LINGUISTICS IN CHINA

No. 3

Edited by Feng Shi and Hongming Zhang

语言学文选

(第三辑)

主编 ◎ 石 锋 张洪明

世界图书出版公司
World Publishing Corporation
北京·广州·上海·西安
Beijing · Guangzhou · Shanghai · Xi'an

Linguistics in China No. 3
Editing
by Feng Shi and Hongming Zhang

Published in China
by World Publishing Corporation Beijing Branch
137 Chaonei Street, Dongcheng District, Beijing, China

First published 2020

Chinese library Cataloguing in Publication Data
Data available
No. (2019) 209611

Printed by Beijing Jianhong Printing Company Limited

ISBN 978-7-5192-6795-7

Linguistics in China

CONTENTS

目　　录

Some Issues on Prosodic Phonology and Chinese Prosodic Studies

Hongming Zhang
Translated by Yuxia Yin
Proofread by Xiaofei Lu *

Abstract: This article introduces and clarifies the basic concepts in prosodic phonology, discusses the principles and methods for determining prosodic units, and examines the application of these methods to Chinese. The article centers on the differentiation of concepts of prosodic phonology, the definition of prosodic structure, current major issues such as what a prosodic unit is and how a specific unit is defined, the mismatch between syntax and phonology and the build-up of the prosodic hierarchy in Chinese.

Key words: prosodic hierarchy, mora, syllable, foot, prosodic word, utterance, Strict Layer Hypothesis

1 Introduction

Over the last two decades, prosodic phonology has received increasing attention among Mandarin Chinese phonologists (Zhang 1992, 1997, 2008a, 2008b, 2014; Chen and Zhang 1997; Feng 1997, 2009, 2013; Duanmu 2000; Cao 2001; Ye 2001; Wu 2006; Wang 2008; Li 2008; Cao 2010; Deng 2010; Dong 2011; Zhou 2001). This line of research has contributed substantially to the development of general prosodic phonological theories, but at the same time created some new problems. This paper firstly introduces and clarifies some basic concepts in prosodic phonology, then discusses the principles for defining prosodic units, and finally examines the application of these principles in Chinese and points out common misunderstandings in current research in Mandarin Chinese prosodic phonology.

Prosodic phonology studies the units of the prosodic structure in human languages and the relations among these units. Natural utterances can be parsed into a finite set of hierarchical prosodic units. From the lowest level to the highest, they are mora, syllable, foot, prosodic word/

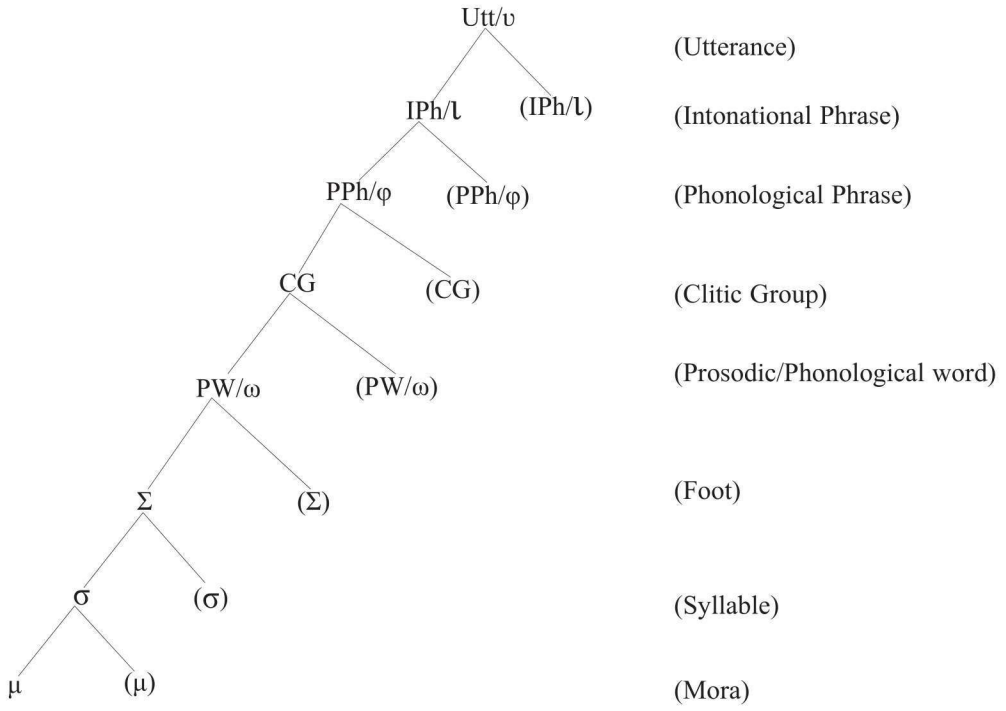
* Hongming Zhang, University of Wisconsin-Madison, Wisconsin, 53706, USA; School of Literature, Nankai University, Tianjin, 300071, China.

