THE RISE OF DISYLLABLES IN OLD CHINESE:

THE ROLE OF LIANMIAN WORDS

by

JIAN LI

A dissertation submitted to the Graduate Faculty in Linguistics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2013

This manuscript has been read and accepted for the
Graduate Faculty in Linguistics in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Date	Juliette Blevins, Ph.D. Chair of Examining Committee
Date	Gita Martohardjono, Ph.D. Executive Officer

Supervisory Committee:

Juliette Blevins, Ph. D

Gita Martohardjono, Ph.D.

William McClure, Ph.D.

Outside Reader:

Gopal Sukhu, Ph.D.

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

THE RISE OF DISYLLABLES IN OLD CHINESE:

THE ROLE OF LIANMIAN WORDS

by

JIAN LI

ADVISER: PROFESSOR JULIETTE BLEVINS

The history of Chinese language is characterized by a clear shift from monosyllabic to disyllabic words (Wang 1980). This dissertation aims to provide a new diachronic explanation for the rise of disyllables in the history of Chinese and to demonstrate its significance for Modern Chinese prosody and lexicalization.

A corpus of 300 Lianmian words in Old Chinese was compiled, including 96 Shuāngshēng words, 172 Diéyùn words and 32 Splitting-sound words. This study builds on previous morphological and phonological research on disyllables in Chinese and looks closely at detailed aspects of Old Chinese sound patterns and their evolution. Based on the analysis of sound patterns of Splitting-sound words and Dieyun words in Old Chinese, evidence from neighboring languages, statistical analysis of the development of Old Chinese, and reconstructed syllable structure, I argue that the simplification of complex onsets in Old Chinese was a central motivating factor for the rise of the earliest disyllabic forms – Splitting-sound words. Monosyllabic words with historic initial CL clusters (L a liquid), undergo fission, surfacing as disyllables where the first syllable has the simple C onset and the second the L onset. The occurrence of the liquid in the second syllable onset preserves consonant identity, which would

otherwise be lost in the onset simplification process. Generalization of this process soon gave rise to another type of mono-morphemic disyllable – Diéyùn. Once onset simplification was complete, around Late Old Chinese to Early Middle Chinese period, phonological motivation for syllabic fission disappeared. Mono-morphemic disyllables lost their productivity at this point. The disyllabic template they defined was preserved, giving rise to productive formation of disyllabic compounds. This word-formation process appears to be responsible for the dominance of disyllables in many modern Chinese languages spoken today.

This diachronic phonological research accounts for issues that previous studies fail to address. It reveals the relation between the rise of disyllables and the creation of Lianmian words, the relation between the creation of Lianmian words and the simplification of Old Chinese phonology. It enriches our understanding of the role of Lianmian words and of Old Chinese phonological development in Chinese historical disyllabicity.

ACKNOWLEDGEMENTS

I would like to express my greatest gratitude to the people who have helped and supported me throughout my five-year doctoral study. First of all, I am grateful to my advisor, Dr. Juliette Blevins, for her continuous and dedicated support of this research, from initial advice and contacts in the early stages of conceptual inception and through ongoing advice and encouragement to its completion as a doctoral dissertation. She introduced me to research of historical phonology. I have learned from her not only linguistics, but also devotion to academics and dedication to students. What's more, she sets a good example for me as to how to strike a balance between family life and academic work, leaving a life-long influence on me.

I am also deeply indebted to my other committee members, Dr. Gita Martohardjono and Dr. William McClure. Dr. Gita Martohardjono provided me with helpful comments on and suggestions for the data collection of secret languages in Modern Chinese. Dr. William McClure guided me in critical thinking. And both of them gave me a lot of support during my Chancellor's fellowship work and my dissertation fellowship application, helping me go through the five-year study program smoothly.

Special thanks also go to Dr. Gopal Sukhu, my outside reader. His insight into the issues of *Fǎnqiē* and his expertise on Chinese etymology contributed significantly to this dissertation.

I am especially grateful to Nishi Bissoondial, our Assistant Program Officer, for taking impeccable care of every administrative need for me.

I wish to acknowledge the encouragement from the following faculty at various stages of my study: Dianne C. Bradley, who guided me into the realm of phonology; Elaine Klein, who helped me with my study plan as my first mentor at the Graduate Center; Marcel den Dikken, who contributed comments to my second qualifying paper from which my dissertation developed.

Without the company of my fellow students at the Graduate Center, this dissertation would not be written so smoothly. Thanks are due to the following: Nazik Dinctopal from the Linguistics Department, for her continuous encouragement when we worked in the computer hub; Seamus Donnelly from the Psychology Department, for his editorial support; Yan Yu from the Department of Speech-Language-Hearing Sciences, for her friendship and emotional support.

I would also like to thank Linda Perrotta, the director of the Child Development Center at the Graduate Center, and her teaching team, who provide a very caring and warm educational environment for my son. Without their efforts, I could not have focused on my research.

Words cannot express the debt I owe to my parents for their enduring patience and indulgence over the course of my study. They sacrificed leisure time to help me with baby-sitting on weekdays without any complaints. Without their support, I could have never completed this dissertation. My special thanks also go to my two sons, Hong-Kang and Hong-Yee, for their smiles and independence. Due to lack of my company, they had to play by themselves most of the time. Last but not least, I would like to take this opportunity to express my deepest gratitude to my husband, Jun Xu, for his unconditional support and earnest encouragement, which accompanied me through my journey of doctoral study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

*	reconstructed forms
>	changed into
[cmn]	Mandarin Chinese. Three-letter codes in square brackets are ISO 639-3 codes from *Ethnologue: Languages of the World* (Lewis, 2009), an online encyclopedic reference work used to uniquely identify the world's languages.
μ	mora
σ	syllable
Ft	foot
.	(low dot) for syllable boundary
+	morpheme boundary
#	word boundary
·	A middle dot is used after a syllabic consonant to indicate sesqui-syllabic onset in Old Chinese
C	consonant or non-syllabic segment
V	vowel or syllabic segment
S	semivowel
O	onset
M	medial
R	rhyme
L	liquid
G	glide

$O_a \ldots O_a$	Indicates identity of subscripted constituents
C_n^m	From n to m consonants (subscript n and superscript m are integers)
OC	Old Chinese
EOC	Early Old Chinese
MOC	Middle Old Chinese
LOC	Late Old Chinese
EMC	Early Middle Chinese
LMC	Late Middle Chinese

CHAPTER 1: INTRODUCTION

Chinese is often misunderstood as a "monosyllabic" language. However, this is only partially correct. In Old Chinese and Middle Chinese, most words correspond to a single syllable and a single character. In Modern Chinese, a morpheme is usually still monosyllabic, but 70% of Modern Mandarin [cmn] words are disyllabic (Duanmu & Zhang, 2010). This historical change from monosyllables to disyllables has become a prominent feature in the development of the Chinese lexicon and aroused many linguists' interest. In this development, there exist two types of disyllabic words – mono-morphemic disyllables and disyllabic compounds, which have different developmental paths and periods. Mono-morphemic disyllables, also called Lianmian words by traditional scholars, were first produced in Early Old Chinese, around 1000 BC, and began to lose their productivity around 6th century AD. There is a phonological relationship between the two syllables of Lianmian words, but no semantic relationship. Disyllabic compounds, which can be subdivided into two morphemic units, began to increase in literature around the Qin and the Han Dynasties (221 BC – 9 AD).

