation of tones in Croatian (Babic, 1988) is subject to the locus of the diacritic accent marked by a star in the underlying representation. The accented syllable is linked with a H tone; the unaccented syllable is associated with a default L tone.

Later accentual analyses of tone have tried to interpret accent as metrical prominence. The metrical accent approach tried to establish connections between tonal system and metrical structure. For instance, in Chizigula (Kisseberth, 1991), the rightmost H tone in a word is

assigned to the penultimate syllable, which is a common position for metrical stress. Other kinds of formal analogies are also sought between tone and metrical systems, such as binary grouping of tones and the quantity-sensitive distribution of H tone (Odden, 1995). However, the metrical accent approach is not really successful. Inspection of metrical tone behavior claimed for some languages shows they are often wrong characterizations (Odden, 1985).

The validity of pitch-accent is questioned in Hyman (2006, 2009). Hyman notes that there is some vagueness and confusion with regard to word prosodic features in the pitch-accent framework. For instance, the most properties in (2) can be found in unambiguous tone systems. Secondly, the essential nature of pitch-accent is in some degree ambiguous and therefore hard to capture in the present tripartite distinction of word prosody. It seems that pitch-accent refers to a system which has both stress and tone. Stated this way, it would not be a third category, but a category where stress and tone co-occur (Hyman, 2006).

3.1.2 Terminological issues: stress, accent, pitch-accent

A well-known issue in prosodic studies is that the use of the terms, such as “stress” and “accent”, is somewhat problematic. The use of these two terms often leads to confusion when comparing different traditions or theories. On the one hand, stress and accent can be used as synonyms. On the other hand, the two terms have acquired their own specialized meanings. In the tradition of Abercrombie (1991) and Fox (2000), accent refers to an abstract lexical property of morphemes or words which marks the location of certain types of observable stress properties. The term stress is often used as a cover term for these observable phonetic properties, such as greater duration, greater intensity, etc. (Van der Hulst, 2014).

The distinctive use of stress and accent leads to the use of compound terms like *stress-accent* and *pitch-accent*. While stress-accent is realized with various phonetic correlates of prominence, the pitch-accent characteristically involves the exclusive use of pitch levels or pitch transitions (Beckman, 1986). This dichotomy is in correspondence with the more familiar terms such as stress and pitch-accent used in word prosody typology.

Although pitch properties were originally thought to be an important part of the set of phonetic exponents of stress (Fry, 1955), "it has been argued that this was often an illusion, arising from the fact that stressed syllables of words 'in focus' position function as anchors for intonational pitch movement" (Van de Hulst, 2014: 7). It has been noted that the descriptions of stress would often be based on the pronunciation of words in isolation. Consequently, the stressed syllable would be in focus and thus be associated with an intonational pitch movement.

More recent work has attempted to tease apart the word level stress and intonational pitch properties, and has shown that when stressed syllables are measured in out-of-focus position they often do not include pitch as a significant cue (Sluijter, 1995; Sluijter and van Heuven, 1996). Rather, stress in out-of-focus position is primarily cued by enhanced intensity, duration, fullness of articulation and spectral tilt (Gordon, 2014; Van de Hulst, 2014). This being so, Van de Hulst (2014:7) concludes that "stress-accent and pitch-accent are almost complementary in their use of phonetic exponents of accent".

There is also some confusion associated with the use of accent and pitch-accent. Apart from dealing with the notions of word-level prominence, the two terms are also used to refer to perceptible intonational prominence, which anchors on the stressed syllable to convey semantic or pragmatic focus in a certain position in a phrase or sentence. In the discussion on word level

stress and intonational prominence in §2.2.3, the term "accent" has been used in this sense, referring to prominence in intonational structures.

Hereafter, this dissertation will define accent as the abstract prominence property associated with certain positions at word level, which is then realized by phonetic means, leaving pitch-accent to be the word-level prominent property cued solely by the pitch patterns in the hybrid prosodic systems.

3.2 NEW CHONGMING

Chinese dialects are high in tone density (Gussenhoven, 2004) as opposed to the pitch-accent languages where tone specification can be sparse. High tone density in Chinese requires a tone for every syllable. Of course, neutral tone is an exception to this requirement, it being underlyingly atonic. However, tones in Chinese dialects can be restricted in their distribution and subject to obligatoriness and culminativity.

Chen (2000) has proposed that not all Chinese dialects are strictly tonal and that a tone-to-accent evolution is emerging in some Chinese dialects. The systems that Chen (2000) has identified as accentual are exclusively Northern Wu dialects as instantiated by Shanghai, New Chongming, Tangxi, and Danyang.[1] Northern Wu dialects are mainly spoken in East China, which roughly covers southern part of Jiangsu Province, Zhejiang Province and southeast Anhui Province.

[1] The languages spoken in different regions of China are more commonly referred to as "dialects". *Language Atlas of China* (1987) classifies the dialects spoken in China (including Taiwan) into ten genetic groups: Mandarin, Jin, Wu, Hui, Gan, Xiang, Min, Yue (Cantonese), Pinghua and Hakka.

Chen's analysis shows that some tonal systems in the Northern Wu group exhibit pitch-accent behaviors. Chen (2000) also tries to reduce accentual structure to metrical structure. However, as Bao (2003, 2004) insightfully points out, the Northern Wu dialects in question do not offer compelling evidence for accentual analysis. To start the discussion, let us first turn to the tonal system of New Chongming.

Hereafter, the geographical name of a city where a dialect is spoken will be use to refer to the dialect. For instance, Shanghai dialect will be referred to as Shanghai in the following sections.

3.2.1 The tone system

New Chongming is a Northern Wu dialect spoken in the Chongming Island under the jurisdiction of Shanghai. New Chongming has an eight-tone system, which is cross-classified by tone register,[2] tonal contour (rising, falling, or level) and syllable type.[3]

Table 3. Tone inventory of New Chongming

	High register	Low register
Smooth tone[4]	T1 H	T2 L
	T3 MH	T4 LM
	T5 HM	T6 ML
Checked tone[5]	T7 H	T8 L

[2] The Middle Chinese voiced/voiceless contrast in the syllable initial consonant leads to the splitting of tones into two registers: *yin* (high) register and *yang* (low) register. Generally, the *yin* register with a voiceless onset has a higher pitch value than the corresponding *yang* register.

[3] Syllables fall into two classes: checked syllables, namely syllables ended by an occlusive coda (/p/, /t/, /k/, or a reduced glottal stop /ʔ/); and, smooth syllables, which are either open syllables (CV) or syllables closed by a nasal stop. Checked tones occur exclusively with checked syllables; smooth tones occur with smooth syllables. Checked tones are remarkably produced with a short duration, which is usually half of the duration of a smooth tone.

[4] Refer to the part on smooth tone in 3.

[5] Refer to the part on checked tone in 3.

3.2.2 Accentual properties

Given the eight citation tones in New Chongming, there are sixty-four (8×8) combinatorial possibilities in a disyllabic word. In fact, only a small number of two tone sequences are attested. Roughly, the sandhi patterns of all disyllabic words can be reduced to just three basic types: [T.o], [o.T] and [T.T], where "T" represents a tone, and "o" represents an atonic syllable which will be assigned a default tone in the sandhi forms. The tones in the basic patterns may contrast in terms of level vs. contour, as in [H.o] vs. [MH.o] and [H.H] vs. [HM.MH]. They may also contrast in high vs. low register, as in [H.o] vs. [L.o] and [H.H] vs. [L.H]. Although the tonal combinations of disyllabic words are reduced to three basic patterns, the tonal distribution in disyllabic words is still unrestricted (Chen, 2000: 234).

When examining longer units, most evident in trisyllabic compounds, "there is exactly one tonic, that is to say, one prominent syllable, marked by its tone-bearing ability" (Chen, 2000: 232). Therefore, Chen (2000) argues that New Chongming has evolved into a pitch-accent system because some unambiguous accent-like tendencies have emerged in polysyllabic compounds in the dialect. The most prototypical pitch-accent characteristics in New Chongming enumerated in Chen (2000) are the following: culminativity, tonicity, leveling, and edgemostness, as shown below:

(4)

Culminativity: One and only one tone-carry syllable per trisyllabic and longer rhythmic unit.

Tonicity: Only accented syllables may carry tone; only tonics may be accented.

Leveling: Underlying tonal contrasts of register and contour are neutralized to a simple high-level tone (H) by the process of leveling.

Edgemostness: The tonic element is located either at the left or the right edge of a rhythmic unit.

The accentual properties enumerated in (4) are best manifested by the tonal patterns of trisyllabic compounds in New Chongming. The trisyllabic compounds may carry any one of the three tone patterns of (5); no other tone sandhi form is permissible:

(5) Well-formed trisyllabic tone patterns

a. H. T. T → *H. T. T → H. σ. σ

b. T. T. E → T. T. *E → σ. σ. H

c. T. T. T → T. *T. T → σ. H. σ

H= high tone *=accent marker E= even tone
T= any tone σ= syllable

The above derivations of pitch-accent are taken from Bao (2004:872), which is based on a study of New Chongming by Chen (2000: 232-233). In the above drivations, "σ" indicates an atonic syllable, which is later assigned a default tone. The derived tone patterns are shown in the rightmost positions in (5). The H distribution shown in (5) is interpreted as a set of ordered statements by Chen (2000):

I. If the first syllable of the trisyllabic compounds bears a H tone, then the first syllable is the tonic syllable and the derived tone pattern is [H.σ.σ], with the last two syllables acquiring default tones which is not relevant to the discussion.

II. When the last syllable bears an even tone, being H or L, the compound finally surfaces as [σ. σ. H], as indicated by (5b).

III. When both underlying H and even tones are not available, the syllable in the middle is the tonic syllable (5c).

Notably, the morphosyntax of the compounds does not play a role in the above tone/accent derivation. Undoubtedly, we see culminativity in trisyllablic compounds in New

Chongming, because trisyllabic compounds have at most one syllable which bears a H tone. This is considered a diagnostic criterion for a pitch-accent system (Hyman, 1977; Odden, 1985, 1995).

Moreover, the underlying tonal contrasts of tone contour and register in New Chongming are greatly reduced by a process known as leveling. New Chongming has eight underlying tones. The leveling process produces only three basic tone patterns for trisyllabic compounds as shown in (5), where a polysyllabic word "is organized around a single prominence peak. Only the prominent syllable is accentuated by a distinctive tone (H), while the recessive syllables are tonally suppressed, reduced to a non-distinct mid or low pitch, transcribed as [o]…" (Chen, 2000: 233).

In addition, Chen (2000) concludes that tonic syllable tends to occur at the leftmost or rightmost edge in polysyllabic expressions in New Chongming, which is referred to as edgemostness (Chen, 2000: 226). Edgemostness is recognized as a function to demarcate positionally restricted accent/stress. In some languages, H tones are located at or near the right edge, e.g. the obligatory word-final H of Choctaw, and the penultimate H of Somali (Hyman, 2009).

3.2.3 Tone-derived accent placement

Given these characteristics, it might seem that New Chongming is a pitch-accent language. However, as Bao (2004) insightfully points out, the accent in New Chongming is derived from the lexically-specified tones rather than a lexical property associated with certain phonological positions. The derivation of tone patterns in (5) clearly shows the locus of accent has to refer to the information of lexical tone at the first place (as shown by the intermediate stage of the

derivation). In typical pitch-accent languages, accent must be specified lexically. Tokyo Japanese is a case in point, where accent cannot be derived, but has to be specified underlyingly, as seen in the following examples cited from Hyman (2009):

(6) Pitch-accent in Tokyo Japanese

a. ma$^{\downarrow}$kura ga ‘*pillow’ + nom*.

b. koko$^{\downarrow}$ro ga ‘*heart’+nom*.

c. atama $^{\downarrow}$ga ‘*head’+ nom*.

d. sakana ga ‘*fish’+ nom*.

In (6), the accent is represented by the down arrow ($^{\downarrow}$), which indicates the place where there is a pitch drop form H to L. As seen above, there can be at most one such drop per word. Since word such as *sakana* “fish” does not exhibit a pitch drop, (6d) only meets the culminativity criterion, but not the obligatoriness criterion. As shown in (6), the placement of H tone is not predictable in that it is not derived by any phonological rules, but based on lexical accentual patterns. In Tokyo Japanese, the pitch-accent tone pattern has to be marked in an underlying domain.

In other pitch-accent languages, tone can be subject to positional restrictions, occurring on the penult or final syllable of a word or stem in languages like Somali, Wolaitta, and Beja (Downing, 2004; Hudson, 1973).

(7) Pitch accent in Wolaitta

a. zaré		‘*lizard*’
b. záre		‘*relative*’
c. zígir-etta	/zígir + ettá/	‘*to gossip*’
d. k’aar-ettá	/k’aar + ettá/	‘*to uncover*’

e. hayttá-tukke /hayttá+ tukké/ ‘*spicy coffee made from coffee leaves*’

Wolaitta allows only one H (indicated by an acute accent) per word, which is located on the penult or final syllable of a stem or word, as shown by the examples listed above (Downing, 2004:108). When a word is formed by combining two stems, only one stem tone can be kept. The preserved stem tone is *morphologically* predictable in cases like (7c), (7d) and (7e). As shown, when the first stem has a tone, the tone should be preserved. Or, the tone on the second stem is kept when the initial stem does not have a tone. However, it is not *phonologically* predictable where the tone will fall. Similar to Japanese, the location of H in a stem or word in Wolaitta has to be specified underlyingly.

Some pitch-accent languages display predictable pitch-accent patterns. For instance, the pitch-accent in Chi Mwi:ni is positionally determined. According to Kisseberth and Abasheikh (1974), accent in Chi Mwi:ni is assigned in the default case to the penult syllable of a word or phrase, or to the final syllable of some grammatical constructions. Compared to these languages, the pitch-accent in New Chongming is very different, which is derived from lexical tones.

Apart from the tone-based accent derivation, Chen (2000) also tries to reduce the accent placement in New Chongming to a metrically driven process. The subsequent analyses of the dialect present the basics of the metical accent approach. This approach attempts to suggest a connection between accent and metrical structure and quantity sensitivity. However, scrutiny of the details of the analyses shows the metrical approach is not successful.

3.2.4 Edgemostness and Weight-to-Accent Principle

According to Chen (2000), accents in New Chongming tend to be demarcative, i.e. signaling the beginning or the end of a phonological unit. Edgemostness is a feature commonly seen in stress

languages. The overwhelming majority of stress systems are head-terminal, i.e. with the stress on the first or the last syllable of a rhythmic unit (Hayes, 1995; Chen, 2000).

With reference to New Chongming, the edgemostness is shown through the accent derivation processes as listed in (5) which are reproduced in below.

(8) Accent placement rules:

a. H. T. T → *H. T. T → H. σ. σ

b. T. T. E → T. T. *E → σ. σ. H

c. T. T. T → T. *T. T → σ. H. σ

H= high tone *=accent marker E= even tone
T= any tone σ= syllable

Two generalizations emerge from the accent derivations shown above. First of all, underlying high level tone tends to attract accent (i.e., realized with H in the sandhi form). Secondly, H tends to appear at the margins: either the leftmost syllable or the rightmost syllable. In Chen (2000:246), the generalizations are expressed in terms of alignment constraints:

(9) AlignL Align accent with the leftmost H.

AlignR Align accent with the rightmost H.

In the OT framework, the appropriate ranking of the above alignment constraints with other markedness constraints (which will not be elaborated here, since they are not relevant to the present discussion) can choose the optimal form from the candidates. However, this ranking is incapable of handling trisyllabic compounds when they neither start with nor end in H, i.e., the case indicated by (8c). Instances are presented in below:

(10) a. [[za-huo] dian] ‘*grocery store*’

L^q MH MH base tone

o H o attested form “^{q}” = marker of a checked syllable

b. [[zhao-xiang] guan] '*photo studio*'

MH	MH	HM	base tone
o	H	o	attested form

The head-terminal alignment constraints in (9) will place the accent on the initial or the final syllable of the trisyllabic compounds in (10), which, however, is not the attested reading [o. H. o]. To obviate the problems noted above, Chen (2000) seeks recourse to the principle of prosodic weight. The basic idea of Weight-to-Stress Principle (Prince, 1990) is that syllables can be ordered on a weight scale of light, heavy, and super-heavy; and a heavier syllable is more likely to attract accent/stress than a lighter one.

Chen (2000) argues that the Weight-to-Stress Principle can be extended to a higher level of metrical organization, i.e. foot structure (Chen, 2000: 255). According to Chen, a binary foot is more likely to attract word-level stress than a degenerate, monosyllabic foot.

However, in the Weight-to-Stress analysis of New Chongming, Chen avoids using foot structure directly. Chen (2000: 257) states that "in effect, stress falls preferably on a longer constituent (in terms of syllable count)". In this analysis, Chen (2000) does not clearly define what is "a longer constituent" in terms of metrical structure. It seems that the "longer constituent" is the constituent that has more syllables in a compound.

The Weight-to-Stress Principle is modified as in (11) by Chen (2000: 257).

(11) **Weight-to-Accent**

Accent (x) } Accent (y) if x > y.

x } y	x is more harmonic than y
x > y	x is heavier than y (in mora, or syllable count)

Tableau 1. [zhao xiang] guan *photo studio*

MH. MH. HM base tone

[o. MH] + HM relevant input

	Tonicity	Align L/R	Weight-to-Accent
a. ☞ o.H.o			
b. ó.o.o	*		
c. o.o.H			*

ó = accented but toneless syllable

The placement of compound-internal accent is instantiated by the OT tableau taken from Chen (200: 259), as shown in Tableau 1. Candidate (b) with an accented but atonic initial syllable is eliminated by Tonicity (only tonics may be stressed). The constraint Weight-to-Accent requires an accent to be placed on the branching (disyllabic) constituent in the candidates. Inside the branching constituent, accent can be placed either on the initial or the final syllable, which won't violate Align L/R. Candidate (c) is removed by Weight-to-Accent, because the singleton (monosyllabic) constituent is incorrectly assigned the accent.

What is worth of attention in the above tableau is that the input to the OT evaluation is not the form with the underlying tones; rather, the input is the outcome of the disyllabic sandhi rules. Without referring to the result of disyllabic tone sandhi, the ranking of the constraints in Tableau 1 is not able to place the accent. To put it more explicitly, Weight-to-Accent only assigns the accent to the branching constituent (longer in terms of syllable count); it cannot, however, determine which syllable inside the branching element can be accented. Similarly, constraint Align L/R is not the decisive constraint either, since it does not really specify which edge is the accent carrier. The constraint Tonicity requires that only accented syllables may carry tone; only tonics may be accented. However, without referring to the tone sandhi result of the

disyllabic constituent,[6] Tonicity cannot possibly distinguish the tonic syllable from the atonic one, and cannot further locate the accent accordingly.

Although Chen (2000) claims to extend the Weight-to-Stress Principle to "a higher level of metrical organization", it is clearly shown in the OT account that Chen does not pursue an analysis based on metrical structure, such as the foot. And, by no means is accent metrically determined in New Chongming as seen in Tableau 1.

In a foot-based OT theory, FtType = Trochee and FtType = Iamb determine whether feet are trochaic or iambic in a language. In a grid-based OT theory, Align-L and Align-R refer directly to stress location, as instantiated by the following two examples:

Tableau 2. Foot-based OT theory

σσσσσσ	FtType=Trochee	Ft-Bin	Parse-Syl
☞(′σσ)(′σσ)(′σσ)			
(σ′σ)(σ′σ)(σ′σ)	*!**		

Tableau 3. Grid-based OT theory

σσσσσσ	Align-L	*Clash	*Lapse
☞ ′σσ ′σσ ′σσ	**,****		
σ′σ σ′σ σ′σ	*,***,**!***		

The above two quasi-tableaux manifest how stress is assigned by relevant constraints: FtType=Trochee determines that foot is left-headed in Tableau 2; similarly, Align-L in Tableau 3 requires stress falls on the leftmost syllable, which is further modified to an alternating stress pattern by *Clash: no adjacent stressed syllables and *Lapse: no adjacent unstressed syllables.

However, as it turns out, the metrical OT account of the accent pattern in New Chongming shows many ad hoc properties in comparison to the more classical metrical OT

[6] For more details, refer to the three basic tone sandhi patterns for disyllabic words presented in §3.2.2.

analyses. First of all, metrical structure (foot, for instance) is not clearly referred to in Chen's analysis. Secondly, tone is more essential than Alignment constraint in accent placement. As shown by Tableau 1, the input tone is still playing the key role in determining where the accent goes.

In addition, the proposed Weight-to-Accent principle is somewhat problematic. By Weight-to-Accent principle, Chen suggests a connection between accent and quantity sensitivity. This connection has long been noted: H tones resist being moved off of heavy syllables, and may move from light syllables to heavy syllables; also, heavy syllables tend to be accented and vice versa. Quantity-sensitive accent appears in a number of languages such as Yao, Kimatuumbi, and Chichewa (Odden, 1995). But, Chen's Weight-to-Accent is not the well-attested H and heavy syllable interaction, because Chen's principle suggests a connection between H distribution and syllable count. This claim is not coherent with the weight hierarchy system. In this system, weight is operative within a syllable, and the syllable weight is dependent on syllable structure. For instance, CVC is considered heavier than CV. The number of syllables in a word, however, is not relevant to metrical weight. The Chen-proposed attraction of H to a word with more syllables is not typically attested elsewhere in pitch-accent languages.