Yuxia Yin, Tianjin Normal University, Tianjin, 300387, China.

Xiaofei Lu, The Pennsylvania State University, University Park, 16802, USA.

phonological word, clitic group, phonological phrase, intonational phrase, and utterance. The hierarchical structure can be represented in (1).

(1) Prosodic Hierarchy



This figure captures the current understanding of prosodic structure of human languages, with all known prosodic units included. Generally speaking, (1) includes the hierarchical prosodic units found in all languages, representing language universality. However, a given language does not necessarily contain all the units in (1), demonstrating language specificity. Linguistic typology determines whether particular prosodic units exist in some languages or not, and whether those that do exist are salient or not. Also, the prosodic hierarchy satisfies the Strict Layer Hypothesis proposed by Selkirk (1984), as shown in (2).

(2) Strict Layer Hypothesis

a. A given nonterminal unit of the prosodic hierarchy, x^p is composed of one or more units of the immediately lower category, x^{p-1}.
b. A unit at a given level of the hierarchy is exhaustively contained in the superordinate unit of which it is a part.

(2a) stipulates that each prosodic unit, with the exception of the terminal moras, must directly dominate the unit under it in the tree diagram presented above. (2b) requires that each prosodic

unit must be parsed into its immediate lower level constituents.

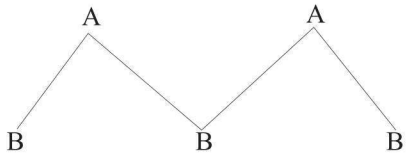
Early studies in phonological prosody were conducted within rule-based theoretical frameworks. And the mapping process between syntactic structure and prosodic structure was mainly based on X-bar theory. Later research, however, has been profoundly influenced by Optimality Theory and the mapping process is analyzed on the basis of various constraints (Selkirk 1995, 2011; Truckenbrot 1995). Within the framework of Optimality Theory, the Strict Layer Hypothesis is decomposed into four constraints as shown in (3).

(3) Strict Layer Hypothesis in Optimality Theory

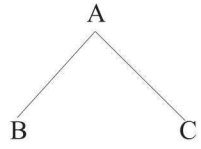
a. Layeredness: no C^i dominates a C^j, iff $j > i$ (e.g. no σ dominates a Σ)
b. Headedness: any C^i must dominate a C^{i-1} (e.g. a ω must dominate a Σ)
c. Exhaustivity: no C^i dominates C^j, iff $j<i-1$, (e.g. no ω immediately dominates a σ)
d. Nonrecursivity: no C^i dominates C^j, iff $j = i$ (e.g. no Σ dominates a Σ)

The Strict Layer Hypothesis is the well-formedness condition for the tree diagram of prosodic hierarchy. The following prosodic structures in (4) are considered ill-formed because of their violating the Strict Layer Hypothesis.