However, since the phonology of Old Chinese is not well established, the motivation for the rise of Lianmian words has never been systematically analyzed, nor has the role of Lianmian words in the trend towards disyllables been discussed. In order to fill this gap, this dissertation will provide an investigation of the phonological motivations for the rise of Lianmian words, thereby establishing a disyllabic template so as to show a clear picture of what really occurred in this development of the Chinese lexicon.

1.1 The history of Chinese

This dissertation focuses on the historical change of the Chinese lexicon from monosyllabic to disyllabic words through the evolution of the Chinese language. Before approaching the core

issues, we need to clarify Chinese periodization and some complex geographic and temporal sub-designations.

In his comprehensive work on Chinese languages, Norman (1988, p.1) observes that few language names are as all-encompassing as that of Chinese. In fact, the Chinese language has always been considered a symbol of a cultural and political unity. It subsumes many historical stages and geographic variants — it may refer to "the archaic inscriptions of the oracle bones, the literary language of the Zhou dynasty (11th century BC – 771 BC) sages, the language of Tang and Song (618 AD –1279 AD) poetry and the early vernacular language of the classical novels, as well as the modern language in its standard and many dialectal forms" (Norman, 1988, p. 1).

In order to conveniently track and label the changes from monosyllables to disyllables in the Chinese lexicon in the process of analysis, this study divides the long history of Chinese phonology into the following periods adapted from Norman (1988):

- Proto-Chinese: the period preceding the earliest written documents
- Old Chinese (OC): 1000 BC – 500 AD (attested in bronze inscriptions, *Shī Jīng* 詩經 and *The Analects of Confucius* 論語)
- Middle Chinese: 500 AD – 1400 AD (attested in *Qiēyùn* 切韻 and *Jīngdiǎnshìwén* 經典釋文)
- Modern Chinese: 1500 AD – present (attested in '*Water Margin*'水滸 and '*Dream of the Red Chamber*'紅樓夢)

Old Chinese (OC) starts at the Zhou period, and is the language of the *Shi Jing* rhymes and the phonetic series (諧聲系列) in the script. Although some oracle bone inscriptions reflecting the Shang period around 1300 BC have been found, this early period is excluded from Old

Chinese, because the Oracles do not provide abundant or systematic evidence bearing on pronunciation (Norman, 1988). OC is further divided into three periods for the convenience of this study: Early OC (EOC), Middle OC (MOC) and Late OC (LOC). EOC is mainly the language of *Shī Jīng* (The Book of Odes), and MOC refers to the Qin and the Han Dynasties (221 BC – 9 AD), while LOC is after the Han period and before the Sui Dynasty (9 AD –580 AD). Early Middle Chinese (EMC) is the language of *Qiēyùn* 切韻 in Sui and *Jīngdiǎnshìwén* 經典釋文 written in early Tang dynasty (550 AD – 600 AD). Late Middle Chinese is the period when the vernacular literary language began to evolve and flourish. Modern Chinese is also divided into two periods: Old Mandarin (1500 AD – 1800 AD) and Modern Mandarin (1800 AD – present). Modern Chinese includes all Modern Chinese dialects, more like a family of languages. They are traditionally called Modern Chinese dialects only because they share the same written language, but in fact, the speakers of different "dialects" typically cannot communicate with each other through their own dialectal forms. Old Mandarin refers to a new common speech which was developed in the Yuan dynasty, based on the dialects of the North China Plain around the capital. New genres of vernacular literature[1] were based on this language, including verse, drama and story forms (Norman, 1988, p. 23). In this dissertation, Modern Mandarin generally refers to the standard official Modern Chinese developed from Old Mandarin unless otherwise specified. This above chronology is rough, but essential for understanding major historical changes in Chinese.

[1] Two distinct traditions exist in Chinese literature: the vernacular and the classical or literary. Classical literature, which can be traced back to 1000 BC, consists originally of verses and later of drama and fiction. This kind of writing placed an emphasis on phrasing. In the 13th century, 'vernacular literature' was created to make distinctions from classical literature. Vernacular literature was longer and more loosely structured, imitating the speech of 'common' people, full of detailed description and lively dialogues.

Historical linguists have tried to reconstruct the phonology, morphology and syntax of Proto-Chinese and its relationships with other languages and language groups. Although Proto-Chinese does not fall into the scope of our discussion, hypotheses on the relationship between Proto-Chinese and other languages and language families have great impact on the phonological and morphological reconstructions of OC. Some historical linguists have hypothesized that an early form of Tibeto-Burman was the language of the Yǎngsháo 仰韶 culture (6900 BC – 6000 BC), in the middle and upper reaches of the Yellow River. The language of the Shāng was influenced by the Austroasiatic language of the Dōngyí 東夷 peoples in Shandong and to the south, and an early form of Miao-Yao was spoken in the middle Yangtze (Van Driem, 1998). Given these hypotheses, it is possible that contact between Sinitic and non-Sinitic languages has a great impact on the development from Proto-Chinese to OC.

From a diachronic perspective, the OC that evolved from Proto-Chinese preserved some features, making it resemble other neighboring related languages or potentially related languages more closely than Modern Chinese does. Given this background of Proto-Chinese, this dissertation will also explore the external evidence relating to the rise of disyllables in OC in neighboring languages, most of which do not belong to the Sino-Tibetan language family.

1.2 The development of disyllables

1.2.1 Identifying Chinese words

Conventionally, Chinese linguists often define 'a Chinese word (*cí*)' with semantic or syntactic criteria. The most widely accepted definition is Wang's: *cí* in Chinese is distinguished as the 'the smallest independently usable part of language' or 'the smallest part of the sentence that can be used independently' (1953, p. 3, translated by Packard, 2000). However, the intuition of many a Chinese speaker is that words simply do not exist in Chinese (Packard, 2000, p. 17);

this comes from the 'fluidity' of the word-morpheme boundary existing in the minds of Chinese speakers (Hoosain, 1992). Hoosain explains that both diachronic and synchronic factors lead to the problem of fluidity: first, many morphemes that are now bound were free in OC and Middle Chinese; second, the bound-free status of morphemes varies across dialects; third, there is great variation in bound-free status of morphemes according to context (Hoosain, 1992, p. 118-20). Due to these three factors, the interpretation of Wang's definition varies across times, dialects and even contexts.

In addition to Wang's definition, Chinese phonologists define 'Chinese words' from a prosodic view. For example, Duanmu (1999a) came up with 'minimal word' (also called S-foot[2]) which is defined as 'a disyllabic trochee'. A minimal word can be either a disyllabic trochee, or a monosyllable followed by a zero syllable. "The zero syllable does not always have to be realized as silence but can be filled by lengthening the preceding monosyllable" (Duanmu, 1999a, p. 19). Although there is still a debate on whether this 'disyllabic unit' should be identified as a foot or prosodic word, Chinese linguists all agree that there is a disyllabic unit that controls the metrics of Chinese speech[3]. This hypothesis coincides with Chinese speakers' intuition about the concept of 'word', which conflicts with the 'word' in morphological or syntactic analysis.