Besides, the claimed "edgemostness" of New Chongming also shows ad hoc properties. The typology of stress system can be roughly divided into two groups based on whether stress rhythmically falls on syllables at regularly spaced intervals within a word or whether it is fixed on a syllable at or near the edge of a word. If New Chongming exhibits accentual "edgemostness" as Chen (2000) claims, obviously, it should be classified as a fixed accent system. In the literature, there are five recognized docking sites for fixed stress/accent in a word: the first syllable, the last syllable, the penultimate (second-to-last) syllable, the antepenultimate

(third-to-last) syllable, and the peninitial (second) syllable (Gordon, 2011). In most languages, only one of these docking sites receives stress/accent. However, the trisyllabic compounds in New Chongming can be accented either on the leftmost syllable, or the rightmost syllable as illustrated by the rules reproduced in (12):

(12) Accent placement rules:

a. H. T. T → *H. T. T → H. σ. σ

b. T. T. E → T. T. *E → σ. σ. H

c. T. T. T → T. *T. T → σ. H. σ

H= high tone *=accent marker E= even tone
T= any tone σ= syllable

There are systems where stress can be alternatively placed at both edges. For instance, Eastern Cheremis stresses the rightmost heavy syllable, else (in words without heavy syllable) the initial syllable, while Komi Jazva stresses the leftmost heavy syllable, else the final syllable (Kager, 1995). But, as just described, the opposite-edge stress placement in these languages is conditioned by the absence of heavy syllable in a word. However, New Chongming is not the case.

Chen's OT analysis of New Chongming still supports that the underlying tone is the key factor in the H distribution. Apart from the culminativity of H tone, New Chongming does not convincingly present other pitch-accent properties. And, the claimed edgemostness is especially ambiguous in its essence, because it is not clear in the New Chongming case whether the H distribution is accent-driven, or is alternatively a true tone phenomenon. As Hyman (2009) points out, tones may also be restricted in their distribution. In Mayo [Uto-Aztecan; Mexico], a /H/ is restricted to occurring on either the first or second syllable of a word, and the placement of the /H/ is usually a lexical property of the roots (Hagberg, 2006).

3.2.5 Summary

There is insufficient evidence to classify New Chongming as a pitch-accent language, because "a strictly tonal analysis is equally plausible, bypassing the accentuation stage altogether" (Bao, 2004: 872). In the more prototypical pitch-accent systems, tonal patterns are derived from the underlying accentual patterns, where accent is a lexical property, as in Japanese. Alternatively, accent is determined metrically, as in Luganda (Bao, 2004). In addition, accent can be positionally determined, as in Chi Mwi:ni. As shown in the forgoing discussion, New Chongming cannot neatly fit into either case. Furthermore, the invoked accentual analyses of New Chongming show properties unusual to accentual systems.

With the neutralization of the contrast of the underlying tonal register and contour, it is more reasonable to assume that New Chongming has evolved into a reduced tone system rather than an accentual system.

Apart from New Chongming, other Northern Wu dialects as Shanghai and Tangxi are claimed to exhibit accentualism where the accent is positionally determined. In Shanghai and Tangxi, the leftmost syllable is singled out as the tonic syllable. Therefore these dialects are often identified as left-prominent (Qian, 1992; Chen, 2000). In the next section, we will reexamine the alleged accentualism in Shanghai.

3.3 SHANGHAI

3.3.1 The tone system

Shanghai is another Northern Wu dialect and is spoken in the city of Shanghai. There are five citation tones in Shanghai. The tones can be divided into two groups according to the rime: *checked tones*, which occur exclusively with rimes ended by a glottal stop, and *smooth tones* with rimes closed by a vowel or a nasal.

The duration of a checked tone is around half of the duration of a smooth tone (Zee and Maddieson, 1979; Zhu, 2005). Tones in Shanghai are also affected by onset voicing. Syllables beginning with a voiceless onset are realized with pitch higher than syllables with voiced onset, yielding a split of tone categories into a high and a low registers. But, if we ignore the register split resulted from onset voicing and the smooth-checked contrast in tonal types, the lexical tones in Shanghai can be phonologically represented as HL and LH.

Table 4. Citation tones in Shanghai

	T1	T2	T3	$T4^{q}$	$T5^{q}$
Contour	HL	LH	LH	LH^{q}	LH^{q}
Register	High	High	Low	High	Low
Tonal Type	Smooth	Smooth	Smooth	Checked	Checked
Tone Duration	Long	Long	Long	Short	Short

"^{q}" indicates a checked tone.

In Table 4, four of the five tones have a LH contour, however, these four LH contours are lexically distinctive since they contrast in terms of tone register or the tonal type. The smooth tones and checked tone are not only differentiated by the presence of the occlusive coda in the latter tone type, but also the inherent duration difference: the smooth tones are usually twice as

long as the checked tones. Therefore, the same contour realized with the different cross-classifications of registers and tonal types makes themselves distinct to native ears.

It is widely agreed that the accent in Shanghai is assigned to the leftmost syllable of a compound, and the lexical tone of the accented syllable is realized over the entire compound. The tones on non-initial syllables are deleted, being irrelevant to the tonal sandhi process (Yip, 1980; Chen, 2000; Bao, 2004).

(13) a. ʔø ʑin ‘*quiet*’ T1 HL (52) + T3 LH (13) → H (55) + L (21)

b. ts^{h}ən fən ‘*Spring Equinox*’ T1 HL (52) + T1 HL (52) → H (55) + L (21)

c. ɕiɔ nø ‘*little girl*’ T2 LH (34) + T3 LH (13) → L (33) + H (44)

d. tsɔ sᴇ ‘*blouse*’ T2 LH (34) + T1 HL (52) → L (33) + H (44)

The tone sandhi derivation of disyllabic compounds is illustrated by the examples in (13). The IPA symbols of the disyllabic compounds are given at the beginning of each line (13a~13d), which is followed by the English glosses. The tone categories of each syllable, and their corresponding tone contours are also given in each tone sandhi rules in (13).

As seen in (13), the output tone pattern is derivable via tone deletion and spreading. Tone deletion eliminates all but the leftmost tone. Tone spreading re-links the initial tone segments with the syllables in a left-to-right, one-to-one fashion. To better illustrate the relation between the derived compound melody and the leftmost tone, the tone values in Chao’s 5 digits system are given in the parenthesis next to the tonal contour in the examples of (13).

3.3.2 Phonetic evidence for iambic feet

Chen (2000) suggests a prominence-based analysis of tone sandhi in Shanghai. Bao (2003) also claims that Shanghai is undeniably an accentual system. For Chen, accent is a marker of prominence; he suggests a scenario in which accent is lexically associated with the first syllable of compounds in Shanghai. Chen (2000) sees a direct link between tonal stability and accentual prominence. In other words, Shanghai represents a left-prominent system, which is manifested by the tonal stability in the left boundary and the spreading of the leftmost tone to the post-accentual syllable. Therefore, the dominant (leftmost) syllable determines the melodic shape of the compound as a whole (Chen, 2000: 222).

In phonological analysis of Chinese dialects, the existence of prominence, including metrical head, is largely determined by the tonal preservation or the ability to resist tonal neutralization associated with certain positions (Yip, 1980, 1999; Wright, 1983; Duanmu, 1995, 1999; Chen, 2000, among others). In recent years, more phonetic studies of tone and stress became available, which helps to identify prominent positions through phonetic evidence. With the detailed phonetic data of syllable duration in disyllabic compounds, Zhu (1996, 2005) argues that Shanghai is not uniformly left-prominent. Based on his rich experimental data, Zhu (1996) pursues an explanation based on a metrical system in Shanghai. He recognizes the prominent syllable as the stressed syllable. In Zhu's analysis, the metrical system in Shanghai allows both left-headed and right-headed feet at the same lexical level.

3.3.2.1 Phonetic correlates of word-level stress

The earliest literature on the acoustic characteristics of stress saw the perceptual prominence primarily as loudness. This is evident, for instance, in Bloomfield (1933), who uses the terms

"stress", "loudness" and "intensity" synonymously (Kunter, 2011). However, such a view proved to be a rather restricted perspective on perceptual prominence by later experimental works such as that presented by Fry (1955, 1958). Fry found length and pitch play a much more important role in the perception of stress than previously thought. His experiments indicate that pitch, duration and intensity are all contributing to prominence, but pitch is capable of overruling contradictory length and loudness cues. Cross-linguistic phonetic studies have shown that the phonetic manifestations of stressed syllable typically associated with one or more of the following properties: raised fundamental frequency, increased loudness, greater duration, and different vowel qualities, e.g. Polish (Jassem *et al.*, 1968), Tai (Potisuk *et al.*, 1996), Pirahã (Everett, 1998), Chickasaw (Gordon, 2004), and Turkish (Levi, 2005).

However, the role of pitch in the perception of word-level stress is argued to be mistakenly overstated (Gordon, 2014; Van der Hulst, 2014). A major limitation of earlier studies is that they were based on words occurring in the focal position of an utterance, a position shown to conflate both word-level stress and phrase-level prominence.[7] Subsequent work in Sluijter (1995) and Slujiter and van Heuven (1996) shows that the two type of prominence are realized through different phonetic means and governed by different prosodic considerations. An important finding in the more recent work is that salient F0 characteristics contribute to perceived prominence primarily associated with syllables carrying phrasal stress (Gordon, 2014). Word level stress, instead, is cued by other properties such as duration, intensity, fullness of articulation and spectral tilt. Ortega-Llebaria *et al*. (2013: 186) also notices that "in sentences with flat F0 melodies such as post-focal sentences and reporting clauses, pitch no long exerted its effect as a cue to stress, causing duration cues to become dominant."

[7] For more details about the conflation of word-level stress and phrasal prominence in intonation structures, see the discussion in §2.2.3.

In a summary of previous studies on stress, Ortega-Llebaria & Prieto (2011:75-76) concludes that " fifty years of research on stress accent languages provide cumulative evidence that duration is a consistent correlate of stress at word level, while F0 excursions covary with stressed syllable at the sentence level. Yet, the relation of amplitude changes in the speech signal to word stress is controversial".

In addition, the perception of word-level stress is subject to language-specific constraints. As Chinese is a tone language, lexical tones in Chinese covary with stress. For instance, Xu (2006) finds that a regular syllable can mostly complete the lexical full-tone target, whereas the F0 of a neutral tone syllable varies greatly according to the preceding tone. Moreover, studies on the acoustics of neutral tone syllable show that the perception of word stress in Standard Mandarin is primarily cued by longer syllable duration, while intensity is proved to be not an important phonetic exponent (Lin & Yan, 1980, 1990). Such being the case, the acoustic studies of word level stress in Chinese have been overwhelmingly centered on the durational characteristics.

3.3.2.2 Phonetic data

In the discussion of the metrical prominence of Shanghai, Zhu (1996, 2005) thoroughly examines the duration of tones in disyllabic sequences and in citation forms, and makes comparisons between these durational data. Based on these data, Zhu infers that Shanghai is not a uniformly left-prominent system. Zhu's study show that when T5, the low-registered checked tone, occurs word initially in a disyllabic sequence, the second syllable is more prominent than the initial one. This view is also shared by Shen *et al*. (1987).

The data in Table 5 show the duration of citation tones of Shanghai. These data are based on the recording of citation tones from four participants (three male, one female). All of the

participants are native speakers of Shanghai. All of them were born and raised in Shanghai. Each participant was asked to read the monosyllabic words for six times. The monosyllabic words were not embedded in carrier sentence when recording was conducted. However, extra "dummy words" were added before and after these words to avoid unnatural "page intonation" (Zhu, 2005: 31) that commonly occurs at the beginning or the end of their reading.

Table 5. The average duration of citation tones in Shanghai

T1 (HL, H register)	**T2** (LH, H register)	**T3**(LH, L register)	**T4**^{q} (LH, H register)	**T5**^{q} (LH, L register)
212 ms	240 ms	238 ms	71 ms	92 ms

Checked tones are suffixed by the symbol "q"
H register = high register
L register = low register

Since the monosyllabic words were not embedded in carrier sentence when recorded, it is hard to confidently determine the beginning of a syllable when it is initiated by a consonant. Thus, the duration of the initial consonant of each word is not measured. Data in Table 5 are the average rime durations (in millisecond) of the tone categories. However, it is justified to call them tone durations, as tone is phonetically realized by the rime part of a syllable.

The durations of tones in disyllabic sequence are presented in Table 6, which are based on the articulation of the same participants as the citation tone task. Similarly, the disyllabic words prepared for recording are not framed in carrier sentence, but dummy words are added before and after these disyllabic test items.

Table 6. The average duration of tones in disyllabic sequences in Shanghai

	T1 + Tx	T2 + Tx	T3 + Tx	T4^{q} + Tx	T5^{q} + Tx	Mean (final tone)
Tx +T1/ T2	158ms 73ms	174ms 121ms	172ms 122ms	55ms 139ms	57ms 169ms	125ms
Tx +T3	187ms 81ms	191ms 135ms	202ms 129ms	77ms 161ms	65ms 193ms	140ms
Tx +T4^{q}	162ms 39ms	174ms 51ms	177ms 55ms	59ms 54ms	55ms 63ms	52ms
Tx +T5^{q}	180ms 46ms	202ms 62ms	211ms 60ms	78ms 66ms	68ms 85ms	64ms
Mean (initial tone)	172ms	185ms	190ms	67ms	61ms	

In Table 6, the four rows of the leftmost column correspond to the tones of the second syllable in a disyllabic sequence. For instance, Tx + T1 indicates a disyllabic sequence with T1 in the final position, and a random tone in the initial position. Likewise, the tones of the first syllable in disyllabic sequence are given across the top. The duration of tones are specified in the cells where the relevant rows and columns intersect. The bottom row shows the mean duration of the initial tone averaged across the second tone. In the rightmost column are the durations of the final tone averaged across the initial tone.

As indicated by the second row of the leftmost column, a complete merger occurs to T1 and T2 when they are the second tone in a disyllabic sequence. This is a result of the tone spreading of the leftmost tone, which determines the melodic shape of the compound as a whole. Therefore, the underlying tone of the second syllable is no longer contrastive. However, merger only happens to T1 and T2 because both of them are smooth tones and they are in the same (high) register. Recall that tone register splitting in Chinese is conditioned by the initial consonant: high register tones with voiceless consonants have higher pitch value than the low register tones with voiced consonants. Thus, T1 and T2 in the second syllable are no longer contrasted by tone, tonal type and tone register. However, other tones in the same position are not subject to complete merger, because there is still contrast either in register (voicing contrast) or in tonal type (smooth/ checked tone contrast).

Data in Table 6 show that tones occurring after T5 have longer duration than those occurring after T1, T2, T3, and T4. For instance, when T3 follows $T5^{q}$, T3 measures 193 ms in duration; T3 measures 81ms, 135 ms, 129 ms and 161 ms respectively when it follows other tones. This fact is better revealed when the duration of the second tone is averaged in the way presented below:

Table 7. The average duration of the second tone in disyllabic sequences in Shanghai

Tone sequence / Duration of the second tone	T1+ **Tx**	T2 + **Tx**	T3 + **Tx**	T4^{q} + **Tx**	T5^{q} + **Tx**
Smooth tone	77ms	128 ms	126 ms	150 ms	181 ms
Checked tone	43ms	57 ms	58 ms	60 ms	74 ms

"^{q}" indicates a checked tone

In the cells of Table 7, one can find the average duration of the second tone for both tonal types. It is easy to see that tones after T5^{q} have the longest duration in both tonal types. According to Zhu (2005: 263), the average duration of smooth tones (181 ms) and checked tones (74 ms) occurring after T5^{q} are significantly longer than the counterparts after other tones (P < 0.01).

However, T5 in the initial position of a disyllabic sequence (T5^{q} + Tx) is not longer than other tones in the same position. On the contrary, initial T5^{q} exhibits a shorter duration. This observation is evidenced by the data shown in Table 8.

Table 8. The normalized duration of word-initial tone categories in Shanghai

	T1 + Tx	**T2** + Tx	**T3** + Tx	**T4^{q}** + Tx	**T5^{q}** + Tx
In word-initial position	172 ms	185 ms	190 ms	67 ms	61 ms
In citation form	212 ms	240 ms	238 ms	71 ms	92 ms
Normalized duration	0.81	0.77	0.80	0.94	0.66

"^{q}" indicates a checked tone

Since the durations of the five tone categories are inherently very different, the duration of each tone category is normalized in order to filter out the influence caused by tone categories. The normalization is conducted in a way shown in (14):

(14) $$\textit{Normalized tone duration} = \frac{\textit{Duration of a tone in word-initial position}}{\textit{Duration of a tone in citation form}}$$

In the above formula, the duration of each tone category is standardized by bringing the average duration of a word-initial tone into proportion with the duration of the same tone in

citation form. The ratio is presented in the bottom row of Table 8. For instance, the word-initial T5 is only 66% of its citation form whereas other initial tones are at least 77% of the citation counterparts. This reveals that $T5^{q}$ in the initial position of a disyllabic word (i.e. $T5^{q}$ + Tx) is shorter than other initial tones.

What has been observed so far suggests that $T5^{q}$ + Tx behaves in a different manner from other tone combinations: tones occurring after $T5^{q}$ are systematically longer; however, domain initial $T5^{q}$ is shorter than other tones in the same context. This observation will become even more apparent, if a comparison is made between two identical tones in a sequence.

Table 9. A comparison of two identical tones occurring in a sequence in Shanghai

	T1	T2	T3	$T4^{q}$	$T5^{q}$
In word-initial position	158 ms	174 ms	202 ms	59 ms	61 ms
In word-final position	73 ms	121 ms	129 ms	54 ms	85 ms
Ratio	0.46	0.70	0.64	0.92	1.40

"^{q}" indicates a checked tone

Table 9 compares the duration of identical tones that occur next to each other, for instance T1 + T1. Again, it is hard to group $T5^{q}$ + $T5^{q}$ with other tone sequences. As manifested by the data, in a $T5^{q}$ + $T5^{q}$ sequence, the second tone is 1.40 times longer than the first tone. Other combinations in Table 9 demonstrate the opposite pattern, where the first tone is always longer than the second tone.

Based on these findings, Zhu (1996, 2005) concludes that $T5^{q}$+ Tx form an iambic foot ([w S]), whereas T1, T2, T3, and $T4^{q}$ + Tx are trochaic ([S w]). Shen *et al*. (1987) also concludes that $T5^{q}$ + Tx is an iambic foot as opposed to the trochaic foot in other tone combinations. Interestingly, Shen *et al*. (1987) finds that the stress pattern can affect the voicing of syllable-initial stop consonants. Shanghai has been traditionally described as having a tripartite division in initial stops: voiceless aspirated, voiceless unaspirated and voiced (Chao, 1928). However, it

has been noted that the voiceless/voiced contrast is not typically found when the consonants are in word-initial position. In fact, all the consonants in this position are voiceless. But, when the consonants are in word-internal position, for instance, when they occur in the second syllable in a compound, Shen *et al*. (1987: 204) contends that the voicing contrast is present and there is a clear distinction between the voiced consonants and the voiceless ones. Surprisingly, perhaps, they find that in $T5^{q}$ + Tx, word-internal stops do not show any voicing no matter whether they are supposed to be voiced or not. They attribute this exceptional behavior to the effect of the iambic stress in $T5^{q}$ + Tx.

Apart from the exceptional duration pattern and stress-consonant interaction reported above, the tone sandhi pattern of $T5^{q}$ + Tx is also different from other tone sequences. Zee and Maddieson (1979), Shen *et al*. (1987) and Zhu (2005) all agree that when $T5^{q}$ is in the word-initial position of a disyllabic word, its entire tone is migrated to the second syllable. The first syllable is later assigned a default low tone.

Table 10. Shanghai tone sandhi patterns in disyllabic domain

Tone sandhi patterns		Note:
T1 + Tx	HL + Tx → H + L	"q" indicates a checked tone
T2 + Tx	LH + Tx → L + H	
T3 + Tx	LH + Tx → L + H	
$T4^{q}$ + Tx	LH^{q} + Tx → L + H	
$T5^{q}$ + Tx	**LH^{q}** + Tx → L + **LH**	

As shown in Table 10, $T5^{q}$+ Tx is different from other sequences in that it does not present a one-to-one type of rightward tone spreading. Rather, the first tone shifts leftwards to the word-final position. This unique pattern shows that the syllable after $T5^{q}$ can attract and capture the neighboring tone, which is considered an indication of phonological prominence although it is not yet the lexical-tone preservation commonly found in the accented positions.

Zhu (2005) does not pursue an accent-based analysis for Shanghai. Instead, Zhu directly proposes a metrical system. In this system, the stress assignment is tone-dependent. While all other tone combinations are trochaic, $T5^{q}$ + Tx is iambic, revealing that $T5^{q}$ is avoided as the carrier of stress. Shanghai is not the only dialect instantiating the interplay of tone and stress. Similar interaction is also found elsewhere in the Northern Wu group. Discussion on the tone-sensitive stress placement is going to be continued in the next chapter.

3.4 SUMMARY

The discussion in this chapter has highlighted the problems of the previous accentual analysis of Northern Wu dialects. Re-examination of previous analyses of New Chongming reveals that the claimed pitch-accent system exhibits a number of ad hoc properties. As stated earlier, the accent placement is not lexically, metrically or positionally determined in New Chongming. Tone still plays the primary role in the distribution of H tone.

Moreover, it is theoretically unnecessary to pursue an accentual analysis in New Chongming. Since tones in Chinese must be specified underlyingly, accent as an intermediate tier between vowel and tone is not needed for predicting the tone-vowel association. Therefore, theoretically and empirically, there is no compelling evidence to pursue an accentual analysis.

The accentual analysis of Shanghai assumes that the accent is determined positionally (leftmost position) in words or compounds. However, phonetic evidence does not support this view. Phonetic evidence suggests that Shanghai dialect has both left prominence and right prominence. The two prominence patterns can be straightforwardly represented as trochaic stress and iambic stress. Compared to the proposed accentual system, positing a metrical system with

both trochaic stress and iambic stress is, first of all, in line with the phonetic facts. Secondly, there is no apparent reason to disfavor such an attempt.

4.0 PERIPHERAL PROMINENCE IN CHINESE

4.1 INTRODUCTION

The distinction between strong and weak positions has long been observed in phonological studies. Different behavior in prominent and weak positions was first identified in the context of positional neutralization (Smith, 2002). Prominent positions license a fuller rang of contrasts than do non-prominent sites in languages which make such a distinction (Beckman, 1995; Steriade, 1995; Zoll, 1997 among others). For instance, typologically marked structure is tolerated in strong positions but neutralized in weak positions (Smith, 2002).