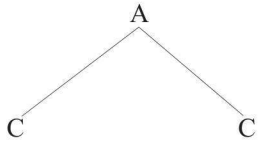
(4) a. Multiple domination



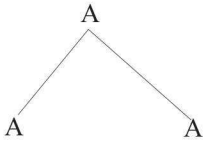
b. Heterogeneous Sisters



c. Skipping of levels



d. Recursion



The Strict Layer Hypothesis determines the organization of the prosodic hierarchical structure and constrains the prosodic constituents that serve as domains of phonological rule application. The structural relationship among different prosodic constituents directly bears up on the principles governing the construction of the prosodic hierarchy. Therefore, the Strict Layer Hypothesis is the core principle in prosodic phonology. The Chinese linguistics literature, however, does not appear to be explicitly aware of principle, as illustrated by the quotations in (5).

(5) a. "Since a foot is a prosodic word, 'X' and 'Y' form a canonical prosodic word." (Feng 2009: 10)

b. "In some languages, two moras can also form a foot (with mora being related to foot directly). That is to say, the falling and rising of rhythm can be realized within a syllable. For instance, "I" in English is pronounced as [ai] , in which both the [a] and [i] have a certain duration and the dynamic process starting from target [a] towards target [i] is rather clear. Therefore, the falling and raising rhythm can be realized in these two moras and they can therefore form a foot." (Feng 2009: 18)

c. "A prosodic word is equivalent to a foot and serves as the minimum domain within which phonological rules operate. … Therefore, a prosodic word constitutes a foot." (Cao 2001: 177)

Of these views, (5a) is inconsistent with the layeredness condition, namely (3a); (5b) is inconsistent with the exhaustivity condition, namely (3b) and also misunderstands the concept of foot (hereafter to be analyzed in section 2.3); (5c) also violates the layeredness condition. Despite current debate over the total number of prosodic hierarchies, the existence of specific prosodic hierarchies, and the principles for constructing the tree diagram, as the core of the Strict Layer Hypothesis, layeredness and headedness are widely accepted as two universal principles, so far with no violations found in known languages.

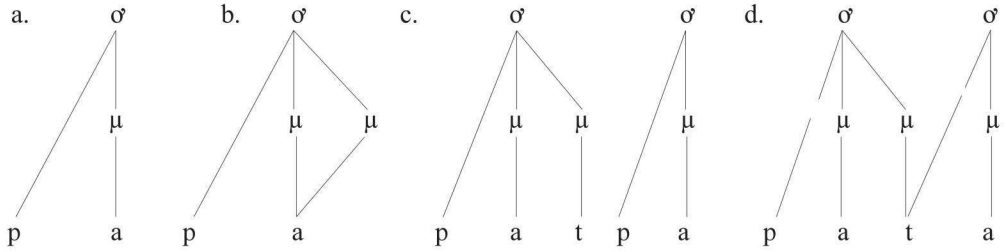
2 Prosodic units and Mandarin prosodic hierarchy

One major aim of prosodic phonology is to investigate prosodic units of human languages and construct the prosodic hierarchy. This raises the following questions: In what ways are these prosodic units defined and constructed? How should we define such units as mora, syllable, foot, prosodic word/phonological word, clitic group, phonological phrase, intonational phrase, and utterance? The definitions and constituents of the aforementioned will be discussed in the present section.

2.1 Mora (μ)

As the minimal unit in the tree diagram of prosodic hierarchy in human languages, mora acts mainly to determine the weight of syllables in some quantity-sensitive languages. It is defined as the elements that make up the rime of each syllable and are used to measure the weight of each syllable. A heavy syllable consists of two moras while a light syllable consists of only one mora. A long vowel or two segments that carry two moras can, therefore, form a heavy syllable, and a short vowel carries only one mora, and thus forms a light syllable. A consonant in the onset position of a syllable does not count as a mora. This is shown in (6) as follows.

(6)



Here, (6a) is a short vowel and consists of only one mora, thus being a light syllable. (6b) is a long vowel and consists of two moras, thus being a heavy syllable. (6c) is composed of a heavy syllable and a light syllable. Despite the fact that the first syllable has only a short vowel, it ends with a consonant and, therefore, consists of two moras. (6d) also has two syllables with the first one *pat* being a heavy syllable and the second one *ta* light. It can be seen here that the concept of heavy or light syllables should be differentiated from that of long or short syllables. A short syllables can be either heavy or light.