Compared with 'word', morphemes are conventionally considered much easier to distinguish in Chinese, for one Chinese morpheme often corresponds to one syllable and one character. However, this view has recently been challenged by a new morphological derivational approach. In his book - The Roots of Old Chinese (1999), Sagart presents a 'radical' morphological reconstruction that focuses on reconstructing OC roots and affixes. Based on Baxter's phonological reconstruction, Sagart believes that most OC words are morphologically

[2] S-foot refers to a syllabic trochee, which is distinguished from a moraic trochee, called M-foot.
[3] Under this prosodic interpretation, a disyllabic unit can be a mono-morphemic disyllable, a disyllabic compound or a disyllabic phrase, in the domain of which tone sandhi can be applied in Modern Mandarin.

complex. For example, he claims that "all OC word-initial consonant clusters are formed of one root-initial consonant plus one or several consonantal prefix(es) and/or infix" (Sagart, 1999, p. 21). He gives further explanations about the morphological functions of these affixes. "*k- added to a verb seems to produce a concrete and countable noun of related meaning" (Baxter & Sagart, 1997, p. 47). "*t- added to a countable seems to produce a derived mass noun" (Baxter & Sagart, 1997, p. 52).

Based on Haudricourt's hypothesis that tones derive from earlier syllable final segmental phonemes, Sagart argues that "besides the phonological function of producing Middle Chinese alternations between forms in qùshēng [departing tone] and forms in other tones, the morphological function of suffix *-s is to derive a noun from a verb or an adjective" (Sagart, 1999, p. 54).

Sagart also proposed that "the function of the infix *-r- was to mark forms that were plural or collective, iterative, durative, or indicating effort (in the case of action verbs), or intensive (in the case of stative verbs and adjectives)" (Sagart, 1999, p. 62).

Many of Sagart's affixation examples are supported by evidence from typologically similar languages, including Phanrang Cham [cjm] (Austronesian/Chamic) and Laven [ibo] and Nyaheun [nev] (Austro-Asiatic/Mon-Khmer/West Bahnaric) and from surviving Modern Chinese dialects, but his view that OC had a great deal of affixation is still controversial. "Out of the first one hundred reconstructions listed in Sagart's 'List of Reconstructions' (235), 70 contain word-initial clusters. 50 contain prefixes and 29 contain the infix *-r-" (Miyake, 2001, p. 261). Furthermore, the morphological function of many 'affixes' is still uncertain or unclear. This study sides with the prevailing view that most consonant clusters in OC should be treated as phonologically complex, but morphologically simple. Since, under all but Sagart's analysis all

Chinese morphemes are free morphemes, and, since the morphological status of potential bound morphemes is not at issue, for the remainder of the thesis, the term "mono-morphemic" will be used to refer to all OC monosyllables.

1.2.2 What are disyllables in Chinese?

Many researchers do not clarify the characteristics of Chinese disyllables before exploring their origin and rise, and thus fall into biased presumptions. There are two morphological types of words in Chinese: mono-morphemic and multi-morphemic words. Mono-morphemic words can be further divided into mono-syllabic, disyllabic and polysyllabic words. All three types are illustrated in (1). Polysyllabic mono-morphemes are rarely seen and most are loanwords, like the polysyllabic word *qiǎokèlì* 巧克力, a borrowing of English *chocolate*.

(1) Mono-morphemic words in Modern Chinese

	Mandarin	South Min[4]	Shanghai	Characters	Glosses
	[cmn]	[nan]	[wuu]		
Mono-syllables	bái	biak8	baq	白	'white'
Disyllables	húdié	o2diap8	wu diq	蝴蝶	'butterfly'
Polysyllables	qiǎo kèlì	ka3kiak7lat8	chio k^{h}eq liq	巧克力	'chocolate'

Mono-morphemic disyllables make up only about 0.3% of the Chinese lexicon[5], approximately 1160 mono-morphemic words. Most of them were created through phonological patterns like total reduplication[6], and partial reduplication[7] (conventionally called Lianmian

[4] The number following each syllable indicates the tone of that syllable in South Min.
[5] According to *Hànyǔdàcídiǎn* [Chinese Dictionary] (1993), there are approximately 375,000 polysyllables.
[6] Normally, total reduplication creates bi-morphemic words. But these reduplicated words in EOC are mostly onomatopoeic and lack base forms. So they are considered mono-morphemic, doubled syllables within one morpheme.
[7] Most Lianmian words are also considered mono-morphemic disyllables, since some of the words do not have base forms, or some of the 'bases' have different, unrelated meanings. Hu, for example, states that "in the beginning, whether for clarification or for repetition, disyllables formed through partial reduplication were used to express one word" (translated from Hu, 1923, p. 30-31, cited in Wang, 1972, p. 47).

words) in EOC. Among the Lianmian words which were formed through partial reduplication, there are two categories traditionally termed as Shuāngshēng 雙聲, and Diéyùn 疊韻. Shuāngshēng refers to words whose two syllables share the same initial (the initial consonant or consonant cluster). Diéyùn refers to words in which the two syllables have the same rhyme, including the vowel and the nasal or approximant coda (Baxter & Sagart, 1997; Sun, 1999)[8].

(2) Lianmian examples

Words	Reconstructed sound	Glosses	Sound pattern	Sources
關關	*kron kron	'the cry of ospreys'	total reduplication	*Shījīng*
參差	*ts^{h}rjum ts^{h}rjaj	'uneven'	Shuāngshēng	*Shījīng*
窈窕	*ʔiwʔ liwʔ	'elegant, beautiful'	Diéyùn	*Shījīng*

Lianmian words also include another type of word – Splitting-sound words, which are often ignored in some researchers' discussions, because they are not formed through partial reduplication. In this thesis, Splitting-sound words are classified as Lianmian words because they are mono-morphemic disyllables, and there are some sound correlations between the two syllables. Their special sound patterns will be examined and analyzed in detail in Chapter Four. In Modern Mandarin Lianmian word-formation is not productive any more[9]. Even so, the generalization that morphemes are consistently monosyllabic in Chinese is incorrect: mono-morphemic disyllables like those in (2) and polysyllables like *qiǎokèlì* 巧克力 in (1) exist.

In Modern Mandarin, bi-morphemic words dominate. Most bi-morphemic words are disyllables originating from disyllabic phrases. Here they are called 'disyllabic compounds'.

[8] The definitions of Shuāngshēng and Diéyùn that are adopted by many Modern scholars and linguists still need further investigation (see details in Chapter 3).
[9] Total reduplication is still productive in Modern Chinese, while Lianmian word-formation ceased to be prodctive in EMC.

(3) Disyllabic compounds:

Word	Glosses	Morpheme1	Morpheme 2
zǒu-shī 走失	to be missing	zǒu- ‘to go’	-shī ‘to lose’
zǒu-sī 走私	to smuggle	zǒu- ‘to go’	-sī ‘illicit’
pá-chǎng 排場	excessive display of splendor	pá- ‘line, to arrange’	-chǎng ‘field’
biān-pái 編排	to arrange	biān- ‘plait, to arrange’	-pái ‘line, to arrange’

In sum, though OC was primarily a mono-syllabic language, disyllables did exist. In Modern Chinese there are now two kinds of disyllables: mono-morphemic disyllables and disyllabic compounds. This study attempts to uncover the diachronic factors giving rise to Lianmian words and for the transition from Lianmian words to disyllabic compounds around LOC.