The early investigations of positional neutralization recognize two positions of "phonological peak" in a word that is able to resist neutralization: the location of stress or accent, or word edges as opposed to medial positions (Smith, 2002; Trubetzkoy, 1939). Subsequent work on positional neutralization has increased the inventory of prominent positions (Smith, 2002). However, the inventory of strong positions often varies considerably from one study to another. Steriade (1993) recognizes initial syllables, final syllables and stressed syllables, in addition to released consonants and long vowels. McCarthy & Prince (1995) proposes that the morphological root (as opposed to affix) is a strong position. Smith (1998, 2001) suggests that nouns can also be considered a prominent position.

The view on the typology of prominence also differs among linguists. Zoll (1997) recognizes a three-way typology of prominence: *organic prominence*, which includes long vowels, low vowels, etc.; *peripheral prominence* associated with initial or final syllable in a domain; and *metrical prominence*, which includes the penult, ante-penultimate positions in a prosodic word in addition to the head in binary feet.

Beckman (1998) divides the inventory of strong positions into *phonetically prominent positions* and *psycholinguistically strong positions*. Phonetically strong positions are those that are realized with prominent acoustic properties. For instance, stressed syllables can be phonetically realized with longer duration and stronger amplitude than other syllables (Lehiste, 1970). Smith (2002) follows the division of Beckman. In addition, they both recognize initial syllable as a psycholinguistically strong position. Smith argues that there are many languages in which positional neutralization is resisted in the initial syllable. Extragrammatical evidence also justifies the special status of initial syllables: the initial syllable is important in speech perception and language processing, and argued to be specifically important in early-stage word recognition (Beckman 1997, 1998; Smith, 2002).

All of the above-mentioned work on prominence typology recognizes multiple prominences, though the type and inventory vary. Li (2003) suggests that multiple prominent positions have to be admitted into the grammar of tone mapping in Chinese; otherwise, some well-known facts of tone sandhi cannot be captured. Following the proposal of prominence-driven tone mapping in Zoll (1997), Li (2003) attempts to include both peripheral prominence and metrical prominence in the phonology of some Chinese dialects.

The recognition of multiple prominences in Chinese is primarily based on the tone sandhi patterns observed in Zhenhai, which is a Northern Wu dialect. The tone retention and migration

in Zhenhai are argued to emerge from the interaction between stress (metrical prominence) and initial prominence (peripheral prominence) (Li, 2003). However, I will argue against this point of view.

4.2 ZHENHAI: DUAL PROMINENCES

4.2.1 Citation tones in Zhenhai

Zhenhai is a Northern Wu dialect spoken in northeast of Zhejiang Province. With sufficiently detailed phonetic data obtained from experiments, Rose (1990) presents an acoustic analysis of citation tones and disyllabic tone sandhi patterns in Zhenhai. In accordance with the native speakers' judgment, Rose (1990) also makes a description of the stress patterns of disyllabic words. Rose notices that native speakers of Zhenhai are well aware of the stress patterns in the dialect, for instance, they can easily tell the stress differences of lexical items between Zhenhai, and Fenghua, a neighboring variety in the Northern Wu dialect group (Rose, 1990:7).

The distribution of stress in disyllabic utterances in Zhenhai is found to be conditioned by both syntactic factors and lexical tones. For instance, unstressed T2 and $T5^{q}$ cannot occur in word-initial position in a disyllabic tone sandhi domain, although they can occur freely in a verb-object construction and auxiliary verb-functive verb construction (Rose, 1990:7). The interplay between stress and tone is not surprising. In Shanghai, similarly, we see tone-dependent stress placement in disyllabic words, for example, stressed $T5^{q}$ (low register checked tone) does not occur in word-initial position.

Table 11. Citation tones in Zhenhai

Tone Type	Register	Tone	Tone Value	
smooth tone (long)	high register	1	441	HL
		2	323	MH
	low register	3	231	ML
		4	213	LM
checked tone (short)	high register	5	5	H^q
	low register	6	23	LM^q

"^{q}" indicates a checked tone

As shown in the above table, there is a dichotomy between smooth tones and checked tones in Zhenhai. The checked tones are closed by a glottal stop, and their duration is phonetically short. Moreover, Zhenhai still retains the register distinction.

4.2.2 Tone sandhi and dual prominences in Zhenhai

Disyllabic expressions in Zhenhai show two metrical patterns: iambic [w S] and trochaic [S w]. As mentioned in §4.2.1, the stress pattern in Zhenhai is related to tone identity, which is evidenced by the restrictions regarding the tone-stress distribution in word-initial position, though any lexical tone can occur in the second syllable regardless of whether it is stressed or not (Li, 2003:114):

Table 12. Zhenhai initial tone distribution in [S w] and [w S] metrical domains

S-w		**w-S**	
σ1-σ2		σ1-σ2	
T1 (high register)	Tone	T1 (high register)	Tone
T2 (high register)		T3 (low register)	
T4 (low register)		T4 (low register)	
$T5^q$ (high register)		$T6^q$ (low register)	

"^{q}" labels a checked tone
Tone = any tone
σ= syllable

The content as shown in Table 12 is reproduced from Li (2003: 114). T1 and T4 can be either stressed or unstressed in word-initial position. However, T2 and $T5^{q}$ must be stressed word-initially. Word-initial T3 and $T6^{q}$ are always unstressed. There is no restriction regarding the tone-stress distribution in the second syllable: tones can occur freely in the second position, being stressed or unstressed (Li, 2003).

Table 13. [w S] disyllabic tone sandhi in Zhenhai

σ2 / σ1	T1 (HL)	T2 (MH)	T3 (ML)	T4 (LM)	$T5^{q}$ (H^{q})	$T6^{q}$ (LM^{q})
T1 (HL) high register	M-HL	M-HL	M-HL	M-HL	M-H^{q}	M-H^{q}
T3 (ML) low register	L-HL	L-HL	L-HL	L-HL	L-H^{q}	L-H^{q}
T4 (LM) low register	L-MH	L-MH	L-MH	L-MH	L-H^{q}	L-H^{q}
$T6^{q}$ (LM^{q}) low register	L^{q}-HL	L^{q}-MH	L^{q}-HL	L^{q}-LH	L^{q}-H^{q}	L^{q}-H^{q}

"^{q}" labels a checked tone
σ= syllable

Tone sandhi patterns in iambic disyllabic words are presented in Table 13. With regard to tone sandhi patterns in an iambic disyllabic domain, Li (2003:122) concludes that Zhenhai presents a case of dramatic positional neutralization in that the first (unstressed) syllable loses its lexically associated tone, and contrasting tone contours only appear on the second (stressed) syllable.[8] The first syllable is then realized with a default level tone, the actual tone value of which is determined by the tone register of the lexical tone originally associated with the first syllable. For instance, if the original tone is in the high register, then M, a high-registered low tone emerges; otherwise, L, a low level tone in the low register, appears in that position.

[8] Specifically, contrasting tone contours are realized on the second syllable only when the domain consists of two smooth tones. Contrasting contours do not seem to be shown on the second syllable when it is lexically associated with a checked tone. Instead, tone neutralization occurs in the second syllable if it underlyingly bears a checked tone.

Rose (1990:28) insightfully points out that the iambic disyllabic domain in Zhenhai is characterized by tone deletion, and tone attraction in the stressed position if the domain consisting of two smooth tones. The weak position displays tone migration: the underlying tone in the weak position is attracted, and migrates to the stressed position. Furthermore, the underlying tone of the weak position (the first syllable) is always realized in the high register when it moves to the strong position (the second syllable).

It is worth noting that these tone sandhi behaviors are not found in domains started by $T6^{q}$, the low-registered checked tone. Scrutiny of the tone sandhi in $T6^{q}$ + Tx, which is presented in the shaded cells in Table 13, highlights the following observations: 1) tone migration from the weak positon to the strong position is not found, 2) the lexical tone of the stressed syllable is retained, but obligatorily realized in high register, and 3) the unstressed syllable loses its underlying tone, and is subsequently neutralized to a low-registered level tone L.

Table 14. [S w] disyllabic tone sandhi in Zhenhai

σ2 / σ1	T1 (HL)	T2 (MH)	T3 (ML)	T4 (LM)	$T5^{q}$ (H^{q})	$T6^{q}$ (LM^{q})
T1 (HL) high register	H-L	H-L	H-L	H-L	$H\text{-}L^{q}$	$H\text{-}L^{q}$
T2 (MH) high register	MH-L	MH-L	MH-L	MH-L	$MH\text{-}L^{q}$	$MH\text{-}L^{q}$
T4 (LM) low register	LM-L	LM-L	LM-L	LM-L	$LM\text{-}L^{q}$	$LM\text{-}L^{q}$
$T5^{q}$ (H^{q}) high register	H^{q}-L	H^{q}-L	H^{q}-L	H^{q}-L	$H^{q}\text{-}L^{q}$	$H^{q}\text{-}L^{q}$

"^{q}" labels a checked tone
σ= syllable

In a trochaic ([S w]) disyllabic domain, the first syllable bears stress. As a consequence of tone-stress interaction, the first tone mostly remains at the original position. Being the unstressed syllable in a trochee, the second syllable loses its underlying tone and register. The second syllable is then assigned a default low level tone, "L". Irregularity is found in T1+ Tx,

where the tone sandhi exhibits a rightward tone spreading and tone re-association. However, rightward spreading of the word-initial tone is only a peripheral phenomenon, seeing that its occurrence is strictly restricted to word-initial T1.

Like other Northern Wu dialects (New Chongming, for instance), Zhenhai is typically characterized by tone reduction. As illustrated by Table 13 and Table 14, lexical tone contours in the unstressed syllable are no longer contrastive as they are reduced to level tones. Notably, the word-initial tone, or the word-initial tone contour in Zhenhai, is mostly realized in the *stressed* syllable. [9] When the first syllable bears stress, the word-initial smooth tone remains at the original site, except for the irregular case of T1 + Tx. When stress falls on the second syllable, tone movement from the first to the second syllable is motivated in most cases except in the sequences of $T6^{q}$ + Tx and T_{smooth} + T^{q}.

The tone sandhi pattern observed in the iambic domain poses problems for a theory where only metrical prominence is recognized. To account for the tone sandhi pattern in this regard, Li (2003) proposes that peripheral prominence, in addition to metrical stress, has to be referred to in Zhenhai phonology. In a tone system which recognizes metrical prominence, stress is usually an important factor regulating tone preservation and tone modification. It is widely attested that the stressed syllable is characterized by the retention of the lexical tone, as opposed to the deletion of the lexical tone triggered by lack of stress in the weak position. However, tone-stress interaction is not theoretically adequate to account for the preservation and migration of the tone in the weak position as observed in Zhenhai.

[9] In cases where tone register promotion happens, it is the word-initial tone contour rather than the word-initial tone that is preserved. See T3 + Tx in Table 13 for example, when the word-initial T3 (ML) is attracted to the stressed position, it is realized as a high register tone "HL". Therefore, the contour of T3 is preserved, while the actual pitch value is not.

Li (2003) argues for a dual-prominence system in Zhenhai: the peripheral prominence and metrical stress jointly manipulate tone preservation and migration. In the iambic disyllabic domain ([w S]), the peripheral prominence associated with the word-initial position retains the initial lexical tone, whereas the metrical stress in the second syllable attracts the word-initial tone and triggers tone migration.

Thus far, it seems that the tone sandhi patterns in Zhenhai present strong evidence for the recognition of peripheral prominence in Chinese dialects. However, in the section that immediately follows, I will argue that the observed peripheral prominence is still metrical stress in nature. Evidence will be presented to show Zhenhai is slowly shifting to a trochaic (left-headed) stress system from an iambic stress system. Over time, the emerging trochaic stress is becoming increasingly dominant in the Northern Wu dialect speaking region. In this transition, the residue iambic stress is still visible in tone sandhi patterns, but its effects are overridden at varying degrees by the emergence of trochaic stress.

4.3 PERIPHERAL PROMINENCE OR METRICAL PROMINENCE?

4.3.1 Shanghai: a case of dual prominences?

The tone sandhi patterns observed in Zhenhai remind us of the tone sandhi patterns of Shanghai. The tone sandhi patterns of Shanghai discussed in Table 10 of §3.3.2 are repeated in below for the convenience of reading.

Table 15. Shanghai tone sandhi patterns in disyllabic domain

Tone sandhi patterns		Note:
T1 + Tx	HL + Tx → H + L	"q" indicates a checked tone
T2 + Tx	LH + Tx → L + H	
T3 + Tx	LH + Tx → L + H	
$T4^q$ + Tx	LH^q + Tx → L + H	
$T5^q$ + Tx	**LH^q** + Tx → L + **LH**	

What deserves attention in Table 15 is the tone sandhi pattern found in the sequence of 5^q + Tx. As discussed in the previous chapter, when a disyllabic word starts with T5 (i.e. $T5^q$ + Tx), the disyllabic word forms an iambic ([w S]) foot, while all other sequences form trochaic feet. Same as the Zhenhai tone sandhi pattern in an iambic domain, the initial $T5^q$ in Shanghai is preserved and migrates to the stressed syllable, while all other tone combinations undergo the tone spreading and re-association processes.

This tone behavior naturally leads to the suspicion that left peripheral prominence should be incorporated in the phonology of Shanghai as well. The logic is: the left-headed stress in words initiated by T1, T2, T3, and $T4^q$ overlaps with the left peripheral prominence, which makes the peripheral strength invisible in the tone sandhi behaviors; however, when metrical stress and peripheral prominence do not co-occur in the same position, the prominence associated with the left edge becomes apparent. In the latter case, the prominence on the left edge is made visible through the word-initial tone preservation. More explicitly, in an iambic ([w S]) domain in Shanghai, the initial tone ($T5^q$) is preserved in the word-initial position because of the peripheral prominence associated with the left word-edge. $T5^q$ is later attracted to the stressed position so that we see tone preservation and migration in the domain of $T5^q$+Tx. This seems to be another strong evidence for the recognition of peripheral prominence in addition to metrical prominence in Chinese dialects.

The sandhi pattern for $T5^{q}$ + Tx in Shanghai seems to provide extra support for the recognition of peripheral prominence in addition to metrical prominence in Northern Wu dialects. Moreover, it seems the domain-initial prominence is competing with metrical prominence with regard to the lexical tone preservation. In an iambic domain, initial prominence keeps the initial lexical tone, however it cannot realize the initial tone. The metrical prominence does not keep its lexically associated tone. Instead, it attracts and realizes the tone in word-initial position.

However, I reject the view that there is peripheral prominence in the Northern Wu group. In the following section, data on Zhenhai's domain-initial vowel duration will be presented, which poses problem for the dual prominence analysis in that the increased duration of the word-initial vowel lacks an explanation. Furthermore, I will show that a stress-shift proposal can provide a better account for the initial tone preservation and migration. In this proposal, the phonetic salience of the word-initial vowel, and some confusing phonological observations, such as tone-sensitive stress placement, can be uniformly accounted for. § 4.3.2 will present the phonetic data on vowel duration in Zhenhai. §4.3.3 will work on the details of the stress-shift analysis.

4.3.2 Zhenhai: increased vowel duration at domain-initial position

It is of interest to note that the domain-initial position in Zhenhai is not only characterized by its ability to preserve the underlying tone, but also realized with phonetic salience. An instrumental study of Zhenhai conducted by Rose (1990) finds that the word-initial vowel in a disyllabic tone sandhi domain is produced with longer duration.

Rose (1990) reports the average duration for the first and second syllable of all possible tonal combinations in disyllabic sandhi context. Phonetic data showed that the vowel in the first syllable was realized with remarkably longer duration than that of the second syllable.

The corpus of Rose (1990) was recorded by one young male native speaker. The corpus consists of 402 disyllabic words with the segmental structure $C_1V_1C_2V_2$, where C indicates a consonant and V a vowel, the subscripted numbers indicate the syllable that a consonant or a vowel belongs to. In order to control the intrinsic effect of vowel quality on F0, samples contained roughly comparable numbers of high vowels and low vowels. Both V_1 and V_2 are monophthongs. The consonants are either voiceless unaspirated obstruents (high register tone), e.g. [p, t, ts, s], a sonorant, or voiced obstruents (low register tone).

Table 16 presents the mean and standard derivation values of durations in disyllabic words of Zhenhai. It is a faithful replica of the data presented in Rose (1990: 10). The duration of each segment is measured in centisecond. Stress is marked with an acute accent (′), so ′1+5 indicates a tonal combination of T1 and $T5^q$, where the first syllable is metrically strong. Rose (1990) does not provide data for the duration of the word-initial consonant, which is probably because the corpus was not embedded in carrier sentence when the recording was made. Thus, it is not always possible to locate the onset of a word-initial consonant.

Table 16. Mean and standard derivation values for durations in Zhenhai disyllabic tone sandhi.

	Tones	n	V_1		C_2		V_2	
			mean	SD	mean	SD	mean	SD
A	′1 + 1,2	16	28.9	7.1	17.4	5.2	15.9	3.4
	′1 + 3,4	21	33.4	8.7	8.1	2.7	16.6	2.7
B	′1 + 5	9	23.9	8.2	15.3	3.4	5.4	1.4
	′1 + 6	8	29.6	8.8	6.3	1.0	8.6	2.6
C	1 + ′1,2	16	32.6	5.2	17.4	3.6	20.0	3.3
	1 + ′3,4	16	31.8	8.4	7.7	1.4	21.0	2.8
D	1 + ′5	10	25.9	8.2	14.0	1.8	8.6	2.1
	1 + ′6	9	32.1	11.5	8.7	0.7	10.8	2.5
E	3 + ′1,2	20	30.4	7.0	16.5	4.7	20.1	2.9
	3 + ′3,4	21	35.3	8.8	8.0	2.4	21.8	3.1
F	3,4 + ′5	18	30.7	8.0	15.5	3.4	9.1	2.1
	3,4 + ′6	9	32.1	5.9	7.9	2.3	10.8	1.7
G	′2 + 1,2	17	25.9	5.2	14.5	2.7	17.0	3.9
	′2 + 3,4	13	29.1	5.4	7.0	1.8	16.1	2.3
H	′2 + 5	9	23.3	6.1	12.7	2.9	7.9	1.7
	′2 + 6	5	33.0	6.5	7.2	2.9	8.1	1.7
I	′4 + 1,2	13	28.8	5.5	14.4	3.5	17.3	2.8
	′4 + 3,4	19	33.1	9.0	7.3	1.8	16.7	2.4
J	′4 + 5	6	28.6	5.1	15.2	5.6	5.4	1.3
	′4 + 6	11	29.5	6.4	7.6	1.9	8.8	2.0
K	4 + ′1,2	15	32.1	6.6	16.3	4.2	26.1	2.7
L	4 + ′3,4	15	31.6	7.7	7.2	2.5	27.6	2.8
M	′5 + 1,2	15	7.4	1.7	12.9	4.1	15.8	3.7
	′5 + 3,4	12	9.1	1.4	7.0	3.8	19.1	2.8
N	′5 + 5	10	6.7	1.3	15.1	4.5	6.6	1.8
	′5 + 6	10	8.7	1.4	7.2	1.9	7.9	1.8
O	6 + ′1	7	6.2	0.4	17.5	1.8	21.9	4.1
P	6 + ′2	9	5.0	1.4	15.7	4.3	24.9	3.9
Q	6 + ′3	10	8.2	1.6	9.1	4.3	26.1	3.9
R	6 + ′4	11	10.3	1.3	8.5	3.2	30.0	1.5
S	6 + ′5	12	6.7	1.7	11.0	5.0	8.8	2.2
T	6 + ′6	11	10.6	2.4	7.7	2.2	12.1	2.5

V_1 = first-syllable vowel; V2 = second-syllable vowel; c = intervocalic consonant; n = number of items in sample.

(Rose, 1990:10)

Phonetic data in Table 16 show a systematic duration difference between the first and the second syllable. Except for the first syllable bearing a checked tone ($T5^{q}$ and $T6^{q}$), the duration of the first vowel is always considerably longer than the second one regardless of whether it is stressed or not. In a smooth-smooth tone sequence, when the second syllable is stressed, its vowel is longer than an unstressed vowel in the same position, but it is never as long as the domain-initial vowel. Checked tones are inherently short. They are always much shorter than the smooth tones regardless of its position and prominence status. When a comparison is made between two checked tones in a sequence, unlike the case of smooth tones, the domain-initial vowel does not seem to be remarkably long, and it is not necessarily longer than the second one.

In order to illustrate the above-noted observations, the following graphs in Figure 1 and Figure 2 are reproduced according to Li (2003:116).