Moras play an important role in some quantity-sensitive languages. For instance, the metrical requirement of śloka in Sanskrit poetry can be achieved by alternation of heavy and light syllables. This can be exemplified in (7) and (8) as follows.

(7)

line 1	x	x	x	x
line 2	˘	(-)	(-)	(˘)
line 3	x	x	x	x
line 4	˘	-	˘	x
line 5	x	x	x	x
line 6	˘	(-)	(-)	(˘)
line 7	x	x	x	x
line 8	˘	-	˘	x

(8)

Odd *pāda*	X	X	X	X	˘	(-)	(-)	(˘)
Even *pāda*	X	X	X	X	˘	-	˘	X

However, Mair and Mei have misinterpreted the prosody of śloka in Sanskrit poetry as the

contrast between short and long syllables (Mair and Mei 1991: 375–470), as shown in (9).

(9) "*śloka* consists of four *pādā*, or quarter verses, of eight syllables each, or two lines of sixteen syllables. Each line allows great liberty except for the 5th, 13th, 14th, and 15th syllable, as in the above schema, where the crosses denote either long or short, the bars long, and the breve signs short."

Due to this misunderstanding, they were unable to explain why the short syllables *vat* and *mos* in (10a) and *ṣot* and *har* in (10b) can be metricalized. It is apparent, however, that the metrical pattern of śloka in Sanskrit poetry is constructed by resorting to the contrast between heavy and light syllables, rather than the contrast between short and long syllables (Mishra 1999: 21–22).

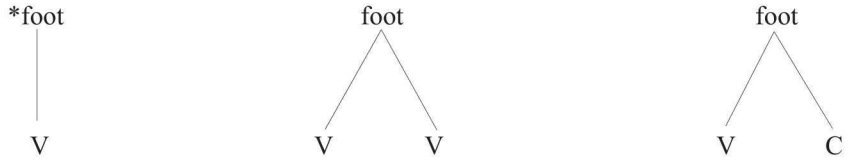
(10) a. nɑmɑste puruṣādhyɑkṣɑ nɑmɑste bhɑktɑ **vɑt** sɑlɑ |
nɑmɑste'stu hṛṣīkeśɑ nārāyɑṇɑ nɑ **mos** tute ||
(Quoted from *Adhyātmɑ Rāmāyɑṇɑ*, 1.5.59.)

b. lokānāṃ tvɑṃ pɑro dhɑrmɑḥ puru **sot** tɑmɑḥ |
śāraṇyɑṃ śɑrɑṇɑṃ cɑ tvām āhurdivyāḥ mɑ **hɑr** ṣɑyɑḥ ||
(Quoted from *Vālmīki Rāmāyɑṇɑ*, V, CXIX, 14.)

Likewise, in the field of Chinese linguistics, there are similar assertions that confuse the contrast between heavy and light syllables with that between short and long syllables. For example, some scholars (Feng 2013: 46) claim that in some longer syllables, the nucleus can be bi-syllabic (VV) while the coda can be a consonant cluster (CC), as shown in (11).

(11) a. Short syllable: V
Long syllable: V(V)C
Super long syllable: VVCC(C)

b. Short syllable　　Long syllable　　Long syllable



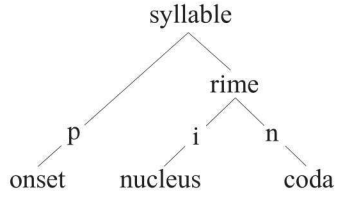
The classification in (11) is incorrect: here, the third type VC is neither a long syllable nor a foot. It is a short syllable, but a heavy one.

Mandarin Chinese is a quantity-insensitive language. Consequently, mora as a unit in the prosodic hierarchy falls into the non-salient category in Mandarin.