1.2.3 Disyllabicity

A prominent feature of the historical development of the Chinese lexicon is the change from monosyllabic to disyllabic words. Here I borrow the term ‘disyllabicity[10]’ from Feng (1998) who coined the word to describe this state. Historical morphological research tells us that in Proto-Chinese, as preserved in Oracle bones, one character basically corresponds with one word[11] (S. Wang, 1992). This strongly suggests that monosyllabic morphemes abound in Proto-Chinese. From the *Shījīng* period (EOC), disyllables began to increase in written documents. Li, Shichun (2007) counted the monosyllables and disyllables in two works, *The Analects of Confucius* 論語 and *Mencius* 孟子, written in EOC. In his study disyllables take up just 15.90% and 27.69% of

[10] Feng used ‘disyllabicity’ as an equivalent for the Chinese term ‘shuāng yīn huà(雙音化)’ ‘*double-sound-change*’ that has been widely adopted by Chinese scholars.

[11] Some linguists argue that in the Oracles there are some compound characters (合體文), expressing a phrase in one character form. Most of these compound characters have been dismantled or reconstructed into two characters in writing (Y. Pan, 1989). But this still does not change the one-to-one correspondence between characters and words.

the corpus respectively[12]. In Middle Chinese, the distribution is quite different, with disyllables going up to 40-44% of the corpus (Qiu, 2010). And in Modern Mandarin, disyllables take up 72% of the lexicon (Duanmu & Zhang, 2010). I compared Li's statistics with those of Qiu's and Duanmu's studies. The result is shown in the following figure:

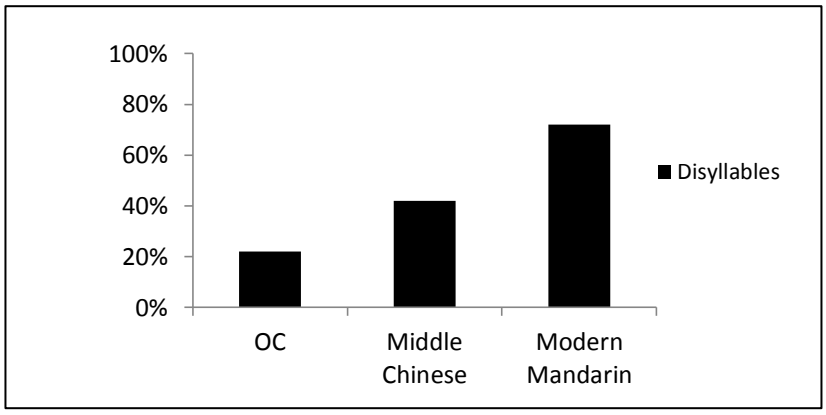


Figure 1.1 Distribution of the percentage of disyllables in the lexicons of OC, Middle Chinese and Modern Mandarin

Some characteristics of this chronological change from mono-syllables to disyllables are shown in the following Figure 1.2:

[12] The average percentage of disyllables in these two works is approximately 22%.

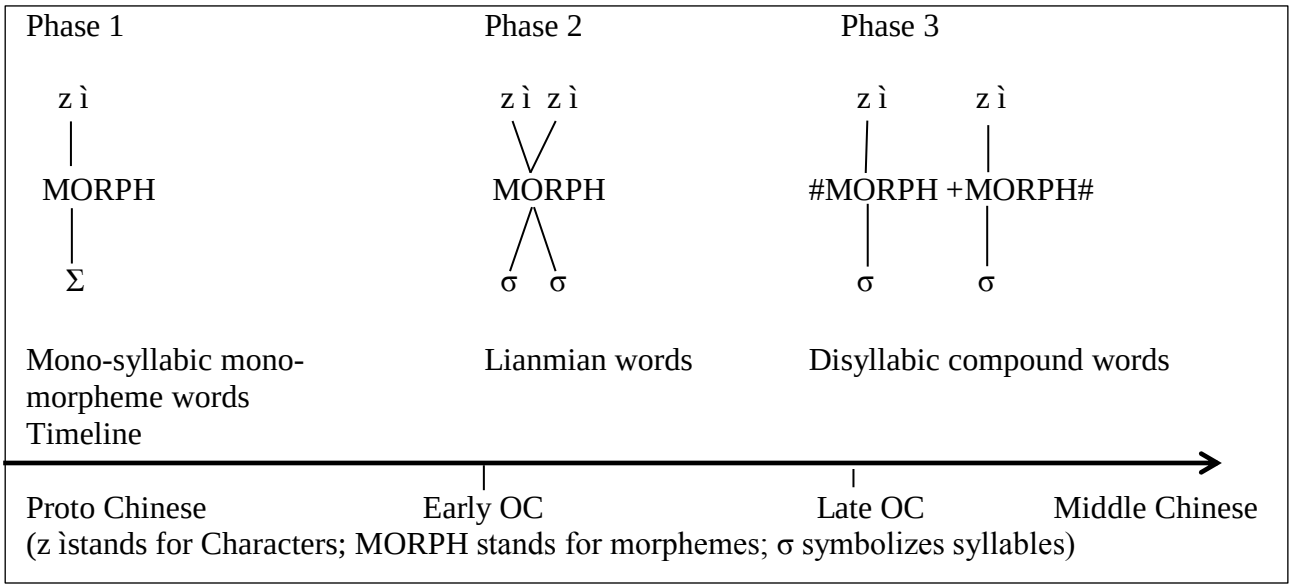


Figure 1.2 Three stages in the evolution of Chinese words

In Figure 1.2, the rise of Lianmian words in Phase 2 and disyllabic compounds in Phase 3 does not result in the total replacement of monosyllables. According to Li's statistics, disyllables were not in the majority until Modern Mandarin (the early 20th century), though they started to rise in OC. Note that figure 1.2 does not incorporate Sagart's morphological analysis. In his view, Phase I includes multi-morphemic mono-syllabic words. Since all of his proposed syllable-internal productive morphology was lost, and does not appear to be relevant in the evolution of lianmian words, it is, for the most part disregarded in this thesis.

This continuous shift from monosyllabic to disyllabic words has been the focus of much linguistic research. Many researchers have attempted to explain its characteristics and its origin, including aspects of: morphology (Pan, 1989; Xu, 2003), semantics (Cheng, 1992; Xu, 2003), phonology (Pan, 1999; Jin, 2002), prosody (Arcodia, 2007; Feng, 1998) and sociolinguistics (Cheng, 1992; Pan, 1989; Xu, 2003). However, few researchers have closely examined the developments of the two kinds of disyllables: Mono-morphemic disyllables and disyllabic compounds. The two kinds of disyllables did not follow the same revolutionary path. As

mentioned above, mono-morphemic disyllables or Lianmian words were created in EOC, much earlier than the formation of disyllabic compounds. But these disyllables formed through certain phonological rules, especially the 'Lianmian' words, were seldom productive after LOC (Cheng, 1992; Feng, 2001). With the decline of Lianmian words, disyllabic compounds developed in MOC and LOC periods. Later, disyllabic compounding became the most productive word-formation tool. As a result, disyllablic compounds make up the majority of Modern Mandarin disyllables and also out-number mono-syllabic words in common usage.