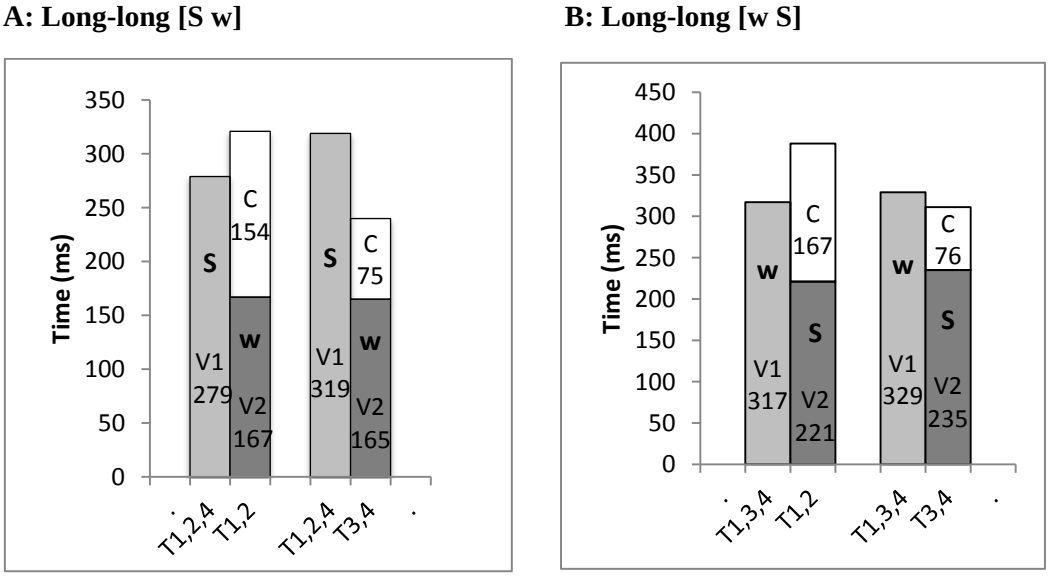


Figure 1. Duration values of vowels and consonants in syllables of long-long (smooth-smooth) tone combination

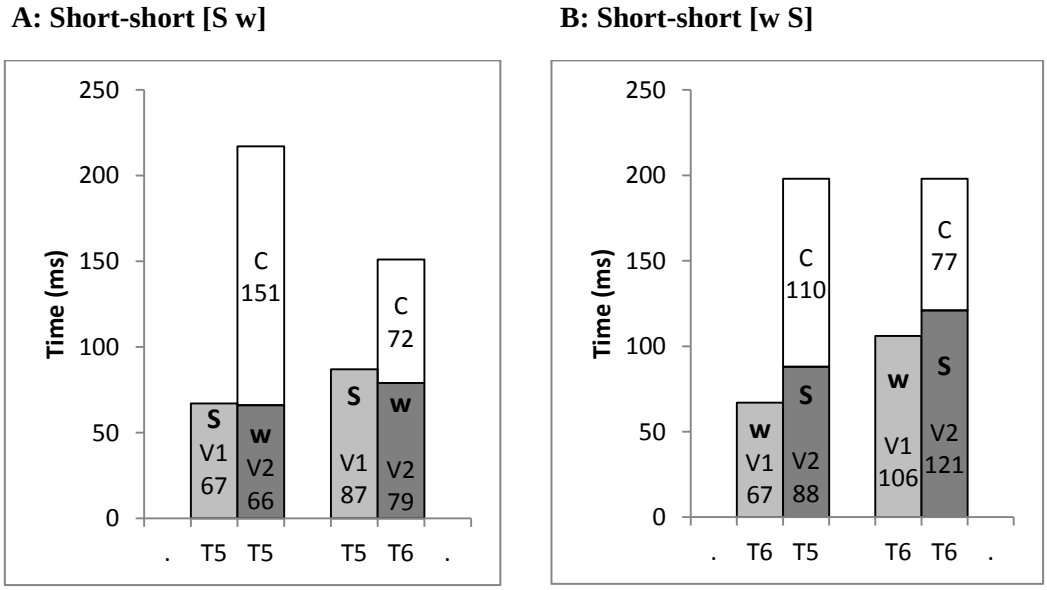


Figure 2. Duration values of vowels and consonants in syllables of short-short (checked-checked) tone combinations

Figure 1 shows segment durations in sequences composed of two long (smooth) tones. Figure 2 shows durations of segments in a sequence of two short (checked) tones. Figures for sequences composed of a long and a short tone are not generated. Since a long tone is always much longer than a short one in all conditions, the duration pattern of long-short/short-long sequences, therefore, is not our concern at the moment.

As shown in the above figures, S (strong) and W (weak) inside each vertical bar are labels indicating the metrical status of a tone. The first bar in each graph indicates the duration (in ms) of the vowel (V1) in the first syllable. The second stacked vertical bar represents the duration of the consonant (C) and vowel (V2) in the second syllable. The value of segment duration shown above is based on the data presented in Table 1, however, the unit of measurement is altered to millisecond (ms) in the above figures.

Take graph A in Figure 1, for example. The first bar shows the average duration of the vowel in the first syllable when it bears T1, T2, and T4 in a trochaic domain. The second bar

indicates the duration value of the intervocalic consonant and the vowel in the second syllable bearing T1 or T2. Graph B shows the duration value in an iambic disyllabic domain, where T1, T3, and T4 are found in the first syllable. Recall that T2 and $T5^{q}$ must be stressed at word-initial position, while T3 and $T6^{q}$ must not be stressed word-initially. T1 and T4 can be either stress or unstressed in word-initial position.

Two comments are in order with respect to the joint effects of tone, stress, and position on the vowel duration, as indicated by Figure 1 and Figure 2:

1) Although Rose (1990) does not conduct a statistical test to prove a significant difference in duration between the first and the second vowel, it is undeniable that the first vowel is extraordinarily long when it bears a smooth tone. As seen in Figure 1, the first vowel in a smooth tone is always much longer than the second vowel, even when it is not stressed.

2) When two checked tones occur next to each other, the long-duration effect is not always apparent in word-initial position. In this case, the influence of stress on vowel duration is observable. For instance, when the first checked tone is not stressed, its vowel is shorter than the following stressed counterpart. When the word-initial position is metrically stressed, the first vowel tends to be longer than the second one, as shown by graph A in Figure 2.

The long duration of vowels at domain-initial position in Zhenhai, naturally, needs an explanation. However, peripheral prominence cannot provide an explanation in this regard. Since peripheral prominence is essentially an abstract property associated with word-edges, it is not supposed to be enhanced by salient phonetic cues.

In order to account for the word initial long-vowel effect in Zhenhai, Li (2003) argues that the word-initial position is phonetically strengthened with extra duration by a process known as initial strengthening. To justify this suggestion, Li enumerates works on initial strengthening such as Barns (2001), Byrd (2000), Dilly, Shattuck-Hufnagel & Ostendrof (1996), Fougeron & Keating (1996), Fougeron (1999) among many others. I will argue immediately in the following parts that the long-vowel effect poses problem for the peripheral prominence proposal in Zhenhai. In addition, the long duration of vowels at word-initial position is not a consequence of initial strengthening; rather, it is a result of metrical prominence.

4.3.3 Peripheral prominence is still metrical prominence in Zhenhai

4.3.3.1 Peripheral prominence, metrical prominence and stress

Before we can proceed to the discussion of the nature of prominence associated with the left boundary of the disyllabic domain in Zhenhai, it is important to clarify some key notions with regard to peripheral prominence and initial strengthening.

According to Zoll (1997), a crucial distinction must be made between *inherent* prominence and *metrical* prominence. Peripheral positions, including the initial and the final syllables in a domain, are recognized as positions for inherent prominence as opposed to the imposed prominence in metrical positions. The prominence associated with peripheral positions is manifested by its ability to license a fuller range of contrasts than do no-prominent sites in languages. This is eminently true for initial syllable. For example, "In !Xõõ, 116 segments, primarily different kinds of clicks, are licensed in prosodic word-initial position, while only six can appear intervocalically and only two word finally" (Zoll, 1997:106-107).

While peripheral prominence naturally supports contrast, the metrically strong positions require further enhancement. The enhancement typically takes the form of high tones or stress (Zoll, 1997). The manifestation of metrically strong positions, therefore, often involves perceptual prominence properties, including pitch, duration and amplitude. In Zoll's theory, stress and metrical prominence are not identical notions. Metrical prominence is often realized with stress. However, metrical prominence should not be confused with stress, because unbounded stress may occur without metrical structure, such as foot (Prince & Smolensky, 1993), and feet may occur without stress (Zoll, 1997).

Therefore, a fundamental distinction between peripheral prominence and metrical prominence is that the former is an abstract property associated with domain boundaries, whereas the latter is often, if not always, realized with salient phonetic cues.

Smith (2002) pursues a different dichotomy of prominence: *phonetically strong positions* and *psycholinguistically strong positions*. While stressed syllable is recognized as *phonetically strong*, boundary prominence, such as word-initial syllable, is defined as a *psycholinguistically strong position*, because the phonological contrasts kept there play a special role in psycholinguistic processing of word recognition (Smith, 2002).

In both theories of prominence stated above, prominence of domain-initial syllables is not overtly manifested by phonetic salience. Rather, peripheral prominence at domain-initial position is expressed abstractly by its ability to resist phonological neutralization. In contrast, stress is either realized with one or more phonetic salient cues (Smith, 2002) or used as a phonetically salient property to enhance metrical prominence (Zoll, 1997).

4.3.3.2 Initial strengthening effect

It has been noted that the phonetic realization of segments is affected by various prosodic factors, such as stress and final lengthening effect. The word-initial position is also examined for the strengthening effect it may have on segment realization. Cooper (1991), for instance, studied English voiceless aspirated stops in word-initial position. He finds glottal gesture is strengthened domain-initially in that word-initial voiceless aspirated stops are realized with a larger glottal opening gesture than the ones in word-medial position. Furthermore, the strengthening effect at domain-initial position is cumulative in that higher prosodic boundaries produce stronger effect. For instance, Jun (1993, 1995) finds that VOT for Korean /p^h/ is longer at phrase-initial position compared to word-medial site, and phrase-initial VOT is longer than word-initial VOT.

Apart from the strengthening of glottal opening gesture, supralaryngeal articulations have also been found to be strengthened domain-initially. A series of studies has found that consonants are generally produced with greater and longer linguopalatal contact (contact between tongue and palatal area) at domain-initial position than domain-final position (Cho & Jun, 2000). Similar findings have been claimed cross-linguistically, for instance in Taiwanese (Hsu & Keating, 1999) and French (Fougeron, 1999).

It may seem that the initial strengthening effect can affect the initial syllable as a whole. However, the most agreed-upon view on initial strengthening is that it enhances the contrast between the initial segment and the neighboring sounds. Specifically, the initial strengthening effect is viewed as the enhancement of CV or VC contrast (Cho & Jun, 2000).

Existing data suggest that what is strengthened in domain-initial position is the consonantality of the initial segment, which is meant to increase the contrast of the onset with the following vowel (Cho & Jun, 2000; Fougeron & Keating, 1996; Hsu & Jun, 1998; Pierrehumbert

& Talkin, 1992). This finding is evidenced by the more extreme oral constriction produced for domain-initial consonants. For instance, voiceless stops are produced with longer closure duration and longer VOT, which makes the domain-initial consonants more consonantal (Pierrehumbert & Talkin, 1992; Jun, 1995). Similarly, nasals are produced with weaker nasal airflow, and thus are less sonorant in domain-initial position (Fougeron, 1999).

Since much of the work on initial strengthening has been focused on domain-initial consonant, a number of subsequent studies attempt to detect the effect of initial strengthening on vowels. Hsu and Jun (1996) finds that closure duration of the phrase-initial consonant is longer than the counterpart in phrase-medial position. However, the vowel after the phrase-initial consonant is actually shorter than that after consonant in phrase-medial position. Similar conclusions are also found in other studies on this topic. Byrd (2000), for example, extends phonetic experiments to boundary adjacent vowels to investigate whether they are lengthened in a comparable degree to boundary adjacent consonants, and whether the relative timing among consonant and vowel gestures is affected by adjacent boundaries. Byrd finds that subjects show substantial lengthening for preboundary vowels (final lengthening) whereas effects on vowel of the postboundary CV syllable (initial lengthening) are negligible (Byrd, 2000:9). In addition, boundaries cause the gesture of postboundary vowel to be later than the control group. Thus, Byrd (2000:11) concludes that "the constrictions for the segments in the phrase initial syllable were less coproduced or overlapped when the syllable was initial in a higher (prosodic) domain".

The findings of Byrd (2000) and other studies are supportive of the enhancement of domain initial CV contrast, whereas they do not support the lengthening effect for domain-initial vowels. These findings also clarify that the scope of initial lengthening and final lengthening is

different. Initial lengthening enhances the domain-initial consonant, while final lengthening works on the domain-final vowel.

In addition, the difference between initial enhancement and final lengthening in terms of their implementation target is confirmed by the study of Zhang (2001). Based on his cross-linguistic survey on the distribution of contour tones, Zhang concludes that the final position of a prosodic domain patterns with stress syllables and syllables with long rimes in contour tone licensing, while domain-initial position does not. Therefore, Zhang (2001) does not recognize domain-initial syllable as a privileged docking site for contour tones. Zhang (2001) suggests that the ability to license contour tones in domain-final position can be interpreted as a result of the increased duration caused by final lengthening effect.

All the above evidence suggests that initial strengthening effect is local to the domain-initial consonant, and vowels are not shown to be affected. This conclusion presents serious challenge to the claim that the long duration of vowels in the left boundary of disyllabic words in Zhenhai is caused by initial strengthening effect. Therefore, some fundamental revisions have to be made to the claim of peripheral prominence suggested by Li (2003). As seen above, the dual-prominence analysis proposed for Northern Wu dialects is shown to be problematic when it cannot convincingly account for some important phonetic observations. Thus, an alternative solution to this problem has yet to be proposed.

4.3.3.3 Peripheral prominence is still metrical in Zhenhai

In conclusion, the increased duration of the initial vowel in Zhenhai cannot be successfully accounted for by appealing to either peripheral prominence or initial strengthening effect. To reiterate, peripheral prominence is an abstract property, which is not supposed to be manifested phonetically. Likewise, initial strengthening is proved to be insufficient to give rise to the long

duration of word-initial vowels. Therefore, in order to adequately characterize the word-initial long vowel phenomenon, further exploration is needed.

I argue in below that the long-vowel effect associated with the word-initial syllable is the result of metrical prominence located word-initially. Evidence to be shown below suggests that peripheral prominence is a wrong characterization. Additional analyses of the phonological and phonetic details of the disyllabic words in Zhenhai, and other dialects in Northern Wu group suggest that iambic feet and trochaic feet may co-occur at lexical level. The co-occurrence of both stress patterns is a result of the emerging trochaic stress in addition to the original iambic stress in Northern Wu dialects.

4.3.3.3.1 Tone-sensitive stress placement: Suzhou, Shanghai, and Zhenhai

As noted in §3.3.2, although there are iambic feet ([w S]) in Shanghai, the more widely distributed stress pattern is trochaic ([S w]). The iambic foot is strictly restricted to disyllabic words that start with a low-registered checked tone ($T5^{q}$). In contrast, trochaic feet are more widely seen, which occur to disyllabic words initiated by all the other tones.

In Zhenhai, similarly, restrictions on tone-stress distribution are found. While some tones can be either stressed or unstressed at domain-initial position, certain tones must or must not be stressed at the same position. As a matter of fact, the tone-sensitive stress distribution is rather common in Northern Wu dialects. With the following analyses on the stress pattern of Suzhou, and the revisit of tone-stress interaction in Shanghai and Zhenhai, the relation between tone and stress in the Northern Wu group will be further revealed.

Suzhou is another dialect in the Northern Wu group, which has a metrical system very close to that of Shanghai. The data of duration ratios presented below provide some insight into the stress pattern of Suzhou.

Table 17. Ratio of duration between the syllables in disyllabic words in Suzhou

σ2 / σ1	Smooth tone (T1,T2, T3, T4,T5)	Checked tone ($T6^q$ and $T7^q$)
T1(smooth tone)	1.2 /1	2.0/1
T2(smooth tone)	1.2/1	2.1/1
T3(smooth tone)	1.3/1	
T4(smooth tone)	1.1/1	1.8/1
T5(smooth tone)	1.5/1	
Average	1.3/1	2/1
$T6^q$ (checked tone)	1/2.7	1/1.2
$T7^q$ (checked tone)	1/2.7	1/1.2
Average	1/2.7	1/1.2

"q" indicates a checked tone
σ= syllalbe

Data in Table 17 is based on an acoustic study of Suzhou in Liao (1994:156), where one can find the duration ratio of the first syllable to that of the second syllable. Duration ratios for some tonal sequences are not available in Table 17, because such data are absent in the original source.

It is of no interest to see smooth tones are always longer than the checked tones in all tone combinations in Table 17, since smooth syllables naturally support longer duration than do checked syllables. However, it is important to note that the word-initial syllable is realized with some extra duration when it bears a smooth tone. This fact is directly observable in the smooth-smooth tone sequence: on average, the smooth syllable is 30% longer than the second one.

On the contrary, according to what the ratio indicates, it is the second syllable in a T^q (checked tone) + Tx sequence that gets some extra duration. This is easy to discern in the T^q + T^q combination: overall, the second checked tone is longer than the first one by 20%, as indicated by the ratio 1/1.2. In the T^q + T_{smooth} sequence, however, this pattern is less observable since the long duration inherently associated with the smooth tone may lead to unintended negligence to

the extra length realized there. To detect it, some additional efforts are needed. To help detect such a pattern, a comparison is made between $T_{smooth} + T^{q}$ and $T^{q} + T_{smooth}$:

Table 18. The smooth tone in $T_{smooth} + T^{q}$ and $T^{q} + T_{smooth}$

a. $T_{smooth} + T^{q}$ Sequence

Smooth Tone	Checked Tone
2	1

b. $T^{q} + T_{smooth}$ Sequence

Checked Tone	Smooth Tone
1	2.7

As shown in Table 18 (a), on average, a smooth tone is twice as long as a following checked tone. One may expect to see a similar duration ratio between the smooth and the checked tone in Table 18 (b). However, data in (b) shows that the word-final smooth tone is almost three times as long as the word-initial checked tone (2.7/1), which suggests that a smooth tone in the word-final position is longer in proportion to a checked tone than in the domain-initial position. Note that domain-initial smooth syllables are already shown to have larger duration ratio.

In view of the data presented above in Table 17 and Table 18, it is evident that smooth tones have longer duration when they occur word-initially. On the other hand, word-final syllables are longer only when occurring after checked tones. One may wonder whether the longer tone duration in the latter case is attributable to final lengthening effect, in view of the fact that Liao (1994) did not specify whether such an effect was eliminated in her study by embedding the corpus in carrier sentence. However, final lengthening effect does not work nicely in the case of Suzhou. Although increased duration occurring at word-final position can be potentially accounted for by final lengthening effect, the long duration of word-initial smooth tones still lacks an explanation. In other words, lengthening effect cannot provide an integrated and unified explanation for the duration pattern in Suzhou.

The duration pattern in disyllabic domain has a simple and natural interpretation if stress is assumed in Suzhou. When metrical stress is posited, the longer duration associated with certain positions can be naturally identified as a salient phonetic property commonly seen in stressed positions. More importantly, as will be shown in below, a tone-stress relation can be captured and depicted with a theory of stress.

Moreover, the tone sandhi patterns in Suzhou seem to be stress-sensitive, which provides additional evidence for the existence of stress. To begin, the citation tones in Suzhou are first presented in Table 19:

Table 19. Citation tones in Suzhou

Tone Type	High register	Low register
Smooth tone (long)	T1 55 HH	T2 23 LM
	T3 51 HL	
	T4 523 HLM	T5 231 LML
Checked tone (short)	$T6^{q}$ 55 H^{q}	$T7^{q}$ 23 LM^{q}

"^{q}" labels a checked tone

As shown in Table 19, Suzhou has a commonly seen tone system in the Northern Wu group. First of all, the division between the smooth tone and the checked tone is kept. Secondly, the tones were split into two registers, though there are not same numbers of tones in each register. When it comes to the tone sandhi patterns in Suzhou dialect, it is interesting to see the tone sandhi behaviors are in parallel with the duration pattern observed in disyllabic words.

Table 20. Suzhou tone sandhi patterns in disyllabic domain

σ2 / σ1	T1 (HH)	T2 (LM)	T3 (HL)	T4 (HLM)	T5 (LML)	T6^{q} (H^{q})	T7^{q} (LM^{q})
T6^{q} (H) High register	H^{q}-LM		H^{q}- HL	H^{q}- HLM	H^{q}- HL H^{q}- HLM	H^{q}- H^{q}	
T7^{q} (LM) Low register	LM^{q}- ML					LM^{q}-H^{q}	
T1 (HH) High register	HH - LL						
T2 (LM) Low register	LL - MM						
T3 (HL) High register	HL - LL						
T4 (HLM) High register	HL - LM						
T5 (LML) Low register	LM - LL						

"^{q}" labels a checked tone
σ= syllable

The tone sandhi patterns presented in Table 20 are based on both an acoustic study (Liao, 1994) and impressionistic studies (Qian & Shi, 1983; Xie, 1982) on the tonology of Suzhou. Obviously, the more pervasive tone sandhi pattern found in Table 20 is the rightward tone spreading and tone re-association, which dominants the disyllabic domain of T_{smooth} + Tx. In this domain, the first tone is retained and part of its melody spreads to the second syllable where the underlying tone is deleted.

There are more complications of tone sandhi in the T^{q} + Tx sequence, which are presented in the shaded cells in Table 20. As seen, the second syllable tends to keep their lexical tones. In addition, rightward tone spreading of the initial tone is not found. This is especially true in the case of T6^{q} + Tx, where more examples of lexical tone retention are present in the second syllable, though tone neutralization has emerged at the same time. For instance, T1 has been merged with T2 in the second syllable. Tone neutralization is more common in T7^{q} + Tx, where only two tones can occur in the domain-final position: a checked tone (H^{q}) and a smooth tone

(ML). To summarize briefly, the T_{smooth} + Tx domain is characterized by second tone loss and initial tone spreading, whereas the syllables in T^q + Tx tends to retain their lexical tones.

The dichotomy of tone sandhi patterns between T_{smooth} + Tx and T^q + Tx parallels the duration pattern observed in the disyllabic domain. Recall that the syllable duration show the same division between T_{smooth} + Tx and T^q + Tx. The same dichotomy in the two sets of patterns should not be just accidental. As stated previously, when stress and metrical structure are posited for Suzhou, it is easy to see the connection between the tone and duration behaviors. To make it more explicitly, the tone sandhi and the syllable duration pattern of disyllabic words can be straightforwardly interpreted as processes driven by the same motivation: stress.

According to the observations on syllable duration and tone sandhi, the metrical patterns of Suzhou can be worked out as follows:

(1) Metrical patterns and their distribution in tones of Suzhou

First syllable	Second syllable	Metrical pattern
smooth tones (T1, T2, T3, T4,T5)	Tx	trochaic [S w]
checked tone ($T6^q$, $T7^q$)	Tx	iambic [w S]

Chen (2000: 296) points out that an important diagnostic of stress in Chinese "is the retention of tone and the ability to determine the shape of a neighboring tone. By the same token, a syllable in weak position tends to undergo tonal modification". Roughly, the stress-sensitive tone modification is what has been observed in Suzhou and Shanghai.