2.2 Syllable (σ)

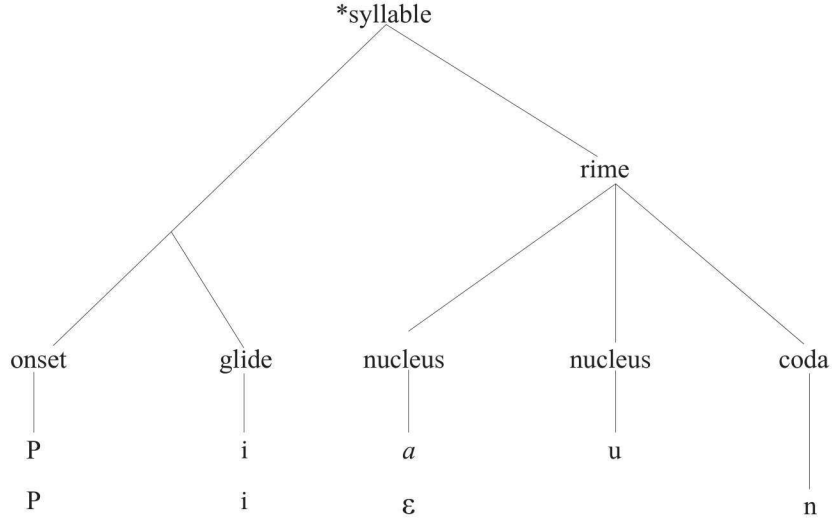
Syllable is the minimal structural unit that can be identified naturally in speech and that can form a foot within words in some stress languages. It has its internal structure and can be decomposed into onset and rime, of which the latter is further divided into nucleus and coda, as illustrated in (12).

(12)



Syllable constitutes the domain in which many phonological rules operate, such as *r*-strengthening, aspiration, and velarization in Spanish. Other syllable-sensitive phonological rules include glottalization in American English, the aspiration of initial voiceless stops in English, final devoicing of obstruents in German, etc. In some languages, syllables can be categorized into strong syllables (i.e. stressed syllables) and weak syllables (i.e. unstressed syllables). A strong syllable and a weak one adjacent to it form a metrical foot. Syllable is the minimal hierarchical prosodic unit in Mandarin Chinese, and its construction must conform to the phonotactic constraints of Mandarin Chinese. The erroneous representation of syllables *biao* /piau/ and *bian* /piɛn/ in Mandarin below in (13) clearly indicates an incorrect understanding of how syllables are defined. The construction of syllables must comply with the sonority sequence principle and, as a result, one syllable can only have one nucleus.

(13)



2.3 Foot (Σ)

Foot is a prosodic unit that is higher than syllable and lower than prosodic word in the prosodic hierarchy. It works as the domain within which some segmental rules apply. In prosodic phonology, the definition of foot is closely related to metrical binary contrast. For instance, in stress languages a foot often consists of a strong syllable and a weak syllable with the strong syllable carrying the primary stress. Many phonological rules and phonotactic constraints can only be represented explicitly by means of foot, such as English aspiration, *l*-devoicing, diphthong shortening, obligatory *n*-velarization, etc. Foot has been classified into two different categories in the literature, namely within-word foot and cross-word-boundary foot. The foot discussed here is within-word foot, which is also referred to as metric-prominent foot. The universal rules can guarantee that only one syllable in a foot is prominent (for example, it is strong, heavy, long or high pitched). As to which syllable is prominent, it falls into the sphere of certain specific linguistic rules. In particular languages, there are strict constraints for syllables on foot formation. There is another type of foot called degenerate foot which is a marked form consisting of only one light syllable in quantity–sensitive systems. Nevertheless, strict constraints have been proposed on degenerate foot in many some languages, which is a feature of universal grammar. In many languages, degenerate foot is allowed due to its appearing in the strong metrical position. Whether foot as a prosodic unit exists in a language or not is determined by the existence of metrical binary contrast such as heavy vs light, long vs short, high pitch vs low pitch, strong and weak. In Mandarin Chinese, however, there is no such contrast, thus denying the existence of foot defined in phonology. A thorough review the literature to date indicates that the discussion of foot within the Chinese linguistics field is fraught with misunderstandings, as illustrated by the quotations in (14).