1.2.4 Exemplification

Disyllabicity refers to the change from mono-morphemic monosyllables to disyllabic words in the historical development of the Chinese lexicon. An example is given below in (4) to illustrate the change and its characteristics.

(4) Diachronic Word change:

a. One character - One syllable - One morpheme - One word

飐 */pew/[13] 'whirlwind' OC

b. Two characters - Two syllables - One morpheme - One word

扶遥 */pa lews/ 'whirlwind' EOC to EMC

c. Two characters - Two syllables - Two morphemes - One compound

飘風 /piao$_{55}$ fəŋ$_{55}$/ 'whirlwind', 'wind' Modern Chinese

(4a) shows a typical mono-morphemic and monosyllabic word in OC. (4b) is a disyllabic mono-morphemic word, representing splitting-sound words, a special kind of Lianmian word. Although

[13] Zhengzhang's reconstruction is adopted here, which is provided by www.eastling.org database. According to Baxter and Sagart's reconstruction (version 1.00 at http://crlao.ehess.fr/document.php?id=1217), 飐 was *praw and 扶遥 was *m-pra law. Based on this reconstruction, Sagart proposed that **-r-* was an infix, and that **m-* was a prefix whose function here was not clear. If so, the word 飐 includes two morphemes (root+infix), while 扶摇 *m-pra law has four morphemes (prefix+root+infix+root).

most splitting-sound words fell out of productivity gradually around LOC and EMC, some are still preserved in classical literatures, like */pa lews/ 扶遙 in (4b). (4c) represents a typical disyllabic compounding case in Modern Chinese (including Old Mandarin and Modern Mandarin). Although /piao$_{55}$/飆 itself means whirlwind, a redundant monosyllabic morpheme /fəŋ$_{55}$/ 风 'wind' is added after it to meet the disyllabic template. /piao$_{55}$/飆 has become a bound morpheme. In Modern Mandarin, most disyllables are such compounds or pseudo-compounds[14].

The above examples illustrate Packard's (2000) 'fluidity' feature of the morpheme-word relationship in the history of Chinese. These examples also illustrate disyllabicity. There are two different processes at work: one turns mono-morphemic monosyllables to mono-morphemic disyllables (Lianmian words); and the other forms disyllabic compounds from combination of mono-morphemic monosyllables. Although the starting point is a language with monosyllabic morphemes, and the ending point is also the stage of monosyllabic morphemes, disyllabicity is not a smooth or simple process. This change not only complicated the concept of 'Chinese words', but also caused a dramatic revolution in the Chinese lexicon, leading to a dual-vocabulary phenomenon in Modern Mandarin (Duanmu & Zhang, 2010): 40% of words in Modern Mandarin have both a disyllabic and a monosyllabic form, with the choice between the two forms dependent on the disyllabic prosodic template mentioned above in 1.2.1.

1.3 Research goals

The primary aim of this study is to investigate possible phonological motivation for the rise of disyllables in OC. In this process, comparative and typological evidence will be used to support the phonological change in OC which is critical in the rise of disyllables. In this way, the current study will contribute to the comparative study of OC and its neighboring languages.

[14] The notion of pseudo-compound was first brought up in Trommelen & Zonneveld (1989) to refer to compound-like words where at least one of the constituents is not a free morpheme and cannot occur outside of compounds.

Identifying the earliest disyllabic form and the starting time of the rise may be the preliminary step to solving the problem. Unlike traditional research on the rise of disyllables, the current study considers Lianmian words (mono-morphemic disyllables) instead of disyllabic compounds as the earliest disyllabic form. One central goal of this dissertation is to investigate closely the phonological patterns of Lianmian words and to provide an explanation for the relation between the decline of Lianmian words and the rise of disyllabic compounds around LOC and EMC. The final goal is to investigate the cause for the rise and development of disyllables on the Modern Chinese lexicon and prosody.

1.4 Outline

The overall geography of this dissertation is as follows. Chapter 2 prepares the reader with the traditional view of Chinese syllable structures and the terminology relevant to Chinese syllable structure that I am going to adopt in the dissertation. It also provides knowledge of OC reconstruction, previous reconstructed OC phonological systems, and the system that is going to be used in this study. The Lianmian word data corpus that serves as a basis of analysis is introduced in Chapter 3. Chapter 4 and 5 constitute the bulk of the dissertation, in which the relation between the simplification of consonant cluster onsets that started in EOC and the rise of Lianmian words is investigated in detail. More specifically, Chapter 4 analyzes the influences of the simplification of phonological structure in EOC on the rise of Splitting-sound words. Chapter 5 discusses the formation of a disyllabic template and how the template spreads from Splitting-sound words to other Lianmian words. Chapter 6 examines the characteristics and the development of disyllabic compounds that appear in MOC and LOC and discusses the role of the disyllabic template in the transition from Lianmian words to disyllabic compounds. This chapter also touches upon the influence of Lianmian words on Middle Chinese and that of the disyllabic

template on Modern Chinese varieties. Chapter 7 reviews previous researches on the rise of disyllables in OC, covering functional explanations (homophony avoidance), phonological explanations and prosodic explanations. After comparing their pros and cons, I conclude that the diachronic approach adopted here is the most satisfactory solution, providing a reasonable and thorough explanation for the range of facts associated with the rise of disyllables. Lastly, Chapter 8 concludes the dissertation with a summary of the major findings and discussion of the contributions of the current study as well as directions for future research.

CHAPTER 2: OLD CHINESE PHONOLOGY

2.1 OC reconstruction

2.1.1 OC reconstruction methods

A native tradition of Chinese phonology developed in the Middle Chinese period. Chinese linguists had long since compiled dictionaries and attempted to clarify the pronunciation of difficult characters by specifying homophonous characters. During the early centuries AD, the more advanced method of fǎnqiē 反切 was developed (see details in section 2.2). By way of fǎnqiē, systematic rhyme dictionaries such as *Qiēyùn* 切韻 (601 AD), were compiled. During the next centuries when Buddhism's influence was increasing in China (7th century – 11th century AD), Chinese linguists borrowed Sanskrit grammatical traditions of analyzing sounds in terms of distinctive features such as place of articulation and phonation type. This led to the creation of rhyme tables. A rhyme table is a Chinese phonological model, usually adopted in rhyme dictionaries. In these dictionaries, characters are grouped by the four tones and then into rhyme groups. Distinct syllables are presented in a number of tabular charts, each with a groups of characters recorded as homophones. The reconstruction of Middle Chinese phonology was established, in part, on the basis of rhyme tables, rhyme dictionaries such as *Qiēyùn* 切韻 (Pulleyblank, 1984, 1992). These reconstructions were widely agreed upon. The reconstruction of OC is more controversial than that of Middle Chinese since it has to be extrapolated from the Middle Chinese data. However, scholars of Chinese historical linguistics agree that reconstruction of Old Chinese is viable. Reconstructions typically make use of at least five kinds of evidence: (1) phonetic series (xiéshēng 諧聲); (2) rhyme groups (yùnbù 韻部); (3) MC phonology and Modern spoken Chinese varieties; (4) phonology of ancient loanwords; (5) non-

Sinitic cognates (including cognates from Sino-Tibetan languages and potential cognates from Austroasiatic languages.