Tone, on the other hand, can affect the distribution of stress. This can be seen in the stress patterns listed in (1). While smooth tones can be stressed domain-initially, stress avoids domain-initial checked tones. In retrospect, similar tone-dependent stress assignment is also found in Shanghai. See below for the tones and metrical patterns in Shanghai:

(2) Metrical patterns and their distribution in tones of Shanghai

First syllable	Second syllable	Metrical pattern
smooth tones (T1, T2, T3)	Tx	trochaic [S w]
checked tone (high register $T4^q$)	Tx	
checked tone (low register $T5^q$)	Tx	iambic [w S]

Although domain-initial checked tone is not strictly excluded from being stressed, it is rather apparent that iambic feet unexceptionally fall on domains initiated by a checked tone in both Shanghai and Suzhou. Reversely, a trochaic foot tends to (but not always) fall on domains with an initial smooth tone.

However, the tone-sensitive stress placement does not seem to exist in Zhenhai. See the tone-stress distribution in Zhenhai listed in (3):

(3) Metrical patterns and their distribution in tones in Zhenhai

First syllable	Second syllable	Metrical pattern
smooth tones (T1, T2, T4)	Tx	trochaic [S w]
checked tone (high register $T5^q$)	Tx	
smooth tones (T1, T3, T4)	Tx	iambic [w S]
checked tone (low register $T6^q$)	Tx	

At first glance, observations from (3) seem to suggest that the tone-dependent stress placement seen in Suzhou and Shanghai does not hold for Zhenhai. For instance, though Zhenhai does show a restrictive tone-stress distribution, stress is not found to be sensitive to the domain-initial tone in a smooth-checked dimension. However, it is interesting to point out that Zhenhai does have a duration pattern similar to that of Shanghai and Suzhou in the disyllabic domain:

Table 21. Ratio of vowel duration between syllables in disyllabic words in Zhenhai

σ2 / σ1	Smooth tone (T1, T2, T3, T4)	Checked tone ($T5^q$, $T6^q$)
T1 (smooth tone)	1.7/1	3.3/1
T2 (smooth tone)	1.7/1	3.5/1
T3 (smooth tone)	1.6/1	
T4 (smooth tone)	1.6/1	4.1/1
$T5^q$ (high register)	1/2.1[10]	1.1/1
Average	1.2/1	3/1
$T6^q$ (low register)	1/3.5	1/1.2
Average	1/3.5	1/1.2

"q" indicates a checked tone

σ= syllable

The data presented in Table 21 are calculated on the basis of duration values provided in Table 16, § 4.3.2. Since the duration of the domain-initial consonant is not measured in the original study, Table 21 has to present ratios of vowel duration between the first and the second syllable, which is meant to roughly reflect the ratio between syllable durations in disyllabic words. Ratio for T3 + T^q is not provided in Table 21, because the data needed for calculation cannot be extracted from the original data source.

As observed in the above table, the duration pattern in the disyllabic domain of Zhenhai behaves exactly like the pattern in Shanghai: in domain initiated by a low-registered checked tone, the second syllable is realized with some extra duration; in the rest cases, the initial syllable is.

Notably, Zhenhai's duration pattern as shown in Table 21 is not reflected in the stress pattern proposed by Rose (1990) and Li (2003). In Shanghai and Suzhou, the syllable duration accurately reflects the metrical pattern. For instance, the syllable gets longer duration in a disyllabic domain is always the stressed syllable in a foot. This is not always true in Zhenhai.

[10] Note that smooth tones occurring after $T5^q$ are actually shorter than they normally are. For instance, a smooth tone is 3.5 times as long as a word-initial $T6^q$, and on average 3.6 times as long as a following checked tone. However, it is only around twice as long as the word-initial checked tone in the $T5^q$ + Tx sequence.

Although the initial smooth tone is always realized with some extra duration, it is not necessarily the stressed syllable in a trochee.

One may argue that the correspondence between the duration pattern and the stress pattern in Shanghai and Suzhou is a matter-of-course, because the metrical pattern in these dialects is primarily determined on the basis of the duration pattern. However, the stress pattern in Shanghai and Suzhou is also independently evidenced by the tone sandhi patterns: unstressed syllables typically undergo tone reduction and deletion, whereas the stressed syllable tends to retain the lexical tone or determine the shape of the neighboring tone.

Yet if one follows the analyses proposed by Rose (1990) and Li (2003), the mismatch between the duration pattern and the stress pattern in Zhenhai will remain perplexing. Must we assume the analyses proposed by Rose (1990) and Li (2003) for Zhenhai? Not necessarily. I thus suggest a stress system for Zhenhai that is very different from the analyses in Rose (1990) and Li (2003). In my analysis presented below, there is no place for peripheral prominence; instead, only metrical stress is recognized. Moreover, the analysis argues for an on-going stress shift in Zhenhai from right-headed stress to left-headed stress. The argument and evidence in support of this proposal will also answer two concerns expressed earlier, 1) why the domain-initial smooth tone in Zhenhai is characterized by increased vowel duration, and 2) why stress tends to be avoided on the domain-initial checked tone. Evidence and argument for the stress-shift analysis suggested for Zhenhai will be given in § 4.3.3.3.2. § 4.3.3.3.3 will elaborate on tone-sensitive stress placement found in the Northern Wu group.

4.3.3.3.2 From right-headed stress to left-headed stress

A crucial problem of the peripheral prominence proposed in Li (2003) for Zhenhai is that it cannot successfully account for the increased vowel duration at word-initial position. This

problem cannot be solved by appealing to the stress pattern proposed by Li (2003) and Rose (1990) either. For instance, the increased initial vowel duration can hardly be interpreted as a stress-driven phenomenon in an iambic disyllabic domain.

The increased vowel duration at word-initial position can be accounted for by proposing that Zhenhai is evolving into a trochee-dominant system from an iambic system. In the progress towards the right-headed stress system, disyllabic domains initiated by checked tones are more conservative than the counterpart started by smooth tones. Moreover, the domain with an initial low-registered checked tone is more conservative than that with an initial high-registered checked tone. Therefore, the disyllabic domain initiated by a low-registered checked tone is the most conservative.

In the residue iambic disyllabic domain, the trochaic stress has already emerged. For instance, the domain-initial vowel is phonetically long. And, the strength of trochaic stress is made visible through the initial tone preservation. Therefore, two competing metrical forces are seen in the iambic domain: the emerging left-headed stress which retains the word-initial tone, and the residue right-headed stress which attracts the initial tone and triggers tone migration.

However, it is important to point out that the emergence of left-headed stress is not found in the iambic domain of $T6^{q}$ + Tx in Zhenhai. As said, the disyllabic domain with an initial low-registered checked tone is most conservative in the progression towards trochaic stress. We are now turning to the specialty of $T6^{q}$ + Tx.

An advantage of assuming stress shift in Zhenhai is that some perplexing phonological and phonetic observations can be accounted for in a unified way. First of all, the stress-shift analysis can provide an explanation to the specific tone sandhi behaviors for $T6^{q}$ + Tx. In order to demonstrate the irregular tone sandhi in $T6^{q}$ + Tx, the content in Table 13 is repeated below:

Table 22. [w S] disyllabic tone sandhi in Zhenhai

σ2 / σ1	T1 (HL)	T2 (MH)	T3 (ML)	T4 (LM)	T5^{q} (H^{q})	T6 ^{q} (L^{q})
T1 (HL) high register	M-HL	M-HL	M-HL	M-HL	M-H^{q}	M-H^{q}
T3 (ML) low register	L-HL	L-HL	L-HL	L-HL	L-H^{q}	L-H^{q}
T4 (LM) low register	L-MH	L-MH	L-MH	L-MH	L-H^{q}	L-H^{q}
T6^{q} (L^{q}) low register	L^{q}-HL	L^{q}-MH	L^{q}-HL	L^{q}-LH	L^{q}-H^{q}	L^{q}-H^{q}

"^{q}" labels a checked tone
σ= syllalbe

As seen above, the tone sandhi behavior in T6^{q} + Tx is very different from tone sandhi patterns shown elsewhere. Specifically, the initial checked tone is not found to be attracted to the stressed syllable; and the lexical tone in the stressed syllable is not deleted as in other cases. This exceptional tone behavior serves as evidence to deny the dual-prominence proposal, because it clearly illustrates that the peripheral prominence is absent in T6^{q} + Tx. Rather, T6^{q} + Tx behaves exactly like a standard iambic foot, where only metrical stress can be seen at work: the stressed syllable is characterized by its ability to retain the lexical tone, and the tone in the weak syllable is neutralized to a default low tone.

In line with the study of Zoll (1997), Li essentially defines peripheral prominence as a position-based inherent property. In Li's framework, there is no clear phonological reason why the domain-initial position in T6^{q} + Tx is excluded as a site for peripheral prominence. Consequently, the absence of peripheral prominence in T6^{q} + Tx sequence is inexplicable.

However, the tone behavior in T6^{q} + Tx will not be a problem, if the emergence of trochaic stress in Zhenhai is taking into consideration. Further, tone-sensitive stress placement should also be assumed. Under this assumption, trochaic stress tends to be avoided on domain-

initial checked tones, especially the low-registered checked tone. This assumption offers an explanation for why the emerging left-headed stress is absent in the domain of $T6^q$ + Tx, and therefore answers why only iambic stress is found to be conditioning the tone sandhi behavior in $T6^q$ + Tx. Moreover, it thus becomes explicit that the lack of increased vowel duration at the word-initial syllable of $T6^q$ + Tx is a result of the absence of the emerging trochaic stress on the low-registered checked tone.

Secondly, when the stress-shift proposal is followed in the analysis of the Northern Wu dialects, the co-existence of iambic stress and trochaic stress in Shanghai, Suzhou and Zhenhai can be accounted for as a transitional state in the shift from a right-headed stress system to a left-headed tress system. See details presented in Table 23 about the different stages of the three dialects in the stress-shift progression:

Table 23. Different stages of Shanghai, Suzhou and Zhenhai in the shift from iambic to trochaic stress system

	Shanghai		
	T_{smooth} + Tx	**T_H^q + Tx**	**T_L^q + Tx**
Iambic stress	no	no	yes
Trochaic stress	yes	yes	yes
	Suzhou		
	T_{smooth} + Tx	**T_H^q + Tx**	**T_L^q + Tx**
Iambic stress	no	yes	yes
Trochaic stress	yes	no	no
	Zhenhai		
	T_{smooth} + Tx	**T_H^q + Tx**	**T_L^q + Tx**
Iambic stress	yes	no	yes
Trochaic stress	yes	yes	no

T_H^q = high-register checked tone
T_L^q = low-registered checked tone
T_{smooth} = smooth tone, Tx= any tone

In Table 23, the disyllabic domain in Shanghai, Suzhou, and Zhenhai are divided into three groups: T_{smooth} + Tx, T_L^q + Tx, and T_H^q + Tx, where T_L^q and T_H^q are representing low-registered and high-registered checked tones respectively . According to Table 23, it is easy to

see that the domain with an initial low-registered checked tone (T_L^q + Tx) is most resistant to stress shift. In Suzhou and Zhenhai, there is no attested phonetic or phonological evidence showing the emergence of trochaic stress in this context.

T_L^q + Tx in Shanghai presents a more complex case. According to the data on syllable duration, there is no evidence that confirms the emergence of left-headed stress in this domain. However, the tone sandhi behaviors observed in T_L^q + Tx do prove the existence of left-headed stress:

(4) $T5^q$ + Tx $\quad HL^q + Tx \rightarrow L + HL^q$

The rule in (4) depicts the tone sandhi pattern occurs in the domain with an initial low-registered checked tone ($T5^q$ + Tx) in Shanghai. As seen in (4), the initial tone retention and migration are exactly like the pattern found in Zhenhai, revealing the co-existence of left-headed stress and right-headed stress. Thus, the emergence of left-headed stress in the domain of T_L^q + Tx in Shanghai is shown through phonological behaviors rather than phonetic correlates.

In the context of T_H^q + Tx, the disyllabic words in Zhenhai and Shanghai have evolved into trochees. However, Suzhou is more conservative, where the left-headed stress has not emerged yet.

In the context of T_{smooth} + Tx, disyllabic words in Shanghai and Suzhou have already switched to trochees. Zhenhai presents a mixed system: according to Rose (1990) and Li (2003), T2 + Tx is a trochaic domain, T3 + Tx is an iambic domain, whereas T1 + Tx and T4 + Tx can be both. However, as mentioned earlier, the emergence of trochaic stress has already been attested both phonetically and phonologically in iambic domains other than $T6^q$ + Tx. The phonetic manifestation for the emergence of trochaic stress, i.e., the increased vowel duration at domain-initial position, suppresses the durational correlate of the original iambic stress. This

situation leads to a consequence: the stressed syllable in iambic feet is no longer phonetically cued by duration; however, iambic stress can still be shown through tone sandhi operations.

Moreover, the distribution of the residue iambic stress in tone shows that the progression toward the trochaic stress system is stratified: in Zhenhai, while T2 + Tx and T5^{q} + Tx have already switched to trochee, T3 + Tx still shows evidence of iambic stress through tone sandhi behaviors. T6^{q} + Tx presents a case where only iambic stress is found. T1 + Tx and T4 + Tx are hybrid cases, where both switch-over to trochaic stress and legacy iambic stress can be found.

As a final note, it is important to mention that the shift from iambic stress to trochaic stress is also evident in other Northern Wu dialects. In the following paragraphs, a brief depiction will be presented for such a dialect, Wuxi.

As stated in Chen (2000), Wuxi exhibits very complicated tone sandhi patterns. Details aside, the tone sandhi patterns in Wuxi involve tone substitution in the leftmost syllable. The left tone acquired through substitution is then extended to the entire tonal domain. Importantly, the tone substitution is conditioned by the tone on the right syllable. The tone derivation in Wuxi proceeds as follows (Chen, 2000:324):

(5) a.

vu	*tɕi*	‘*weapon*’
HHL	LHH	base tone
LLH	LHH	**pattern substitution**
LLH	0	tone deletion
L	LH	tone spreading

b.

na	*dã*	‘*milk candy*’
HHL	LHL	base tone
LHL	LHL	**pattern substitution**
LHL	0	tone deletion
LH	HL	tone spreading

According to Chan & Ren (1989), the left tone substitution as illustrated in (5) is determined by the tone on the right syllable. In the tone derivation, the right syllable first retains its underlying tone. At the point when tone substitution happens, the tone category of the right syllable determines the tone contour which replaces the original tone on the left syllable. Note that the difference of the tones on the right syllable causes the different left-tone substitution in (5a) and (5b), even though (5a) and (5b) originally have the identical tone on the left syllable. Once the left tone substitution finishes, the newly acquired tone spreads onto the entire tone sandhi domain, wiping out the original tone that is on the remaining syllable of the span.

Chan & Ren (1989) takes the tonal derivation in Wuxi as evidence of last syllable dominance at some stage in the tonal history of Wuxi. The last syllable dominance is now replaced by a system of initial syllable dominance. Chen (2000: 325) agrees with Chan and Ren's stress-shift theory, and he further argues that the endpoint of this long evolution was already reached by other better known Northern Wu dialects such as Shanghai, and Suzhou among others. However, as our analyses show, the endpoint of the evolution has not been finally reached yet in Shanghai and Suzhou, since the last-syllable dominance is still visible more or less in the disyllabic domain initiated by checked tones, through the tone sandhi and duration patterns.

4.3.3.3.3 Why is stress placement tone-sensitive in the Northern Wu group?

As seen, stress tends to be avoided on domain-initial checked tones. Moreover, stress tends to be avoided on low-registered tone. In particular, low-registered checked tone is most resistant to be stressed at domain-initial position as instantiated by the Northern Wu dialects discussed so far.

Chen (2000: 222) proposes that the phonetics of the smooth vs. checked syllables gives some clue to this tone-sensitive stress placement. According to duration measurement provided

by Zee and Maddieson (1979: 54), Zhu (2005: 215) and Liao (1994:153) among others, smooth syllables in citation are about twice as long as checked syllables in citation. Therefore, Chen (2000) suggests that smooth syllables and checked syllables are heavy and light respectively in terms of syllable weight. The distributional restriction on domain-initial stress is easy to account for if WSP (Weight-to-Stress Principle) is assumed in the shift from iambic stress to trochaic stress:

(6) tone-sensitive stress placement

WSP: Avoid stressing an (initial) light syllable.

By further stipulating that left-headed stress placement is subject to WSP, it then becomes apparent that a stressed checked tone in a trochee ($'T^q + \sigma$) is disfavored, because checked syllables are light in terms of syllable weight and therefore are not preferred to bear stress.

The avoidance of stressing the domain-initial checked tone, as seen in Suzhou, is thus explicable by assuming WSP in the stress assignment process. Tone register, in addition to syllable type, can affect stress assignment as well. Low-registered tones are less likely to be stressed domain-initially than high-registered tones. This fact is detectable in both smooth syllables and checked syllables. See Zhenhai's initial tone distribution in trochaic and iambic domains reproduced in Table 24:

Table 24. Initial tone distribution in metrical domains in Zhenhai

S-w		**w-S**	
σ1-σ2		σ1-σ2	
T1 (high register)	Tone	T1 (high register)	Tone
T2 (high register)		T3 (low register)	
T4 (low register)		T4 (low register)	
$T5^{q}$ (high register)		$T6^{q}$ (low register)	

"^{q}" labels a checked tone
Tone = any tone

The tone-stress distribution in Table 24 shows that high-register tones proceeds low-register tones in the shift from iambic stress to trochaic stress. T2 + Tx and $T5^{q}$ + Tx are cases where the stress shift has been completed. T1 + Tx and T4 + Tx are examples of partial switch-over. The most important observation in Table 24, for our current concern, is that domains initiated by low-register tones tend to be more conservative in the stress shift process. Similarly, in Shanghai, although impact of tone register on stress placement is no longer visible on smooth tones, the low-registered checked tone is still shown to be more resistant to trochaic stress than the high-registered counterpart.[11]

The effect of tone register on stress assignment can be explained by De Lacy's theory of tone-stress interaction. As De Lacy (2002) points out, apart from high-sonority nuclei and long vowels, tone can also influence stress placement. Higher-toned syllables attract stress over low-toned syllables. Tone-driven stress systems are reported in languages such as Golin, Tibetan and Vedic Sanskrit (De Lacy, 2002). To deal with the tone-sensitive stress system, De Lacy (2002: 2) proposes that there is a hierarchy of tone prominence, analogous to the sonority hierarchy:

(7) Tone prominence scale

H > M > L

[11] All disyllabic domains initiated by smooth tones are already trochees in Shanghai dialect. Moreover, disyllabic domain initiated by the high-registered checked tone also forms a trochee. Domain initiated by low-registered checked tone is a more conservative case, where iambic stress is still discernable.

In parallel to the scale in (7), the prominence hierarchy of tone register can be worked out in (8):

(8) Tonal register prominence scale

$R_H > R_L$

The prominence of tone register scale can combine with the structural position, foot head (H_D), to form constraints in a fixed ranking as shown in (9).

(9) $*H_D / R_L >> *H_D / R_H$

In (9), $*H_D / R_L$ invariably outranks $*H_D / R_H$. Thus, low-registered foot-heads are predicted to be universally less desirable than high-registered ones. The assumptions in (8) and (9) can account for the tendency to avoid trochaic stress on low-registered tones as found in the Northern Wu dialects.

Moreover, the joint effect of WSP and $*H_D / R_L >> *H_D / R$ further narrows down the possibility for a trochaic stress to fall on a low-registered checked tone, which explains why the disyllabic domain initiated by a low-registered checked tone are most resistant to left-headed stress.

4.3.4 Summary

An examination of the phonological and phonetic details in Zhenhai, Shanghai and Suzhou reveals a rich and complex array of facts. These facts, such as the increased vowel duration in the domain-initial position, avoidance of stress on checked tones, and avoidance of stress on low-registered tones cannot be successfully accounted for by the dual-prominence analysis. It thus appears that the dual-prominence analysis is neither necessary nor desirable. The stress-shift proposal, on the other hand, allows us to account for these facts.

4.4 SUMMARY

To sum up, this chapter refutes the dual-prominence account for the tone sandhi patterns in Zhenhai. This chapter further argues that the claimed peripheral prominence is metrical in nature, which is based on at least three important observations. First of all, peripheral prominence cannot explain the increased vowel duration at domain-initial positon. Secondly, if peripheral prominence is pursued in Zhenhai phonology, the absence of peripheral prominence in $T6^{q}$ + Tx domain is inexplicable. Thirdly, many phonological facts in Shanghai, Suzhou and Zhenhai involve the interaction of syllable weight, tone register and metrical stress, which cannot be theoretically covered by the notion of peripheral prominence.

Finally, the account for tonal behaviors in Northern Wu dialects crucially depends on assumptions on metrical prominence and the historical shift of stress. The case study of Zhenhai persuasively illustrates the need to incorporate stress in the description of tone sandhi facts in some Chinese dialects.

5.0 TONE AND STRESS INTERACTION

We have seen in the previous chapter that tone can condition stress placement. In this chapter we will focus on the effects of metrical system on tones. Roughly speaking, metrical system is playing a bipartite role in conditioning tonal operations. On the one hand, prosodic constituents, such as foot, phonological word, and phonological phrase can constitute the domain of tone sandhi rules (Selkirk, 1980; Chen, 2000). On the other hand, stress effects can manipulate tonal behaviors. While full-fledged tonal contrasts tend to be realized on stressed syllables, tone modification, neutralization, or complete loss can often be induced by lack of stress (Chan and Ren, 1989; Chen, 2000; Duanmu, 1995, 1999; Kennedy, 1953; Yip, 1980).

5.1 METRICAL DOMAIN AND TONE SANDHI DOMAIN

Many generative studies on Chinese tonology have argued that metrical prominence can determine tone sandhi behaviors. Specifically, tone sandhi domains are argued to be metrical domains (Chan, 1985; Chen, 2000; Duanmu, 1995, 1999, among others). The Northern Wu dialects, in particular, exhibit a common feature: unstressed syllables lose their underlying tone whereas the stressed syllable extends its lexical tone rightward to span over the entire domain. Duanmu (1995, 1999) and Chen (2000) have made attempts to reduce tone sandhi domain to the

metrical foot, where tone sandhi rules keep the lexical tone in the foot head and spread it to the rest of the syllable(s).