(14) a. "The basic foot in Chinese is disyllabic. That is to say, disyllabic foot is the most prevalent form, despite the existence of monosyllabic and trisyllabic foot. …We take disyllabic foot as the minimal and most fundamental foot form, thus naming it 'standard foot". Other types of foot are considered to be variants of bi-syllabic foot with monosyllabic foot being degenerate foot and trisyllabic foot being super foot. However, constraints exist in terms of the appearance of both degenerate and super foot. The difference among standard, degenerate and super foot lies in the following. Under normal circumstances, disyllabic foot has the priority to be realized in language since it is the most prevalent and fundamental form. After that, if there are still monosyllabic elements left, it will be attached to its adjacent bi-syllabic foot to form a trisyllabic foot. As for degenerate foot, it only occurs in monosyllabic words that function as independent intonational groups, where it meets the requirements of foot by means of pause or vowel lengthening. "（Feng 2009: 2–3）

b. "A foot is a prosodic word. In Chinese a syllable is usually a 'syllabic word, ' thus forming a 'two-syllabic-words-combined' unit. … In some languages, moras can form a foot directly. For example, cat [kæt] functions as an independent intonational group and can be a foot in and of itself. In other languages, however, moras are not allowed to form a foot. … Another important concept in the prosodic hierarchy is prosodic word. … What is a prosodic word? This concept is very important because this is the level at which prosody meets morphology (interface). This can only be understood if you have some knowledge of hierarchy. For instance, the units below foot in the prosodic hierarchy are essentially phonetic elements. Of course, these units are suprasegmental in nature (i.e. they can be measured by weight)…" (Feng 2013: 46)

c. "A prosodic word is also a foot." (Cao 2001: 177)

d. "A prosodic word is equivalent to a foot." (Cao 2001: 177)

e. "The basic foot or standard foot in Mandarin Chinese consists of two standard stresses." (Cao 2001: 177)

The problem in (14a) lies in that the definition of foot is not based on the number of syllables, but on the metrical prominence principle. The statements in (14b–c) are also problematic because they confuse foot with prosodic word. Foot is not prosodic word as the two constitute different units in the prosodic hierarchy. The analysis of *cat* as well as the units lower than foot in (14b) is erroneous because *cat* is just a syllable, not a foot. Moreover, the claim "the units below foot in the prosodic hierarchy are essentially phonetic elements" is also problematic. The unit below foot is syllable, whereas the unit below syllable is mora. The "phonetic elements" that "can be measured by weight" are never suprasegmental, but rather segmental (that is, they are related to syllable weight). (14e) goes against the definition of foot since two syllables within the same foot cannot carry two standard stresses.

The phonological structures in Mandarin Chinese lack metrical binary contrast. Therefore, there is no foot in the prosodic hierarchy of Mandarin in terms of prosodic phonology.

2.4 Prosodic word (ω)

Prosodic word, also called phonological word, is the minimal prosodic unit that is higher than foot in the prosodic hierarchy and that directly dominates foot. When the words defined in morphosyntax cannot correspond to those defined in phonology, the concept of prosodic word becomes extremely essential. Prosodic word constitutes the lowest unit in the prosodic hierarchy constructed on the basis of mapping rules that make use of non-phonological notions. Prosodic word represents the interaction between the phonological and morphological components. As required by the Strict Layer Hypothesis, all the feet in a sound string must be grouped to form prosodic words. Every foot must be included in a prosodic word and the syllables belonging to the

same foot cannot be dominated by different prosodic words. The domain of prosodic word should be the terminal constituents in syntax trees. The domain of prosodic words consists of 1) a stem; 2) any element identified by specific phonological and/or morphological criteria; 3) any element marked with the diacritic [+w] or any attached elements within the terminal constituents of a syntactic tree which forms part of the adjacent prosodic word closest to the stem; and if no such prosodic word exists, then the independent elements form a prosodic word on their own (Nespor & Vogel 1986: 141). Prosodic words across languages have demonstrated a great similarity and their minor differences can be accounted for by parameters. Many phonological rules such as stress assignment and vowel harmony take prosodic words as the application domain.