A phonetic series in Chinese is based on phonetic components of the written characters. In contrast to the popular notion of Chinese as a primarily pictographic or ideographic language, the vast majority of Chinese characters (about 95 percent of the characters in the *Shuowen Jiezi* 說文解字) are constructed as either logical aggregates[15] (hu ì y ì z ì會意字) or, more often, phonetic complexes which are also called pictophonetics (x íng-shēngzì 形聲字) (DeFrancis, 1984, p. 84). Phonetic complexes or pictophonetics are characters containing two parts, one of which indicates a general category of meaning and the other the sound. For example, in the character 跑 pǎo, 'to run', the left-hand component 足 'foot' suggests the meaning, while the right-hand part 包, a character pronounced 'bāo' by itself, lends the pronunciation. If this process of forming phonetic complexes is continued, a phonetic series of characters can be formed, where all the characters in a series share a common phonetic element, but have different meanings. We may compare words whose characters share a phonetic element (a phonetic series) to yield much information about pronunciation. Usually the characters in a phonetic series are still pronounced alike. For example, the character 中 (zhōng, "middle"), was adapted to write the word chōng ("pour", 冲) and zhōng ("loyal", 忠). In other cases the words in a phonetic series are pronounced quite differently in all known varieties of Modern Chinese or Middle Chinese, but are assumed to have been similar in OC at the time the characters were chosen (Norman, 1988, p. 43-44).

[15] Characters of logical aggregates combine pictograms to indicate an abstract concept. For example, 人 (**njin* in OC r én in Pinyin) is a pictogram of a person, and putting two 人 together makes 从 (**zloŋ* in OC c óng in Pinyin), meaning 'to follow'. Combining 日 (**njig* in OC r ìin Pinyin) 'sun' and 月 (**ŋod* in OC yuè in Pinyin) 'moon' makes 明 (**mraŋ* in OC míng in Pinyin) 'bright', which is traditionally interpreted as symbolizing the combination of sun and moon as the natural sources of light.

Another traditional source of evidence in Chinese historical linguistics is the rhyming practice of the *Shījīng* 詩經 (The Book of Odes), a collection of songs and poetry compiled in the 6th century BC. In this book, some of the lines still rhyme today, but many do not. This fact had been attributed to lax rhyming practice, until the late-Ming Dynasty scholar Chen Di 陳第 argued that a former consistency had been obscured by sound change from OC to Middle Chinese. The systematic study of these rhymes began with Gu, Yanwu (顧炎武) in the 17th century. He divided them into ten rhyme groups (yùnbù 韻部). Other scholars later kept refining these groups until in the 1930s these groups culminated in a standard set of 31 rhyme groups. One of these scholars, Duan, Yucai (段玉裁) stated the important principle that characters in the same phonetic series would be in the same rhyme group, making it possible to assign rhyme groups to almost all characters (Baxter, 1992, p. 150-170; Norman, 1988, p. 42-44).

Middle Chinese phonology was established on the basis of rhyme dictionaries and rhyme tables. *Qiēyùn* is the first important rhyme dictionary written in 601, with many revisions and expansions over the following centuries. It reflects the codification of an elite standard pronunciation common to educated speakers from both north and south in the late Nánběicháo (420-589) period (Sun, 1999, p. 10). "Early Middle Chinese" refers to this system. Rhyme tables (děngyùntú 等韻图) originally designed as keys to *Qiēyùn* by Buddhists, present a more detailed analysis of initials and finals in terms of place of articulation ("seven sound divisions"), manner of articulation, the height of vowels and the different properties of main vowels. It provides a synchronic description of the standard pronunciation in mid- and late Tang, which represents the Middle Chinese pronunciation system. Traditional scholars of Chinese historical phonology usually based their reconstruction of OC phonology on the comparison of the phonetic series and rhyme groups with the phonological system of Middle Chinese.

The correspondences and sound changes reflected in Chinese varieties can also be considered as internal evidence for Middle Chinese and OC reconstructions. "When a language undergoes changes, traces of the changes are often left behind in the language's structure, as allomorphic variants or irregularities of some sort" (Campbell, 2004). Of all Modern Chinese varieties and dialects, Min language appears to be the most important since as a language of people of the coast and mountains who have long been isolated from the political and cultural center of China, it preserves some of the OC features that disappeared in Middle Chinese. For example, in Middle Chinese and Modern Mandarin, 寒 hán 'cold', 糊 hú 'paste' and 懷 huái 'hold in the bosom' are all pronounced as syllables with a fricative initial. But in Min, their initials are all the plosive k-. Based on this, as well as other evidence, we can reconstruct OC onset *g- for that syllable. Then we can say that this *g- onset is well preserved as a plosive in Min[16] (Sun, 1999:10). Another example is from the Pingnan dialect of South Min, which still preserves the OC voiceless nasal *ŋ^{h}- in its consonant inventory. This sound has disappeared in Middle Chinese (Li, 1999).

Scholars also attempted to determine the phonetic content of OC by examining pronunciation of loans in Korean [kor], Japanese [jpn] and Vietnamese [vie] (Sinoxenic materials). Foreign words that were taken into OC also provide helpful data for comparative research on OC reconstruction. For instance, the famous loanword found in *Hànshū* 漢書 by Ban Gu 班固 (AD 32-92) - the proper noun "Alexandria" - was transcribed in Chinese character as 烏弋山離(wū y ìshān l íin Pinyin) in OC period. The four Chinese characters have been well preserved from OC to Modern Chinese, but the sounds underwent great changes. 烏弋山離, according to *Qiēyùn*, is easily reconstructed as **ʔɔ jik ʂəɨn lia* in EMC, which does not match the

[16] It should be noted that the voiced obstruents have been systematically devoiced in Min.

donor sounds very well. These discrepancies reflect the sound changes from OC[17] to Middle Chinese (Li, 1980, p. 13; Sun, 1999, p. 14). Liquid *l exists in OC inventory. So we have the reason to assume that the second syllable onset of OC pronunciation of 'Alexandria' is *l. In the EMC pronunciation of this loanword, the second syllable onset is *j. This reflects the palatalization of OC *l: *l > *j.

The fifth source is external evidence from Chinese's sister language – Tibetan [bod] and other related languages. In Tibetan, there are some hundreds of proposed cognate words, including highly basic vocabulary (Coblin, 1986). For example, Sino-Tibetan comparison proves very helpful in resolving the OC reconstruction of liquid onsets *l- and *r-. It had been accepted that these two liquid onsets existed in OC. But how to assign these two onsets to OC syllables was a problem. Pulleyblank (1962, 1992) and Baxter (1992) discovered the developmental paths of the two liquid onsets from OC to MC on the basis of Jeon's cognates in Tibetan and Chinese (Jeon, 1996). They postulate that the Middle Chinese *l- should be traced back to the OC *r- and the Middle Chinese *j- to the OC *l- (Sun, 1999; Boltz, 2002). This conclusion could also be reflected in the above mentioned loanword case – "Alexandria".