5.1.1 Shanghai: stress-foot as tone sandhi domain

In the Northern Wu dialects, Shanghai is much more accessible, and far better documented than other varieties in the group. Likewise, the tone sandhi in Shanghai is more thoroughly analyzed from a metrical perspective. In Shanghai dialect, left-headed foot is argued to shape the domain of tone spread at word/compound level (Chen, 2000; Duanmu, 1995).

5.1.1.1 Metrical analysis of word/compound tone sandhi in Shanghai

Essentially, the metrical analysis of Shanghai compound tonology is based on the following metrical rules, where stress clash, stress reduction and optional avoidance of a degenerate foot are claimed to be better predictors of tonal domains than previous analyses (Chen, 2000:307; Duanmu, 1995, 1997):

(1) a. Morpheme level:

Line 0: trochee, left to right, ignore degenerate foot

Line 1: left-headed, unbounded stress

b. Word/compound level: Assign cyclic left-headed stress

c. Phrasal level: Assign cyclic right-headed stress

d. Stress Reduction: Optionally delete Line 1 stress

e. Clash Resolution: Remove the stress column next to a higher column

The basic metrical elements of the rules listed in (1) are borrowed from Halle & Vergnaud (1987). Line 0 and line 1 are both formal tools devised to represent stress and express stress derivation. Line 0 consists of syllables. Line 1 is the stress line, where we find stress represented by asterisks. A degenerate foot is a monosyllabic metrical domain, which is not preferred though not necessarily forbidden.

The rules listed above can correctly predict the tone sandhi pattern of multisyllabic morphemes without hierarchical morphosyntactic structure, for instance, multisyllabic proper names of foreign origin. Examples are given in (2):

(2) a.	HL LH	Underlying tone
	(HL 0)	Tone deletion
	(H L)	Tone spreading
	(pa - li)	Line 0
	x	Line 1
	'*Paris*'	
b.	HL LH LH LH LH	Underlying tone
	(HL 0) (LH 0 0)	Tone deletion
	(H L) (L H L)	Tone spreading
	(ka - li) (-fo' - ɲi - ya)	Line 0
	x x	Line 1
	'*California'*	

The multisyllabic morphemes given in (2) are transliterations of foreign names with flat morphological structure. The underlying tones of foreign names come from the syllables chosen to represent the name in Chinese. In line with the metrical rule set up for morphemes in (1a), we find the metrical derivation in (2): the multisyllabic morpheme forms trochees, from left to right. When there is a degenerate foot, it merges with the preceding foot. When the metrical rules work out the metrical pattern, tone sandhi rules work in each trochaic domain in the familiar way: the

initial tone is realized over the entire domain, in a one-to-one, left-to-right manner. In the second domain in (2b), the initial tone re-association cannot reach the last syllable; therefore, the last syllable gets a default low tone. The left-headed metrical domain in (2) is a correct prediction of the tonal domain.

The strongest evidence for cyclic application of stress rule (1b) and clash avoidance (1e) is shown by tone sandhi in compounds, which is illustrated by the following examples:

(3) a. Underlying tone HL LH LH → Tone sandhi (H L L)

*(tçi) (ts^{h}z - pã) → (tçi ts^{h}z - pã)

Word stress x x Clash resolution x

Compound stress x

'*chicken wing*'

b. Underlying tone LH HL HL

Tone sandhi (L H) (HL) Tone sandhi (L H L)

(lu - sõ) (t^{ɦ}ã) or (lu - sõ t^{ɦ}ã)

Word stress x x Stress reduction x

Compound stress x

'*Russian soup*'

In the examples of (3), following the common assumption in metrical phonology, word boundaries are projected as metrical boundaries (Duanmu, 1995: 231). In the case of (3a), the two word boundaries provide two metrical domains with stress assigned on the first and the second syllable, which creates a classic stress clash. To resolve stress clash, the stress in the second foot is deleted, then the unfooted syllables merge into the first one. Consequently, the de-stressed syllable *ts^{ɦ}z* loses its underlying tone, and gets a low tone from tone spreading. Just as in (3a), the word boundaries in (3b) create two domains. However, the foot structure of (3b) does not produce stress clash, therefore (3b) can either keep two left-headed feet, or merge the monosyllabic foot with the preceding one according to the optional stress reduction rule stated in (1d): a degenerate foot is not preferred although it is not forbidden (Duanmu, 1995: 231).

In accordance with the observation from the examples in (3), Duanmu (1995) argues that Shanghai has a metrical stress system, and that stress at word and compound level is left-headed, otherwise the contrast between (3a) and (3b) is unexpected. Moreover, the tonal domain in Shanghai is a metrical domain. As seen, the metrical domain can correctly predict the domain for tone sandhi.

5.1.1.2 Problems of metrical analysis of Shanghai compound tonology

The metrical account presented above rests crucially on the left-headed metrical structure in Shanghai. However, as argued previously, feet in Shanghai are not strictly left-headed. Zhu (1996, 2006) points out that a foot is right-headed, when the first syllable in the foot bears a low-registered checked tone ($T5^{q}$). See the following contrasting examples demonstrated by Zhu (1996):

(4) a. Underlying tone LH LH LH → Tone sandhi (L H L)
*(gou) (er- duo)[1] (gou er-duo)
Word stress x x Clash resolution x
'*dog ear*'

b. Underlying tone LH **LH**^{q} LH
Tone sandhi (LH) (L H)[2] Tone sandhi (L H L)
(gou) (bi - tou) or (gou bi-tou)
Word stress x x Stress reduction x
'*dog nose*'

c. Underlying tone HL LH LH
Tone sandhi (H L) (LH) Tone sandhi (H L L)
(san-hao) (men) or (san-hao men)
Word stress x x Stress reduction x
'*gate No. 3*'

[1] Since IPA transcriptions are not available in the data source cited, the Shanghai syllables in (4) are represented by Pinyin, which is the romanization of Chinese writing system based on Beijing phonology.
[2] Note that the sandhi rule $T5^{q}$ (LH^{q}) + Tx→ L + LH does not apply in this case. A regular rightward tone spreading and one-to-one tone association occur in the second foot.

d. Underlying tone	**LH**^{q} LH LH		Tone sandhi	(L L **LH**)^{3}
	*(liu - hao) (men)	→		(liu- hao men)
Word stress	x x		Clash resolution	x
	'*gate No. 6*'			

Examples in (4a) and (4c) are assigned left-headed stress, where stress clash arises in the case of (4a), thus de-footing is triggered by the obligatory rule of clash resolution. Example in (4c) is metrically unproblematic. Alternatively, a single foot can be created by de-footing when the optional stress reduction rule applies. Since the metrical domain constitutes the tone sandhi domain in the current analysis, (4c) have two alternative tone readings.

The matter becomes a bit more complicated in (4b). In the metrical approach proposed by Duanmu (1995) and Chen (2000), the feet in (4b) is presumably left-headed. Since (4b) and (4a) are both right-branching and have the identical initial tone, one may expect to see the same tone sandhi domain and same tone sandhi result in (4b) and (4a). (4a) has only one reading "L H L", (4b) have two alternative readings. The extra reading "LH L H" suggests that (4b) must have split into two tone domains, where the underlying tone is kept in the first domain, and the second domain undergoes tone spreading and re-association. Since the tone domain is determined by the metrical domain, the extra sandhi pattern attested in (4b) reveals that there can be two corresponding metrical domains. Thus, abutting stresses should not occur in (4b) in order to allow the occurrence of two stress domains in a right-branching structure ([σ[σ σ]]).

Based the tone sandhi patterns as shown in (4b), and the detailed phonetic analysis on the duration pattern of Shanghai disyllable words,[4] Zhu (1996, 2005) concludes that metrical structure in Shanghai is not uniformly left-headed. A foot with an initial T5^{q} syllable is assigned

[3] Zhu (1996) adopts a "L L H" sandhi pattern for (4d) when stress reduction creates a single foot. However, Zee and Maddieson (1979) reports "L L LH" for trisyllabic compounds with initial T5^{q} syllable, where the initial tone is attracted to the last syllable. Zee and Maddieson's pattern is adopted for the current discussion, since it is more widely agreed and cited.

[4] See §3.3.2 for details.

a right-headed stress. This claim is further supported by the tone sandhi pattern found in (4d). If left-headed stress is followed, (4d) as a left-branching compound can support two metrical domains. Optionally, foot merger can lead to a single domain when stress reduction applies. However, only one reading is attested for (4d). A [[T5^{q} σ]σ] compound supports only one tone sandhi domain, suggesting that foot merger must apply as a result of abutting stresses, which is a strong evidence for right-headed stress in a [T5^{q} σ] domain.

In conclusion, the tone sandhi contrast between (4a) vs. (4b), (4c) vs. (4d) supplies strong support for Zhu's claim that right-headed stress is assigned to a foot with an initial low-register checked tone (T5^{q}).

As Bao (2003, 2004) points out, another crucial problem of Duanmu's metrical analysis of Shanghai is that the size of the metrical foot varies greatly, ranging from monosyllables to polysyllabic compounds. Consequently, variable size of the foot can be produced. See the compounds in (5) for illustration (Bao, 2004: 882):

(5) a. [[(σ̲σ)](σ̲)]→ (σ̲σσ)

b. [[[(σ̲σ)](σ̲)](σ̲)]→ (σ̲σσσ)

c. [[(σ̲σ)][(σ̲σ)]]→ (σ̲σσσ)

Following the convention of Bao (2003), parentheses are used to delimit the domain of tone sandhi in the examples of (5), and the head syllable in a foot is marked with an underscore symbol. According to Bao (2004), all compounds in (5) exhibit a single sandhi domain regardless of the internal morphosyntactic structure. (5a) and (5b) are left-branching compounds, where there is no pressure from clash resolution to obligatorily create a single metrical domain through foot merger. In view of the optional stress reduction rule posited by Duanmu (1995, 1997) and Chen (2000), (5a) can optionally create a single foot though foot merger, because a

monosyllabic foot is not preferred although it is not completely blocked. Similarly, (5b) can be optionally parsed into a single domain, creating a foot of four syllables. However, it is still not clear why (5c) supports a single domain. From a metrical point of view, (5c) has the optimal morphosyntactic structure, since the morphosyntactic structure produces the metrically perfect structure. It is unexpected to see (5c) creates a single tone sandhi domain. Bao (2004:881) points out that "the variable size of the foot deprives the metrical analysis of any explanatory force. Although the so-called unbounded foot is allowed in some parametric metrical theories …, I shall adopt a stricter notion of metrical foot, … where the size of the foot is capped at no more than two syllables or moras".

According to the metrical rules listed in (1), the compounds such as (5a) and (5b) should have two tonal readings because stress reduction rule can be optionally applied. However, language data in published studies (Xu *et al.*, 1981; Zee and Maddieson, 1979) do not confirm that compounds with the morphosyntactic structure of (5a) and (5b) have two readings. The data collected in these studies, with left-branching structures as shown in (5), support that there is a single tone sandhi domain. If this claim is further proved to be accurate, then metrical structure will not be the key to the derivation of the tone sandhi domain as previously thought.

Instead, Bao (2003:163) argues that compound tone sandhi in Shanghai "can be readily accommodated in an account that makes no reference to metrical notions". In an end-based account, the tone melody of a tone sandhi domain is largely determined by the tone of the first syllable of a polysyllabic compound demarcated by an end-setting parameter as shown in (6). Essentially, the rule in (6) parameterizes the syntax-phonology mapping in Shanghai. It requires the left edge of a syntactic structure (i.e. Lex^{max}, which stands for the maximal projection of a

lexical category N, A and V) to coincide with the edge of constituent in prosodic structure (Selkirk and Shen, 1990):

(6) {left, Lex^{max}}

In conclusion, the metrical account of Shanghai compound tonology faces some problems, and therefore remains controversial. There are metrical effects which reveal the role of metrical structure in the tone sandhi of Shanghai, such as the tone spreading and re-association delimited by the feet in foreign proper names. However, the details of the metrical analysis often induce disagreements. And, the predictions the metrical account makes are often claimed to be not supported by empirical language data.

5.1.2 Metrical analysis of tone sandhi in Xiamen and Danyang

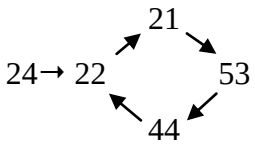
Xiamen, also called Amoy, is a southern Min dialect. It is mainly spoken in Xiamen, a region in southern part of Fujian Province. Closely related varieties of Xiamen are also spoken in Taiwan. Xiamen presents a fairly complex tone system. Like other Chinese dialects, Xiamen divides syllables into smooth type and checked type.

(7) Smooth tones	44	H	Checked tones	4^{q}	H^{q}
	24	MH		32^{q}	ML^{q}
	22	M			
	21	ML			
	53	HM			

Xiamen tone sandhi distinguishes two positons: domain-final position and “everywhere else” (Chen, 2000: 432). Lexical tones are preserved in domain-final positon, whereas “everywhere-else” is the sandhi positon where tone substitution happens, creating a pattern widely known as the Min Circle:

(8) Xiamen Tone sandhi

smooth tones



checked tones $4^{q} \rightarrow 21^{q}$

$32^{q} \rightarrow 4^{q}$ for syllables ending in p, t, k
53^{q} for syllables ending in ʔ

To focus on the theme of the current discussion, the nature of the tone substitution pattern shall not concern us here. With regard to the tone sandhi domain of Xiamen, Duanmu (1995) proposes that the tonal domain is subject to metrical effects. Specifically, Duanmu exploits the same set of metrical rules proposed for Shanghai in Xiamen tone-stress relation,[5] except that metrical foot in Xiamen is right-headed. On this account, Xiamen's metrical system is the exact mirror image of Shanghai.

However, the metrical approach does not fit into the tonology of Xiamen gracefully. The metrical approach encounters exactly the same problems as it does in Shanghai, and succeeds no better. The following data from Chen (2000:436) highlight the problem that the metrical approach often makes false predictions of Xiamen tonal sandhi domain.

(9) a.

	*(ts'ian- li)(be)	→	(ts'ian-li be)	Stress resolution
Word stress	x x		x	
Compound stress	x			
	'*winged steed*'			

b.

	(ts'u) (ting-bin)	or	(ts'u ting-bin)	Optional stress reduction
Word stress	x x		x	
Compound stress	x			
	'*housetop*'			

c.

	*(ang) (tsng-t'au)	→	(ang tsng-t'au)	Stress resolution
Word stress	x x		x	
Compound stress	x			
	'*red brick*'			

[5] See the rules listed in (1) in this chapter.

Left-branching compounds like (9a) create stress clash. Clash resolution, therefore, de-stresses the second syllable. As a result, there is only a single foot in (9a), and thus a single corresponding tone sandhi domain. Data in (9b) has two alternative readings. Like Shanghai, an optional stress reduction rule must be posited in Xiamen to deal with this situation. Otherwise, the alternative reading cannot be accounted for. Data in (9c) has the same morphosyntactic structure as (9b). Thus, two alternative readings are expected for the compound in (9c). Chen (2000: 437) points out that the prediction is not true: the compound in (9c) only supports a single domain. Importantly, (9c) is not an idiosyncratic case. Similar examples can be found easily in Xiamen (Chen, 2000).

Chen (2000) concludes that Xiamen tone sandhi does not yield nicely to a metrical analysis, and thus resumes the theory of syntax-phonology interface, where an end-based account was first proposed in Chen (1987). In the end-based approach, the end-setting rule shown in (10) derives the phonological phrase, i.e. the domain of tone sandhi, by demarcating the right edge of a maximal projection (X^{max}).

(10) {Right, X^{max}}, X^{max} not an adjunct

In a comparative sense, the application of the metrical approach to the derivation of tonal domain in Danyang, another Wu dialect, seems to produce more encouraging results. However, as we shall see below, the metrical analysis of the data in Danyang still shows ad hoc properties.

Lü (1980) first presents a detailed description and analysis of the tones of Danyang. Details aside, in addition to the lexical tones, Danyang has six word melodies that must be specified independently of the lexical tones:

(11) Danyang word melodies

	Basic melody	disyllabic	trisyllabic	tetrasyllabic
a.	H	H-H	H-H-H	H-H-H-H
b.	M	M-M	M-M-M	M-M-M-M
c.	L	L-L	L-L-L	L-L-L-L
d.	HL	HL-L	HL-L-L	HL-L-L-L
e.	LH	LH-H	LH-H-H	LH-H-H-H
f.	HLH	HL-LH	HL-HL-LH	HL-HL-HL-LH

The initial syllable of a polysyllabic compound selects a melody from the inventory of the basic word melody as shown in (11). The word melody is then extended over the entire compound in the pattern demonstrated above.

Chen (2000: 328) suggests a metrical approach to determine the tonal domain in Danyang by exactly the same metrical rules applied in the analysis of Shanghai. The clash-induced stress deletion and optional application of stress reduction make correct predictions of the tone sandhi pattern in Danyang.

However, a problem is encountered when the metrical account is applied to tetrasyllabic compounds as illustrated by the examples in (12) (Chen, 2000:334).

(12) a. [(σ)[(σσ)]](σ) → [(σσσ)](σ) → (σσσ)(σ)

b. [[(σσ)](σ)](σ) → [(σσ)(σ)](σ)→(σσ)(σ)(σ)

For (12a), clash-induced foot merger correctly predicts the attested tone sandhi pattern. However, in (12b), clash is unexpectedly tolerated between the third and the fourth syllable. To stick with the stress-based account, Chen (2000) resorts to directional asymmetry which is stated as follows:

(13) Asymmetric Clash Resolution

a. Leftward clash must be resolved by de-stressing
b. Rightward clash is tolerated

The asymmetric clash resolution rule can lead to two consequences. First, the application of clash resolution is no longer obligatory. Second, the clash resolution is cyclic in some compounds but noncyclic in others. Since the metrical structure at compound level is subject to morphosyntactic structure, multiple morphosyntactic brackets will lead to cyclic stress assignment and stress clash configuration. The asymmetrical clash resolution rule may block the cyclicity. These problems make the asymmetric clash resolution an ad hoc rule.

5.1.3 Summary

This part has reviewed the application of metrical approach to delimit the tonal domain of a couple of Chinese dialects. On the one hand, metrical approach has advantages over the end-based analysis. For instance, it makes an explicit connection between prominence and tone stability. On the other hand, such attempts have been met with limited success, and often show ad hoc properties.

5.2 STRESS DRIVEN TONE MODIFICATION

Although the metrical analysis of tonal domain derivation has been not very successful, it is undeniable that tone and stress can co-exist in some dialects of Chinese, and interact in a systematic fashion. For instance, metrically strong positons are characterized by the retention of tone, ability to attract tone or the ability to determine the shape of a neighboring tone, as instantiated by the tone sandhi behaviors of dialects in the Northern Wu group. On the other

hand, tone can affect stress-placement. For example, stress is not attracted to syllables associated with a low tone or short tone, as seen in Zhenhai.

5.2.1 Stress-induced tone neutralization: an example from Cantonese

It has long been noted that stress may condition tone neutralization. The second part of this chapter will present a case of Cantonese where tone neutralization is induced by tone-stress interaction.

5.2.1.1 Tone-stress interaction in OT

De Lacy (2002) systematically examines the relationship between tone and prosodic positions in OT. His study shows that prosodic heads prefer higher tones over lower tones, whereas non-heads exhibit the opposite preference. Likewise, tone can also influence main stress placement. For instance, stress tends to be attracted to higher-toned syllables over the lower-toned ones.

De Lacy (2002) defines the tone-dependent stress placement as "tone-driven stress". To deal with tone-driven stress in the framework of OT, De Lacy proposes a tone prominence hierarchy. In this scale, high tone is more prominent than lower tones (De Lacy, 2002:2):

(14) Tonal prominence scale

H > M > L

(15) a. *HD/ L≫ *HD/ M

b. *Non- HD/ H ≫ *Non- HD/ M

Constraints listed in (15) combine the tonal prominence scale with the structure positions, such as foot head (Hd) and foot non-head (non-Hd). The constraints are in a universally fixed ranking. Therefore, heads with low tones always incur more significant violations than mid-

toned ones. In the same spirit, constraint *Non- H_D/ H punishes high-toned non-heads. By constraints in (15), the OT based tone-stress interaction claims affinity between foot heads and high tone, and foot nonheads and low tone.

A similar tone-stress interaction is found in Cantonese. In Fengkai dialect, a variety of Cantonese spoken in the west part of Guangdong (Canton) Province, the avoidance of high tone in nonheads has been attested. Therefore, stress-conditioned tone neutralization emerges in the nonhead position.

5.2.1.2 Stress-conditioned tone neutralization in Fengkai Cantonese

Cantonese is generally believed to lack tone sandhi. However, Hou (2011a) presents a detailed description of stress-conditioned tone sandhi in Fengkai Cantonese.

Fengkai Cantonese is spoken in the mid-west region of Guangdong (Canton) Province, and the adjacent regions in Guangxi Province, such as Wuzhou and Yulin. Fengkai presents a very complicated tone system. As reported in Hou (2011a), the tone system of Fengkai retains the four Middle Chinese tonal categories, which are evenly split into a high register and a low register. Fengkai exhibits an intricate inventory of checked tones, as shown in Table 25.

Table 25. Tone system in Fengkai

Middle Chinese Tone Category	Even (Smooth tone)	Rising (Smooth tone)	Falling (Smooth tone)	Entering (Checked tone)
High register	44 (T1)	52 (T3)	32 (T5)	55^{q} ($T7^{q}$)
				32^{q} ($T8^{q}$)
Low register	24 (T2)	242 (T4)	21 (T6)	34^{q} ($T9^{q}$)
				21^{q} ($T10^{q}$)

"^{q}" labels a checked tone

Four tone categories, referred to by the traditional terms as Even, Rising, Falling and Entering, are fully kept in Fengkai. Apart from the commonly seen division of the high and low

tone registers, checked tones in Fengkai show further split conditioned by vowel height. In each tone register, the checked tone with a high vowel is higher-pitched than the checked tone with a non-high vowel. As shown in Table 25, this multiple splitting leads to an inventory of four checked tones in Fengkai.