Although the concept of prosodic word is not new, it has never been truly understood in the Chinese linguistics field. as evidenced by the quotations in (15) are quoted from various published works on foot.

(15) a. "'Prosodic word, ' from the perspective of prosody, can be defined as 'the smallest language unit that can be used freely.' In prosody, 'language unit' means 'prosodic unit' and, therefore, prosodic words are based upon the prosodic units in languages." (Feng 2009: 1).

b. 'The smallest unit that can be used freely is foot." (Feng 2009: 1)

c. "In prosodic morphology, the smallest prosodic unit that can be used is 'foot.' Therefore, a prosodic word must contain at least one foot. And if a foot consists of two syllables, then a prosodic word naturally must contain at least two syllables. … Regardless of the relations between the elements forming foot, so long as they fulfill the basic requirements of foot, prosodic words can still hold water." (Feng 2009: 2)

As stated here, (15a) claims that prosodic word is "the smallest prosodic unit that can be used freely." But then in (15b) and (15c), the same author proposes that foot is "the smallest prosodic unit that can be used freely, " thus confusing prosodic word with foot. From earlier discussions, it may be concluded that prosodic word and foot are two kinds of completely different prosodic hierarchical units. Foot is determined on the basis of phonological notions while prosodic word constitutes the lowest units in the prosodic hierarchy constructed on the basis of mapping rules that make use of those notions which represent the interface between phonological and morphosyntactic features.

The tone sandhi in Mandarin Chinese can provide evidence for the proposal of establishing prosodic word as a prosodic hierarchical unit, but there is no foot as a prosodic unit in Mandarin. So, prosodic word should be the smallest prosodic unit that is higher than syllable in the prosodic hierarchy and that dominates syllable directly.

2.5 Clitic group (CG)

Clitic group is the prosodic unit that is located between prosodic word and prosodic phrase in the prosodic hierarchy and that joins a clitic element together with the lexical hood, which is its host. Clitic group consists of a prosodic word a directional clitic and a plain clitic that has no possible host with which it shares more category memberships. How clitic elements choose their lexical hood, that is, when a clitic chooses as host the word on its left or right is determined by the syntactic properties of a particular language. Let us take the stress rule in Latin as an example to show that clitic group is an independent prosodic hierarchical unit. The distribution of the main stress rule in Latin can be stated as follows: within polysyllabic words, the main stress rule assigns primary stress either to the penultimate or to the antepenultimate syllable. If the penultimate syllable is heavy then the primary stress falls on it, as illustrated in (16a) and (17a); if the penultimate syllable is light, then the primary stress falls on the antepenultimate syllable, as illustrated in (16b) and (17b). Nevertheless, the examples in (18) appear to be exceptions. In (18a), despite being light, the penultimate syllable still receives the stress.

(16) a. strat**é́**gus "chief"
b. p**ó**pŭlum "people"

(17) a. stomach**ő**sus "irritated"
b. hom**ú**cŭlus "little man"

(18) a. ros**ǎque** "and the rose (nom.)"
b. ros**ā́que** "and the rose (abl.)"

Actually, the application of the main stress rule in Latin is determined by prosodic units. The examples in (16) and (17) are all prosodic words while those in (18) are clitic groups. The application domain of stress assignment of prosodic words depends on whether the penultimate syllable is heavy or light. In contrast, in clitic groups, stress is always assigned to the penultimate syllable regardless of whether it is heavy or light. From this, it may thus be concluded that prosodic word and clitic group, being different prosodic units, are different domains for applying different phonological rules.

The tone sandhi in Mandarin Chinese can also provide evidence for the proposal of establishing clitic group as an independent prosodic hierarchical unit. In Mandarin, syntactic structures determine that the clitic element will always choose the word on its left as hood to form a clitic group.

2.6 Phonological phrase (φ)

Phonological phrase is a prosodic hierarchical unit that is higher than prosodic word or clitic group in the prosodic hierarchy and that consists of one or more prosodic words or clitic groups. The prosodic phrase is a phonological unit that is established on the basis of mapping rules that