2.1.2 The OC reconstruction system adopted in this study

The most widely-accepted OC reconstruction system is based on the work of Li, Fang-Kuei (1971, 1980) and William Baxter (1992). Their studies, especially Baxter's, differ dramatically from Karlgren's pioneering work and "now make OC look like a very run-of-the-mill Tibeto-Burman language from the Himalayan perspective" (Van Driem, 2005, p. 301). Li and Baxter's studies are in substantial agreement on such core issues as pre-initials[18], consonant-cluster onsets, main vowels, codas, and post-codas. This study adopts an updated reconstruction system –

[17] 烏弋山離(wū yì shān lí in Pinyin) is reconstructed as **qaa lɯg sreen reels* in OC by Zhengzhang and **ŋa ljək sren rejs* by Baxter, according to the database provided by www.eastling.org.
[18] This term is used by Sagart to indicate the prefix consonant before the complex onset.

Zhengzhang's system, which accepts Li's reconstructed inventory as "basically correct" and adds a few modifications (Zhengzhang, 2003). However, whenever there are significant differences among different reconstructed systems, especially with respect to the complex onset in my data, I will lay them out for comparison.

Focusing on the development of the consonant-cluster onsets in OC, the present dissertation is going to discuss Li, Baxter and Zhengzhang's reconstructions of the complex onsets in detail and clarify why Zhengzhang's reconstruction is used in the study.

Table 2.1 Li's reconstructed consonant inventory for OC

	Bilabial	Alveolar	retroflex	Palatal	Velar	glottal
Plosive	p b	t d			k g k^w g^w	ʔ $ʔ^w$
	p^h	t^h			k^h k^{hw}	
Nasal	m	n			ŋ	
	m^h	n^h			$ŋ^h$ $ŋ^{hw}$	
Affricative		ts dz ts^h				
Fricative		s				h h^w
Lateral approximant		l l^h	ɻ			

(Based on Li, 1980)

Li noticed the existence of consonant-cluster onsets in OC, but he thought that the problem was so complicated that he only reconstructed *s- clusters in onset position. Li also hypothesized that *s- could be considered as a prefix, playing an important role in OC morphology (Li, 1980, p. 25).

Baxter added three more consonants (*x, *ɦ and *z) to Li's reconstruction of the OC consonant inventory in his 1992 book. His reconstruction enriched the complex onset types. He

reconstructed *s-, *S- *ɦ-, and *N- clusters. He wrote capital *S- in those occasional cases where *s- clusters appear to metathesize to form affricates in Middle Chinese, such as *Sg^{w}- > dz-. *N- here represents a nasalizing element which produces a nasal reflex in Middle Chinese, e.g. *Nk- > ŋ-.

Although not referring to Baxter's system, Zhengzhang's study arrived at a surprisingly similar conclusion to Baxter's[19]. In Zhengzhang's reconstructed consonant inventory, *ts, ts^{h}, dz disappeared. He added two more fricatives *s^{h}, *z. He also hypothesized that glottal obstruents *ʔ, h, ɦ developed in LOC and gradually replaced the uvular *q, q^{h}, G (Zhengzhang, 2003, p. 85). In his consonant system, there are 25 basic consonants plus two glides *w and *j.

Table 2.2 Zhengzhang's reconstructed consonant inventory in OC

	Bilabial	Alveolar	retroflex	Palatal	Velar	Uvular	glottal
Plosive	p b	t d			k g	q G	ʔ
	p^{h}	t^{h}			k^{h}	q^{h}	
Nasal[20]	m	n			ŋ		
	m^{h}	n^{h}			ŋ^{h}		
Fricative		s z					h ɦ
		s^{h}					
Approximant	w	ɹ		j			
		ɹ^{h}					
Lateral approximant		l					
		l^{h}					

(Based on Zhengzhang, 2003)

Among the 27 consonants, 11 consonants *-m, -n, -ŋ, -ʔ, -s, -b, -d, -g, -l, -w, -j could appear in the coda position. They also constitute 10 coda clusters: *-mʔ, -nʔ, -ŋʔ, -ms, -ns, -ŋs, -bs, -ds,

[19] Before China carried the reform and opening-up policy, any academic exchange between China and the outside world had been blocked. Baxter's works were not accessible to Zhengzhang, and Zhengzhang's studies could not be disseminated either until 1982 (Zhengzhang, 2003).

[20] Zhengzhang established a set of voiceless nasals based on dialects of Miao and Mien, OC dialects, and on Li's reconstruction. In Li's account, *m^{h}, n^{h}, ŋ^{h}, were written as *hm, hn, hng clusters (Zhengzhang, 2003).

-gs, -wg.

This dissertation adopts Zhengzhang's reconstruction system, mainly because he is the first phonologist who provides a clear structural typology of complex onset among the proponents of 'consonant cluster' view of OC. It has been widely accepted by recent reconstruction studies that complex onsets exist in OC (Baxter, 1992; Ding, 1979; Li, 1980), but few OC phonologists analyzed the phonological structure of the complex onsets. Zhengzhang presents a clear structural typology of complex onsets based on their different origins and developmental pathways. He divides the complex onsets into three components: 'Crown', 'Head' and 'Tail'.

Table 2.3 Complex onset structure in OC

	Onset		
	Crown (optional)	Stem	
		Head (obligatory)	Tail (optional)
Consonant	*s-, h-, ʔ-, ɦ-, m-, n-, ŋ-, r-, p-, t-, k-	All consonants except *j, w, h, ʔ,-ɦ	*-l, -r, -j, -w, -wl, -wr, -wj, -wlj, -lj

(Based on Pan, 2002; Zhengzhang, 2003)

Compared with Baxter's reconstruction of complex onsets, Zhengzhang greatly enriched the 'Crown' part by adding three plosives *p-, t-, k-. In Zhengzhang's reconstruction, 'Crown' is considered as a reduced minor syllable of sesqui-syllables. Sesqui-syllable is a term coined by James Matisoff (1973), meaning 'one-and-a-half syllables.' A sesqui-syllabic word is composed of a major syllable proceeded by a minor one. Sesqui-syllables are most commonly found in Mon-Khmer [mkh] (Austroasiatic) languages, but other languages from other families (Austronesian, Sino-Tibetan [sit]) have them, too (Thompson, 2010). Proto-Tibeto-Burman reconstructed by Matisoff (2003) and Proto-Tai reconstructed by Pittayaporn (2009) also include

this syllabic structure. Zhengzhang's reconstruction is based on the assumption that Proto-Chinese is also a language with sesqui-syllables, and that the minor syllable in the sesqui-syllables underwent reduction in the transition from Proto-Chinese to OC: the vowel in the minor syllable was deleted, and only a syllabic consonant remained. This hypothesis is supported by the evidence gleaned by Sun, H. (1982) and Pan (1999) from neighboring languages, such as Darang Deng [dat] (Tibeto-Burman branch of Sino-Tibetan family) and Palaung Shwe [pll] (Mon-Khmer branch of Austroasiatic family), where the sesqui-syllables are usually reduced to what some analyze as a syllabic consonant or syllable appendix (Vaux & Wolfe, 2009). For example, *mˈphrŭh* 'to blow fire pipe', *kˈphru* 'to spray' and *s-la* 'pants' in Palaung; *g-don* 'bald', *s-mon* 'fennel' and *k-lno* 'chisel' in Burmese [mya] (J. Wang, 1984). Zhengzhang's sequi-syllable interpretation of the Crown solves the problem of sonority hierarchy of a syllable onset. Table 2.3 tells us that nasals *m-, *n-, *ŋ- and liquid *r- could appear in Crown position. If interpreted as a common consonant cluster onset, the combinations of *mpl-, *nkw- and *rt^h- would violate the sonority hierarchy constraint - SSP[21] (Pan, 1999; Zhengzhang, 2003). When Crowns are considered as reduced minor syllables, Head + Tail clusters always obey the sonority hierarchy. Therefore, clusters with Crowns may be called sesqui-syllabic clusters.