Apart from a complicated tone system, Fengkai is reported to have iambic feet (Hou, 2011a, 2011b). Hou (2008) gives a detailed measurement of syllable duration in disyllabic words of Fengkai, as shown in Table 26:

Table 26. Syllable durations in disyllabic words of Fengkai

	Duration of the first syllable (ms)	**Duration of the second syllable (ms)**
T1+T1	137	220
T2+T2	164	234
T3+T3	153	179
T4+T4	172	182
T5+T5	154	251
T6+T6	158	216
$T7^{q}$+$T7^{q}$	51	77
$T8^{q}$+$T8^{q}$	62	107
$T9^{q}$+$T9^{q}$	84	182
$T10^{q}$+$T10^{q}$	97	190

"^{q}" labels a checked tone

All the disyllabic words measured in Hou (2008) are composed of two identical tones in order to filter out the possible influence on syllable duration caused by tone category. It is apparent in the above table that the second syllable is always considerably longer than the first one.

Except for the consistent correlation between stress and duration, the existence of right-headed stress in Fengkai is further evidenced by the role stress plays in tone sandhi processes. The details are presented in Table 27.

Table 27. Fengkai tone sandhi patterns in disyllabic domain

σ2 / σ1	Smooth Tones						Checked Tones			
	T1	T2	T3	T4	T5	T6	$T7^{q}$	$T8^{q}$	$T9^{q}$	$T10^{q}$
T1 (44)	32-T1	32-T2	32-T3	32-T4	32-T5	32-T6	$32\text{-}T7^{q\,q}$	$32\text{-}T8^{q}$	$32\text{-}T9^{q}$	$32\text{-}T10^{q}$
T2 (24)	21-T1	21-T2	21-T3	21-T4	21-T5	21-T6	$21\text{-}T7^{q}$	$21\text{-}T8^{q}$	$21\text{-}T9^{q}$	$21\text{-}T10^{q}$
T3 (52)	32-T1	32-T2	32-T3	32-T4	32-T5	32-T6	$32\text{-}T7^{q}$	$32\text{-}T8^{q}$	$32\text{-}T9^{q}$	$32\text{-}T10^{q}$
T4 (242)	21-T1	21-T2	21-T3	21-T4	21-T5	21-T6	$21\text{-}T7^{q}$	$21\text{-}T8^{q}$	$21\text{-}T9^{q}$	$21\text{-}T10^{q}$
T5 (32)										
T6 (21)										
$T7^{q}$ (55^{q})	32^{q}-T1	32^{q}-T2	32^{q}-T3	32^{q}-T4	32^{q}-T5	32^{q}-T6	$32^{q}\text{-}T7^{q}$	$32^{q}\text{-}T8^{q}$	$32^{q}\text{-}T9^{q}$	$32^{q}\text{-}T10^{q}$
$T8^{q}$ (32^{q})										
$T9^{q}$ (34^{q})	21^{q}-T1	21^{q}-T2	21^{q}-T3	21^{q}-T4	21^{q}-T5	21^{q}-T6	$21^{q}\text{-}T7^{q}$	$21^{q}\text{-}T8^{q}$	$21^{q}\text{-}T9^{q}$	$21^{q}\text{-}T10^{q}$
$T10^{q}$ (21)										

"^{q}" labels a checked tone

Based on a detailed acoustic analysis of tones in Fengkai by Hou (2011a), all the disyllabic tone sandhi patterns in Fengkai are given in Table 27. The shaded area marks the combination where no tone sandhi process is found. The sandhi patterns shown in Table 27 demonstrate two regularities. First, tone sandhi only affects tones in the first syllable; the second tone remains unchanged. Second, high tones in the word-initial position are altered to low tones. For instance, T1 with tone value 44 is lowered to 32 in word-initial position. Low tones are not affected word-initially.

Since word-initial high tone is prohibited, the ten lexical tones in Fengkai are no longer contrastive at word-initial position. Only two tones, with the value of 32 and 21, can appear word-initially, being smooth or checked. Since T5, T6, $T8^{q}$ and $T10^{q}$ are underlyingly with the tone value of 32 or 21, they do not seem to undergo further tone modification at word-initial position. However, as a consequence of tone lowering, the underlying 32 and 21 has merged with the derived ones.

Tonal stability is exhibited in the second syllable of disyllabic words, where the lexical tone is always faithfully retained. This is not surprising, if the tonal behavior of disyllabic words

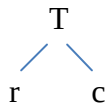
in Fengkai is viewed as influenced by stress. Specifically, when the tone-stress connection is assumed, the disyllabic tone sandhi domain can be parsed into an iambic foot. As a result, the tone stability and the tone modification in disyllabic words can be readily explained as conditioned by stress effects.

In §5.2.1.3, attempts will be made to account for the tone stability and tone lowering instantiated by the Fengkai dialect. This time, I will exploit De Lacy's OT theory of tone-stress interaction.

5.2.1.3 OT analysis of stress-conditioned tone neutralization in Fengkai

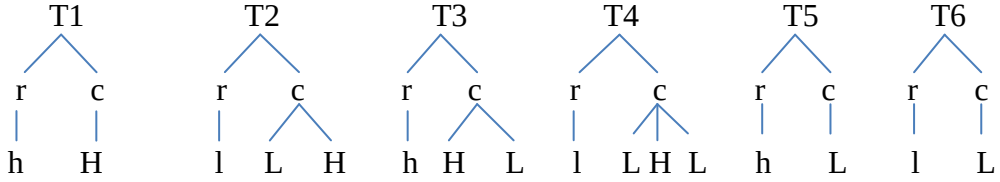
Following Bao (1999) and Chen (2000), we assume that a tone consists of a tone register node and a contour node in a way shown by the following representation:

(16) Tone representation

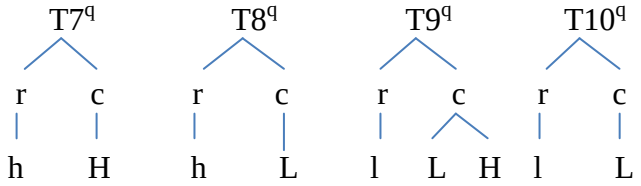


The "r" node in (16) specifies a tone register, whereas "c" specifies the tonal contour of a given tone, for instance HL or LH. Thus, the tone inventory and the tone sandhi rules in Fengkai can be represented as follows:

(17) Smooth tones

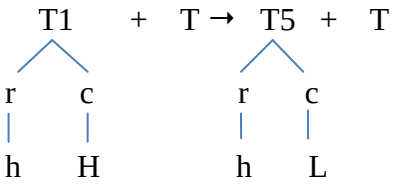


(18) Checked tones

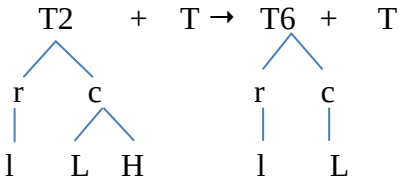


(19) Tone sandhi rules

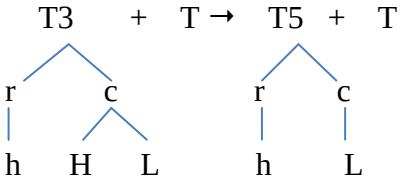
T1 (44) + T → T5 (32) + T



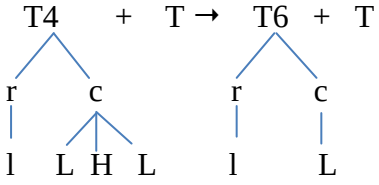
T2 (24) + T → T6 (21) + T



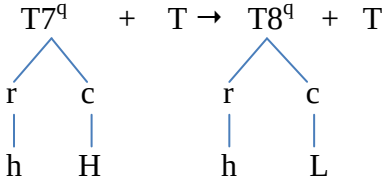
T3 (52) + T → T5 (32) + T



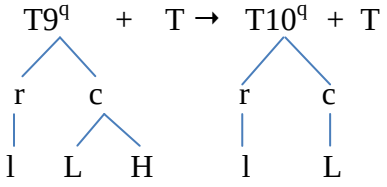
T4 (242) + T → T6 (21) + T



T7q (55q) + T → T8q (32q) + T



T9q (34q) + T → T10q (21q) + T



The inclusion of tone register in the tone representations of Fengkai is based on the observation that tone register has a crucial influence over the tone neutralization in the disyllabic domain. To put it more explicitly, the word-initial high tone in Fengkai is uniformly lowered to a certain tone value according to the tone register it lexically associates with. As seen in Table 27, high tones in the high register are lowered to 32, while high tones in the low register are lowered to 21.

In OT terms, neutralization is produced when a markedness constraint outranks relevant faithfulness constraints. When it comes to the tone neutralization in Fengkai, we need a markedness constraint *Non-H$_D$/ H to punish high tones in unstressed position. Thus, the lack of high tone at word initial position can be handily explained. Moreover, we still need an undominated faithfulness constraint to manifest the fact that high tone remains unchanged at word-final (stressed) position in Fengkai. With the formulation of the following constraints, we will be able to cover the Fengkai tone sandhi processes in OT:

(20) Ident (T): If mora x bears tone T in the input, then the output correspondent of x bears T.

Ident (R): If a tone is in tone register R in the input, then the output correspondent of the tone in still in the R tone register.

*Non-H$_D$/ H: No high tones in the non-heads of feet (De Lacy, 2002).

Iamb: Every foot is right-headed.

FtBin: Every foot is binary at syllabic level (McCarthy and Prince, 1993).

Parse-σ: Every syllable is contained inside a foot (Prince and Smolensky, 1993).

To ensure avoidance of register modification for Fengkai tones, Ident (R) must be undominated. The constraint *Non-H$_D$/ H should also be undominated, because *Non-H$_D$/ H is never violated. Iamb, FtBin and Parse-σ are constraints that make sure every disyllabic word

parses into a binary iambic foot. There is no evidence that Iamb is ever violated, therefore, Iamb, FtBin and Parse-σ are not outranked by other constraints. The ranking of these constraints is explicated by the tableau in below:

Tableau 4. T1(44) + T → T5 (32) + T

T1 T r c h H	*Non-HD/ H	Iamb	FtBin	Parse-σ	Ident (R)	Ident (T)
☞ a. (T5 ′T) r c h L						*
b. T1 T r c h H				*!		
c. (T5 ′*T*) r c h L						**!
d. (′T1 T) r c h H		*!				

In the above tableau, foot boundaries are delimited by parentheses, and stress is indicated with an acute accent (′). T represents any tone, therefore, T1 + T indicates a disyllabic words with an initial T1 following by a random tone. Since tone modification in Fengkai does not affect the tone in foot head position, the tonal specifications of the second tone is therefore not shown in the tableau. Candidate (c) demonstrates a case where the lexical tone in the head position is

not preserved, which is distinguished from other candidates by an underlined and italicized "*T*". Candidate (c) is therefore ruled out by two violations of Ident (T). The joint effect of FtBin, Iamb, and Parse-σ requires that the disyllabic word parses into an iambic binary foot. Thus, candidate (b) is ruled out. Candidate (d) fatally violates the right-headed stress constraint Iamb.

The constraint ranking shown in Tableau 4 can successfully account for the absence of high tones in nonhead position caused by the lack of stress. To illustrate why low tones in nonhead position are not affected, the evaluation of the candidates of T5 + T is given below:

Tableau 5. T5(32) + T→ T5(32) + T

T5 T r c h L	*Non-H$_D$/ H	Iamb	FtBin	Parse-σ	Ident (R)	Ident (T)
☞ a. (T5 ˈT) r c h L						
b. (T1 ˈT) r c h H	*!					*
c. (T5 ˈ*T*) r c h L						*!
d. (T6 ˈT) r c l L					*!	*

Candidate (a) is the winning form, since it does not incur any violation of the constraints. Candidate (d) is associated with a low tone in nonhead position, it does not incur violation of *Non-HD/ H, but it violates Ident (R) since the low tone is not realized in the original tone register. Candidate (c) is removed from the competition at the point it makes changes to the right tone. Candidate (b) is ruled out by *Non-HD/ H, since it has a high-toned nonhead.

To summarize, the theory presented in De Lacy (2002) allows for an easy account of metrically influenced tone systems as instantiated by Fengkai. The influence of foot head and nonhead on tone is clearly manifested, when tone-head constraints outrank tone preservation constraints (De Lacy, 2002:28).

However, the validity of this OT grammar needs to be tested with a larger corpus. For instance, the spreading of the stressed tone in Shanghai may violate both *Non-HD/H and *HD/L. See the sandhi rule for T3 + Tx in Shanghai listed below:

(21) ′T3 + Tx ′LH + Tx → ′L + H

As we already know from the discussion of Shanghai tonology, T3 + Tx in Shanghai creates a trochaic foot where the tone sandhi rule spreads the word-initial tone over the entire domain in a left-to-right, one-to-one fashion. As seen in (21), the foot head (which is indicated by an acute accent) keeps a low tone. At the same time, the nonhead syllable acquires a high tone (H) from tone spreading. The redistribution of the word-initial tone violates both *Non-HD/ H and *HD/ L. To address such a tone operation, De Lacy's OT grammar may need some amendment.

5.3 SUMMARY

This chapter has looked into the metrical effects on tonal operations in some Chinese dialects. The metrical analysis of tonal behaviors has produced mixed results. On the one hand, metrical approach was shown to be technically inadequate, in that it often cannot correctly parse the metrical domain into a tonal domain. On the other hand, as seen in the second part of this chapter, tone-stress interaction in OT offers a preliminary, but exceedingly simple interpretation of stress-driven tone modification. However, the validity of OT grammar needs to be further improved before it can cover more language data.

6.0 CONCLUSION

This dissertation has reviewed the major work on prominence-related issues in Chinese, much of which is already familiar to linguists of Chinese phonology. It was shown that when the phonetic and phonological evidence is examined, some previous conclusions on phonological prominence in Chinese are invalid. In particular, this dissertation offers strong evidence to refute the generalized stress view of Standard Mandarin, accentual prominence in New Chongming and the dual-prominence claim in the Zhenhai dialect.

We have posited two main types of tone-stress interaction. First, we have shown that tone and stress are two independent phonological properties that can co-exist in Chinese. Underlying tone contrasts are preferentially preserved in prominent positions, although the lack of prominence can often induce tone modification, tone loss, and tone neutralization.

Secondly, tone can condition stress placement. Stress placement in the Northern Wu dialects suggests that stress assignment depends on tone. The distribution of stress in these dialects shows that stress is not found in low-toned syllables and short-toned syllables. Therefore, Northern Wu dialects present a case where stress placement is not only tone-sensitive, but also quantity-sensitive.

The original contribution of this dissertation has been in two areas. First, it makes a contribution to what is known about the typology of prominent positions at word level in Chinese. To my knowledge, no published work has investigated word-level strong positions in

Chinese with an explicit categorization of phonological prominence. This study has examined three types of prominence, i.e., metrical prominence, pitch-accent prominence and peripheral prominence. The analysis of the data has proved that metrically strong position is the only type of prominence attested so far at word level in Chinese. Second, this study makes a contribution to the understanding to the tone-stress relation. By analyzing data from Northern Wu dialects and a Cantonese variety, this dissertation revealed that tone properties can constraint the process of stress evolution and that stress can trigger tone modifications. Meanwhile, it points out that the current metrical approach, either rule-based or OT-based, needs to be further improved before it can more revealingly interpret tone-stress relation.

Some research topics may arise from the discussion in this dissertation. Among them, two are particularly intriguing. First, what is the motivation of the on-going stress shift? Second, how has a long history of the contact between Chinese dialects affected the prosodic phonology of these dialects? I leave these topics for future work, in view of the fact that they deserve long and independent studies.

The discussion presented in this dissertation is limited to some representative Chinese dialects. With more than a decade of research in both phonological and phonetic properties, these dialects are now more accessible than others. The lack of studies with phonetically sufficient details of Chinese dialects greatly limits the practice of more principled research on prominence of Chinese in general. At the same time, even the analyses of the phonetically well-studied Chinese dialects, the Shanghai dialect, for instance, can vary considerably because of the great discrepancy in the first-hand language data collected.

Prominence and prominence-related effects on phonology of Chinese are still an understudied field. Prominence-related issues are, by and large, not well understood. While

several cases have been discussed in this dissertation, no doubt many more are yet to be discovered. Observations and studies on prominence phenomena in Chinese made in this dissertation are provisional and await future research.

BIBLIOGRAPHY

Abercrombie, D. (1991). *Fifty years in phonetics*. Edinburgh, U.K.: Edinburgh University Press.

Babic, Zrinka. (1988). Accent systems in Croatian dialects. In Harry van der Hulst and Norval Smith (ed.) *Autosegmental studies on pitch accent*. Dordrecht: Foris. 1-10.

Bao, Zhiming. (1999). *The structure of tone*. Oxford: Oxford University Press.

Bao, Zhiming. (2003). Tone, accent and stress in Chinese. *Journal of Linguistics* 39. 147-166.

Bao, Zhiming. (2004). Accentualism in Chinese. *Language and Linguistics* 5. 863-889.

Barnes, Jonathan. (2001). Domain-initial strengthening and phonetics and phonology of positional neutralization. In Mako Hirotani (ed.) *Proceedings of NELS 32*.

Beckman, Jill N. (1995). Shona height harmony: Markedness and positional identity. In J. Beckman, L. Walsh Dickey and S. Urbanczyk (ed.) *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*. 53-75.

Beckman, Jill N. (1997). Positional faithfulness, positional neutralization and Shona vowel harmony. *Phonology* 14. 1-46.

Beckman, Jill N. (1998). *Positional faithfulness*. Ph.D. dissertation, University of Massachusetts, Amherst.

Beckman, Mary. (1986). *Stress and non-stress accent*. Dordrecht, Holland: Foris Publications.

Blevins, Juliette. (1994). A place for lateral in the feature geometry. *Journal of Linguistics* 30. 301-348.

Bloomfield, Leonard. (1933). *Language*. New York: Holt.

Bybee, Joan L., Paromita Chakraborti, Dagmar Jung, and Joanne Scheibman (1998). Prosody and segmental effect: some paths of evolution for word stress. *Studies in Language* 22. 267-314.

Byrd, Dani. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica* 57. 3-16.

Cao, Jianfen. (1989). Putonghua shuangyinjiexi de shichang fenbu tezheng fenxi [An analysis of the patterns of duration distribution in disyllabic words in Standard Mandarin]. In *Proceedings of Conference on Acoustics, Speech and Signal Processing* 4. Beijing, China.

Cao, Jianfen. (1992). Temporal structure in disyllabic word frame: evidence for relational invariance and variability from Standard Chinese. In *Proceedings of ICSLP 1992*. Alberta, Canada.

Cao, Jianfen. (1994). The effects of contrastive accent and lexical stress on temporal distribution in a sentence. In *Proceedings of ICSLP 1994.* Yokohama, Japan.

Cao, Jianfen. (1995). Liandu biandiao yu qingzhong duili [Tone sandhi and stress contrast]. *Zhongguo yuwen* 4. 312-320.

Cao, Jianfen. (2007). Cong yuyin hecheng kan hanyu de zhongyin jiegou [Viewing the stress pattern from the perspective of speech synthesis]. In Jianfen Cao (ed.) *Xiandai yuyin yanjiu yu tansuo* [Studies on Chinese phonetics]. Beijing: Commercial Press. 273-288.

Cao, Jianfen. (2008). Fayin zengqiang yu suojian [Strengthening and reduction of articulation]. *Zhongguo yuyin xuebao* 1. 90-96.

Chan, Marjorie and Hongmao Ren (1989). Wuxi tone sandhi: from last to first syllable dominance. *Acta linguistica Hafniensia* 21. 35-64.

Chan, Marjorie. (1995). An autosegmental analysis of Danyang tone sandhi: some historical and theoretical issues. In Eric Zee (ed.) *Studies of Wu dialects*. Hong Kong: Chinese University of Hong Kong. 145-184.

Chao, Yuenren. (1928). *Studies in the modern Wu dialects*. Beijing: Qinghua Research Institute.

Chao, Yuenren. (1968). *A grammar of spoken Chinese*. Berkeley and Los Angeles: University of California Press.

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

Chen, Ning-ping. (1987). Xiandai hanyu mingcilei de kuoda [The enlarged nominal categories in Modern Chinese]. *Zhongguo yuwen* 5. 379-389.

Cho, Taehong and Sun-Ah Jun (2000). Domain-initial strengthening as enhancement of laryngeal features: Aerodynamic evidence from Korean. *UCLA Working Papers in Phonetics* (2000). 57-70

Cooper, Andre Maurice. (1991). *An articulatory account of aspiration in English*. Ph.D. dissertation, Yale University.

De Lacy, Paul (2002). The interaction of tone and stress in Optimality Theory. *Phonology* 19. 1-32.

Dilly, Laura, Stefanie Shattuck-Hufnagel and Mari Ostendrof (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of phonetics* 24. 423-444.

Downing, Laure. (2004). What African languages tell us about accent typology. *ZAS Papers in Linguistics* 37:101-36.

Duanmu, San. (1990). *A formal study of syllable, tone, stress and domain in Chinese languages*. Ph.D. dissertation, MIT.

Duanmu, San. (1995). Metrical and tonal phonology of compounds in two Chinese dialects. *Language* 71. 225-259.

Duanmu, San. (1997). Recursive constraint evaluation in OT: evidence from cyclic compounds in Shanghai. *Natural language & Linguistic Theory* 15. 465-507.

Duanmu, San. (1999). Metrical structure and tone: evidence from Mandarin and Shanghai. *Journal of East Asian Linguistics* 8. 1-38.

Duanmu, San. (2000). *The phonology of Standard Chinese*. 1st Edition. Oxford: Oxford University Press.

Duanmu, San. (2007). *The phonology of Standard Chinese*. 2nd Edition. Oxford: Oxford University Press.

Duanmu, San and Bing-fu Lu (1990). Word length variation in Chinese two-word construction. Ms., MIT and University of Connecticut.

Fougerson, Cecile and Patricia Keating (1996). Articulatory strengthening in prosodic domain-initial position. *UCLA Working Papers in Phonetics* 92. 61-87.

Fougerson, Cecile. (1999). Articulatory properties of initial segments in several prosodic constituents in French. *UCLA Working Papers in Phonetics* 97. 74-100.

Fox, Anthony. (2000). *Prosodic features and prosodic structure*. Oxford: Oxford University Press.

Everett, Karen. (1998). Acoustic correlates of stress in Pirahã. *The journal of Amazonian languages* 1. 104-162.

Fry, D.B. (1955). Duration and intensity and physical correlates of linguistic stress. *Journal of the acoustical society of America* 27.4. 765-768.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and speech* 1. 126-152.

Gao, Mingkai and Anshi Shi (1963). *Yuyanxe gailun* [Introduction to linguistics]. Beijing: Zhonghua Publishing House

Goldsmith, John. (1976). *Autosegmental phonology*. Ph.D. dissertation, MIT.

Goldsmith, John. (1984). Tone and accent in Tonga. In G. N. Clements and John A. Goldsmith (ed.) *Autosegmental studies in Bantu tone*. Dordrecht: Foris. 19-51.

Gordon, Matthew. (2004). A phonological and phonetic study of word-level stress in Chickasaw. *International journal of American linguistics* 70. 1-32.

Gordon, Matthew. (2011). Stress systems. In John Goldsmith, Jason Riggle and Alan Yu (ed.) *The handbook of phonological theory*. Oxford: Wiley-Blackwell. 141-163.

Gordon, Matthew. (2014). Disentangling stress and pitch accent: a typology of prominence at different prosodic levels. In Harry van der Hulst (ed.) *Word stress: theoretical and typological issues*. Cambridge: Cambridge University Press. 83-118.

Gussenhoven, Carlos. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Halle, Morris and Jean-Roger Vergnaud (1987). *An essay on stress*. Cambridge: MIT press.

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

Hagberg, L. (2006). The place of pitch-accent in a typology of phonological prominence. Paper presented at BeST conference. University of Leiden, Ms. SIL, Mexico.

Hoa, Monique. (1983). *L'accentuation en Pekinois*. Paris: Editions Language Croises.

Hou, Xingquan. (2008). Guangdong Fengkai fangyan shengdiao shiyan yanjiu[An experimental study of tones in Fengkai Cantonese]. Paper presented in 8th National Conference of Phonetics. China, Beijing.

Hou, Xingquan. (2011a). Goulou pian yueyu de liangzi liandu biandiao [Tone sandhi of disyllabic words in Cantonese spoken in Goulou region]. *Fangyan* 2. 132-141.

Hou, Xingquan. (2011b). Guangdong Kaijianhua qingzhong yinjiebu [Iambic foot in Kaijian Cantonese]. *Journal of Jinan University* (Philosophy and social sciences) 4. 117-124.

Hsu, Chai-Shune and Patricia Keating (1999). Domain-initial strengthening: A follow-up study. *UCLA Working Papers on Phonetics* 97. 152-156.

Hsu, Chai-Shune and Sun-Ah Jun (1996). Is tone sandhi domain part of the prosodic hierarchy in Taiwanese? Paper presented at the 132nd meeting of the ASA and the 3rd joint meeting of the ASA and the ASJ. Honolulu, Hawaii.

Hsu, Chai-Shune and Sun-Ah Jun (1998). Prosodic strengthening in Taiwanese: Syntagmatic or paradigmatic? *UCLA Working Papers in Phonetics* 96. 69-89.

Hu, Yushu. (1981). *Xiandai hanyu* [Contemporary Chinese]. Shanghai: Shanghai Education Press.

Hu, Yushu and Bingfu Lu (1988). Guanyu zhiyue hanyu yuxu de yixie yinsu [Constraints on the word-order of Chinese]. *Journal of Yantai University* (Social sciences) 1. 55-60

Hudson, R. A. (1973). Syllables, moras and accents in Beja. *Journal of Linguistics* 9. 53-63.

Hyman, Larry and Russell Schuh (1974). Universals of tone rules: evidence from West Arica. *Linguistics Inquiry* 5. 81-115.

Hyman, Larry M. (1977). Tone and/or accent. In D.J. Napoli (ed.) *Elements of tone, stress and intonation.* Washington, D.C.: Georgetown University Press. 1-20.

Hyman, Larry M. (2006). Word prosodic typology. *Phonology* 23. 225-257.

Hyman, Larry M. (2009). How (not) to do phonological typology: the case of pitch-accent. *Language Sciences* 31. 213-238.

Hyman, Larry. (2011). Tone: is it different? In J. Goldsmith, J. Riggle and A. C. L. Yu (ed.) *The handbook of phonological theory.* Oxford, U.K: Wiley-Blackwell. 197-239.

Jassem, Wiktor, John Morton and Maria Steffen-Batog (1968). The perception of stress in synthetic speech-like stimuli by Polish listeners. *Speech Analysis and Synthesis* I. 289-308.

Jun, Sun-Ah. (1993). *The phonetics and phonology of Korean prosody*. Ph. D. dissertation, Ohio State University.

Jun, Sun-Ah. (1995). Asymmetrical prosodic effects on laryngeal gesture in Korean. In B. Connell and A. Arvaniti (ed.) *Phonology and phonetics evidence: papers in laboratory phonology.* Cambridge: Cambridge University Press. 235-253.

Kager, Rene. (1995). The metrical theory of word stress. In John Goldsmith (ed.) *The handbook of phonological theory*. Oxford, U.K.: Wiley-Blackwell. 367-402.

Kennedy, George A. (1953). Two tone patterns in Tangsic. *Language* 29. 367-373.

Kisseberth, Charles. (1991). *Metrical structure in Chizigula tonology*. MS, University of Illinois.

Kisseberth, Charles and Mohammed I. Abasheikh (1974). Vowel length in Chi Mwi:ni: a case study of the role of grammar in phonology. In A. Bruck, A. Fox and M. W. La Galy (ed.) *CLS papers from the parasession on natural phonology*. 193-209.

Kunter, Gero. (2011). *Compound stress in English: the phonetics and phonology of prosodic prominence*. Berlin: Walter de Gruyter.

Ladd, D. Robert. (1996). *Intonational Phonology*. Cambridge, U.K.: Cambridge University Press.

Ladd, D. Robert. (2008). *Intonational Phonology*. Cambridge, U.K.: Cambridge University Press.

Language Atlas of China. (1987). Complied jointly by Australian Academy of the Humanities and Chinese Academy of Social Sciences. Hong Kong: Longman.

Lehiste, Ilse. (1970). *Suprasegmentals*. Cambridge, Mass.: MIT press.

Levi, Susannah. (2005). Acoustic correlates of lexical accent in Turkish. *Journal of the International Phonetic Association* 35.1. 73-97.

Li, Charles and Sandra A. Thompson (1981). *Mandarin Chinese: a functional reference grammar*. Berkeley, Calif.: University of California Press.

Li, Linding. (1990). Dongci fenlei yanjiu suolue [A brief discussion on verb classification]. *Zhongguo yuwen* 4. 248-257.

Li, Weimin. (1981). Shilun qingsheng yu zhongyin [A preliminary discussion on neutral tone and stress]. *Zhongguo Yuwen* 1. 35-40.

Li, Zhiqiang. (2003). *The phonetics and phonology of tone mapping in a constraint-based approach*. Ph.D. dissertation, MIT.

Li, Ya, Yingchao Lu, Xiaoying Xu and Jianhua Tao (2011). Lianxu yuliu zhong yunlu cengji he diaoxing zhuhe dui zhongyin ganzhi de yingxiang [Influence of prosodic hierarchy and tone combination patterns on Mandarin stress perception in continuous speech]. *Journal of Tsinghua University* (Sci and Tech) 9. 1239-1248.

Liao, Rongrong. (1994). Suzhouhua danzi he shuangzizu de yinchang guilu [Duration patterns of syllables in monosyllabic and disyllabic words in Suzhou dialect]. In Feng shi and Rongrong Liao (ed.) *Yuyin Conggao*. Beijing: Beijing Language College Press. 150-160.

Lin, Maocan and Jingzhu Yan (1980). Beijnghua qingsheng de shengxue xingzhi [Phonetic qualities of neutral tone in Beijng Mandarin]. *Fangyan* 3. 166-178.

Lin, Maocan, Jingzhu Yan and Guohua Sun (1984). Beijinghua liangzizu zhengchang zhongyin chubu shiyan [A preliminary experiment of stress in regular disyllabic words in Beijing Mandarin]. *Fangyan* 1. 57-73.

Lin, Maocan and Jingzhu Yan (1990). Putonghua qingsheng yu qingzhongyin [Neutral tone and stress in Standard Mandarin]. *Yuyan Jiaoxue Yu Yanjiu* 4. 88-104.

Lin, Tao. (1983). Tantao Beijinghua qingsheng xingzhi de shubu shiyan [A preliminary experiment on the nature of neutral tone in Beijing Mandarin]. *Yuyanxe Luncong* 10. 16-37.

Liu, Feng-His. (1992). Verb and syllable in Chinese, paper presented at the 25th international conference on Sino-Tibetan languages and linguistics. Berkeley, California.

Liu, Fu. (1924). *Sisheng siyan lu* [Experimental study of four tones]. Shanghai: Yiqun Publishing House.

Lu, Bingfu. (1989). Hanyu dingyu de feilei ji yuxu [Classes and order of nominal modifiers in Chinese]. *Huawen shijie 4. 44-52.*

Lu, Bingfu. (1998). *Left-right asymmetries of word order variation: a functional explanation.* Ph.D. dissertation, University of Southern California.

Luo, Changpei. (1930). Xiamen yinxi [The phonology of Xiamen]. *Bulletin of the Institute of History and Philology, Academia Sinica.* (Reissued by the Chinese Academy of Science, Shanghai, 1956).

Luo, Changpei and Jun Wang (1957). *Putong yuyanxe gangyao* [An introduction to general linguistics]. Beijing: Science Publishing House.

Lü, Shuxiang. (1963). Xiandai hanyu danshuang yinjie wenti chutan [An investigation of the issue of monosyllabicity and disyllabicity in Contemporary Chinese]. *Zhongguo Yuwen*. 10-22.

Lü, Shuxiang. (1980). Danyang fangyan de shengdiao xitong [The tone system of Danyang]. *Fangyan* 2. 85-122.

Lü, Shuxiang. (2001). *Xiandai hanyu babai ci* [Eight hundred words in Contemporary Chinese]. Beijing: Commercial Press.

Lynch, J. (1983). On the Kuman 'liquids'. *Languages and Linguistics in Melanesia* 14. 98-112.

McCarthy, John and Alan Prince (1986). *Prosodic morphology*. Ms, University of Massachusetts, Amherst, and Brandeis University.

McCarthy, John and Alan Prince (1993). Generalized alignment. *Yearbook of Morphology 1993.* 79-153.

McCarthy, John and Alan Prince (1995). Faithfulness and reduplicative identity. In J. Beckman, L. Walsh Dickey, and S. Urbanczyk (ed.) *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*. 250-384.

Odden, David. (1985). An accentual approach to tone in Kimattmbi. In D. Goyvaerts (ed.) *African linguistics: studies in memory of M.W.K. Seminenke*. Amsterdam: John Benjamins. 345-419.

Odden, David. (1995). Tone: African languages. In John A. Goldsmith (ed.) *The handbook of phonological theory*. Oxford, U.K.: Wiley-Blackwell. 444-475.

Ortega-Llebaria, Marta and Pilar Prieto (2011). Acoustic correlates of stress in Central Catalan and Castilian Spanish. *Language and Speech* 54 (1). 73-97.

Ortega-Llebaria, Marta, Hong Gu and Jieyu Fan (2013). English speaker's perception of Spanish lexical stress: context-driven. *Journal of Phonetics* 41. 186-197.

Pankratz, Leo and Eunice V. Pike (1967). Phonology and morphotonemics of Ayutla Mixtec. *IJAL* 33. 287-299.

Pearce, Mary. (2006). The interaction between metrical structure and tone in Kera. *Phonology* 23. 259-286.

Pierehumbert, Janet and David Talkin (1992). Lenition of /h/ and glottal stop. In G. Docherty and D. R. Ladd (ed.) *Papers in Laboratory Phonology II: Gesture, Segment, Prosody* Cambridge, UK: Cambridge University Press. 90-117.

Potisuk, Siripong, Jackson Gandour and Mary P. Harper (1996). Acoustic correlates of stress in Thai. *Phonetica* 53. 200-220.

Prince, Alan. (1980). A metrical theory of Estonian Quantity. *Linguistic Inquiry* 11. 511-562.

Prince, Alan. (1990). Quantitative consequences of rhythmic organization. In Michael Ziolkowski, Manuela Noske, and Karen Deaton (ed.) *Papers from the 26th regional meeting of the Chicago Linguistic Society* 12. 355-398.

Prince, Alan and Paul Smolensky (1993). Optimality theory: constraints interaction in generative grammar. Ms, Rutgers University and University of Colorado.

Qian, Nairong and Shi Rujie (1983). Suzhou fangyan liandu biandiao tanlun zhi 2[A discussion on the tone sandhi of Suzhou 2]. *Fangyan* 4.275-296.

Qian, Nairong. (1990). *Xiandai hanyu* [Contemporary Chinese]. Hubei: Hubei Education Press.

Qian, Nairong. (1992). Dangdai wuyu yanjiu [Contemporary Studies on Wu dialects]. Shanghai: Jiaotong Publishing House.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, Jan Svartvik (1972). *A grammar of Contemporary English*. New York: Seminar Press.

Rose, Phil. (1990). Acoustics and phonology of complex tone sandhi. *Phonetica* 47. 1-35.

Selkirk, Elisabeth. (1980). Prosodic domains in phonology: Sanskrit revisited. In Mark Aronoff and M. L. Kean (ed.) *Juncture*. Saratoga, California: Anma libri.

Selkirk, Elisabeth O. and Tong Shen (1990). Prosodic domains in Shanghai Chinese. In Sharon Inkelas and Draga Zec (ed.) *The phonology-syntax connection*. Chicago: University of Chicago Press. 313-338.

Shao, Jingmin. (1995). Shuangyinjie V + N jiegou de peijia fenxi [An analysis on the valence of disyllabic V + N structure]. In Shen Yang and Zheng Ding'ou (ed.) *A study on valence grammar of contemporary Chinese.* Beijing: Beijing University Press.

Shen, Xiaonan. (1990). Tonal co-articulation in Mandarin. *Journal of Phonetics* 18. 281-295.

Shen, Z.W., C. Wooters, and W. S-Y. Wang (1987). Closure duration in the classification of stops: a statistical analysis. *The Ohio State University Working Papers in Linguistics* 35. 197-209.

Shih, Chilin. (1986). *The prosodic domain of tone sandhi in Chinese*. Ph.D. dissertation, University of California, San Diego.

Sluijter, Agaath M. C. (1995). *Phonetic correlates of stress and accent.* Ph.D. dissertation, Leiden University.

Sluijter, Agaath M. C and Vincent J. van Heuven (1996). Spectral balance as an acoustic correlate of linguistics stress. *Journal of the Acoustica Society of America* 100. 2471-2485.

Smith, Jennifer. (1998). Noun faithfulness and word stress in Tuyuca. In Jennifer Austin and Aaron Lawson (ed.) *Proceedings of ESCOL* 97. 180-191.

Smith, Jennifer. (2001). Lexical category and phonological contrast. In Robert Kirchner, Joe Pater, and Wolf Wikely (ed.) *Papers in experimental and theoretical linguistics 6: workshop on the lexicon on phonetics and phonology*. Edmonton, Alta.: University of Alberta. 61-72.

Smith, Jennifer. (2002). *Phonological augmentation in prominent positions*. Ph.D. dissertation, University of Massachusetts, Amherst.

Steriade, Donca. (1993). Positional neutralization. Handout presented at NELS 23. University of Massachusetts, Amherst.

Steriade, Donca. (1995). Positional neutralization. Ms, UCLA.

Trefty, D. (1969). A comparative study of Kuman and Pawaian. *Pacific Linguistics* B-13. Canberra, Australia: The Australian National University.

Trubetzkoy, Nikolai Sergeevich. (1939). Principles of phonology. Translated by Christiane A. M. Baltaxe. Berkeley, Calif.: University of California Press. 1969.

Van der Hulst, Harry. (2011). Pitch accent systems. In Van Oostendorp, Marc, Colin Ewen, Elizabeth Hume and Keren Rice (ed.) *The Blackwell companion to phonology, vol. 2*. Malden, MA & Oxford, U.K.: Wiley-Blackwell. 1003-1026.

Van der Hulst, Harry. (2014). The study of word accent and stress: past, present, and future. In Harry van de Hulst (ed.) *Word stress: theoretical and typological issues*. Cambridge, U.K.: Cambridge University Press. 3-55.

Wang, Guangquan. (1993). Dongci zhijie zuo dingyu shi de weizhi [The order of bare form verbs functioning as attributives]. *Zhongguo yuwen* 1. 70-72.

Wang, Hongjun. (2001). Yinjie danshuang, yinyu yanzhan yu yufa jiegou leixing he chenfen cixu [The relation between disyllabification, pitch range and the grammatical structure in Chinese]. *Dangdai yuyanxue* 4. 241-252.

Wang, Jialing. (2002). Youxuanlun he Tianjinhua de liandubiandiao ji qingsheng [OT, tone sandhi and neutral tone in Tianjin dialect]. *Zhongguo Yuwen* 4. 363-372.

Wang, Jing and Wang Lijia (1993). Putonghua duoyinjie ci yinjie shichang fenbu moshi [The syllable duration patterns of polysyllabic words in Standard Mandarin]. *Zhongguo yuwen* 2. 112-116.

Wang, Li. (1931). *Une prononciation chinoise de Po-pei, etudiee a l'aide de la phonetique experimentale*. Imprimerie des Presses Universitaires de France.

Wang, Maochun. (2003). Xiandai hanyu VP + zhe chengli de jige tiaojian [Conditions on the well-formedness of VP + zhe in Contemporary Chinese]. *Journal of Sichuan Normal University* 6. 68-73.

Wang, Yunjia. (1993). *Beijinghua shengdiao weiguan bianhua shiyan yanjiu* [A study on the micro change of tones in Beijing Mandarin]. Ph.D. dissertation, Beijing University.

Wang, Yunjia, Min Chu, lin He and Yongqiang Feng (2003). Lianxu huayuzhong shuangyinjie yunluci de zhongyin ganzhi [Stress perception of disyllabic prosodic word in connected speech]. *Acta Acustica* 6. 534-539.

Weise, Richard. (1996). Phrasal compounds and the theory of word syntax. *Linguistic Inquiry* 27. 183-193.

Wright, Martha S. (1983). *A metrical approach to tone sandhi in Chinese dialects*. Ph.D. dissertation, University of Massachusetts, Amherst.

Xie, Zili. (1982). Suzhou fangyan liangzizu de biandiao [Disyllabic tone sandhi in Suzhou]. *Fangyan* 4. 245-263.

Xu, Baohua, ZhenZhu Tang, and Nairong Qian (1981). Xinpai Shanghai fangyan de liandu biandiao [Tone sandhi in New Shanghai]. *Fangyan* 2. 145-155.

Xu, Baohua, ZhenZhu Tang, and Nairong Qian (1982). Xinpai Shanghai fangyan de liandu biandiao [Tone sandhi in New Shanghai]. *Fangyan* 2. 115-128.

Xu, Baohua, ZhenZhu Tang, and Nairong Qian (1983). Xinpai Shanghai fangyan de liandu biandiao [Tone sandhi in New Shanghai]. *Fangyan* 3. 197-201.

Xu, Shirong. (1982). Shuangyijie ci de yinliang fenxi [A quantitative analysis loudness of disyllabic words]. *Yuyan jiaoxue yu yanjiu* 2. 4-19

Xu, Yi and Yiya Chen (2006). Production to weak elements in speech-evidence from F0 patterns of neutral tone in Standard Mandarin. *Phonetica* 63. 47-75

Yin, Zuoyan. (1982). Guanyu Putonghua shuangyin changyongci qingzhongyin de chubu kaocha [A preliminary study on stress in commonly-used disyllabic expressions in Putonghua]. *Zhongguo Yuwen* 5. 168-173.

Yip, Moria. (1980). *The tonal phonology of Chinese*. Ph.D. dissertation, MIT.

Yip, Moria. (1999). Feet, tone reduction and speech rate at the word and phrase level in Chinese. In Rene Kager and Wim Zonneveld (ed.) *Phrasal phonology*. Nijmegen: Nijmegen University Press. 171-194.

Zee, Eric and Ian Maddieson (1979). Tones and tone sandhi in Shanghai: Phonetic evidence and phonological analysis. *UCLA Working Papers in Phonetics* 45. 93-129.

Zhang, Guoxian. (1989). V + N jiegou zhong danshuang yinjie dongzuo dongci gongneng chayi chutan [A preliminary discussion on the different grammatical functions of monosyllabic/disyllabic verbs in the V + N structure]. *Zhongguo yuwen* 3. 186-190.

Zhang, Jie. (2001). *The effect of duration and sonority on contour tone distribution: Typological survey and formal analysis*. Ph.D. dissertation, UCLA.

Zhang, Min. (1998). *Renzhi yuyanxue yu hanyu mingci duanyu* [Cognitive linguistics and Chinese nominal phrases]. Beijing: Chinese Social Science Press.

Zhang, Xunru. (1947). Guoyu qingzhongyi bijiao [A comparison of the stressed syllable and unstressed syllable in Standard Mandarin]. *Guowen yuekan* 57. 11-15.

Zhang, Zhigong. (1983). *Xiandai hanyu* [Contemporary Chinese]. Beijing: People's Education Press.

Zhu, X. S. (1996). Two stress patterns of Shanghai compounds. In P. McCormack and A. Russell (ed.) *Proceedings of the sixth Australian international conference in speech science and technology*. Canberra, Australia: Australian Speech Science and Technology Association. 581-586.

Zhu, X. S. (2005). *An experimental study on Shanghai tones*. Shanghai: Shanghai Education Press.

Zoll, Cheryl. (1997). A note on multiple prominence and tone mapping. In Benjamin Bruening, Yoonjung Kang and Martha McCinnis (ed.) *PF: papers at the interface*. *MIT Working Papers in Linguistics* 30. 97-111.