Four different combinations of elements are allowed in this consonant cluster scheme: C (which stands for a simple Head consonant), sC (a Head consonant proceeded by a Crown – sesqui-syllable clusters), Cl (a Head consonant plus a Tail) and sCl (which represents the most complex onset: a Head consonant with a Crown and a Tail – sesqui-syllable clusters). Crowns

[21] A sonority hierarchy is a ranking of speech sound by amplitude. The sonority ranking, from the highest to the lowest, is the following: vowels > approximants (glides and liquids) > nasals > fricatives > affricatives > stops. It is important when analyzing syllable structure. Sonority S Principle (SSP) is formulated to control why segments may appear in onsets or codas together in terms of their sonority values. SSP states that sonority has to fall toward both edges of the syllable. That is to say, the onset clusters should constitute a progressively increasing sonority value toward the syllable peak – the syllable nucleus (Selkirk, 1984).

are further divided into five categories: sibilants *s-, laryngeals *ʔ-, *h-, *ɦ-, nasals *m-, *N-, liquids *r- and stops *p-, *t-, *k-. The possibilities of sesqui-syllabic clusters are listed in the following tables:

Table 2.4 Sesqui-syllabic clusters with Crown *s- in OC (sC structure)

Crown	Head		Reconstructed OC (*)	Character	Pinyin	Gloss
*s-	nasals	-n-	*s-njo*	需	xū	'await'
		-m-	*s-mewʔ*	小	xiǎo	'small'
		-ŋ-	*s-ŋaa*	蘇	sū	'a species of thyme'
	liquids	-r-	*s-rɯg*	嗇	sè	'miserly'
		-l^h-	*s-l^{h}um*	心	xīn	'heart'
		-l-	*s-l^{h}ods*	兑	dùi	'one of the eight trigrams of baguah'
	plosives	-k-	*s-keeb*	浹	jiā	'full'
		-k^h-	*s-k^{h}ooŋ*	譔	zhuàn	'to teach wholeheartedly'
		-g-	*s-geŋʔ*	阱	jǐng	'trap'
		-b-	*s-baŋs*	匠	jiàng	'carpenter'
		-t-	*s-tɯɯʔ*	載	zǎi	'age'

Table 2.5 Sesqui-syllabic clusters with Crown *s- in OC (sCl structure)

Crown	Head	Tail	Reconstructed OC (*)	Character	Pinyin	Gloss
*s-	-k^{h}-	-r	*s-k^{h}ruɯb*	笈	j í	‘bamboo box for storing books’
	-p-		*s-proob*	眨	zhǎ	‘blink’
	-g-		*s-grɯm*	岑	c én	‘small but high hill’
	-q^{h}-		*s-q^{h}ra*	所	hǔ	‘the sound of cutting trees’
	-G-		*s-Grɯʔ*	俟	s ì	‘big’
	-m^{h}-		*s-m^{h}reew*	吵	chǎo	‘noisy’
	-k^{h}-	-l	*s- k^{h}lam*	僉	qiān	‘the whole’
	-k-		*s-klɯg*	稷	j ì	‘millet’
	-g-		*s-gluulʔ*	猣	xi è	‘female raccoon dog’
	-G-		*s-Gloŋ*	松	sōng	‘pine tree’
	-ɦ-	-w	*s-ɦweds*	慧	hu ì	‘intelligent’
		-wl	*s-ɦwlin*	旬	x ún	‘a period of ten days’
			s-ɦwlan	旋	xu án	‘to turn’

Table 2.6 Sesqui-syllabic clusters with Crown laryngeals (*ʔ-, *h-, *ɦ-) in OC (sC structure)

Crown	Head	Reconstructed OC (*)	Character	Pinyin	Gloss
*h-	-n	*h-n*	漢	hàn	'ripple, overflow'
	-m	*h-maaŋ*	慌	huāng	'nervous'
	-ŋ	*h-ŋeews*	嬈	ráo	'weak'
*ɦ-	-ŋ	*ɦ-ŋoŋs*	院	yuàn	'courtyard'
	-l	*ɦ-laaw*	号	háo	'howl'
	-r	*ɦ-raw*	鴞	xiāo	'owl'
*ʔ-	-l	*ʔ-leg*	益	yì	'overflow'
	-r	*ʔ-reegs*	隘	ài	'narrow'

Table 2.7 Sesqui-syllabic clusters with Crown laryngeals (*ʔ-, *h-, *ɦ-) in OC (sCl structure)

Crown	Head	Tail	Reconstructed OC (*)	Character	Pinyin	Gloss
*h-	-m-	-l	*h-mluɯg*	黑	hēi	‘black’
	-m^{h}-	-j	*h- m^{h}jew*	秒	miǎo	‘beard of corn or grain’
	-n-		*h-njuw*	手	shǒu	‘hand’
	-ŋ-		*h-ŋjeds*	勢	sh ì	‘power’
	-l^{h}-		*h- l^{h}jin*	伸	shēn	‘to reach’
	-ŋ-	-wr	*h-ŋwraals*	化	hu à	‘to change’
*ɦ-	-b-	-r	*ɦ-brim*	廩	lǐn	‘store house for rice’
	-g-		*ɦ-grig*	理	lǐ	‘to carve and polish’
	-b-	-l	*ɦ-glam*	盐	y án	‘salt’
	-g-		*ɦ-bluŋ*	孕	y ùn	‘be pregnant’
	-l-	-j	*ɦ-ljuŋ*	繩	sh éng	‘rope’
***ʔ-**	-s-	-r-	*ʔ-sru*	齺	zōu	‘uneven (teeth)’
	-l-	-j-	*ʔ-ljums*	枕	zhěn	‘pillow’
	-s-	-l-	*ʔ-sleeg*	迹	j ì	‘mark’

Table 2.8 Sesqui-syllabic clusters with Crown nasals (*m-, *N-) in OC (sC structure)

Crown	Head	Reconstructed OC (*)	Character	Pinyin	Gloss
*m-	-q	*m-qeeŋʔ*	瀴	yīng	‘far’
	-G	*m-Gaam*	姏	m án	‘elderly ladies’
*ŋ -	-g	*ŋ-gaans*	岸	àn	‘bank, shore’
*n-	-l	*n-luunʔ*	跻	ji àn	‘to trample’
	-d	*n-denʔ*	碾	niǎn	‘to grind’

Table 2.9 Sesqui-syllabic clusters with Crown nasals (*m-, *N-) in OC (sCl structure)

Crown	Head	Tail	Reconstructed OC (*)	Character	Pinyin	Gloss
*ŋ-	-g-	-r	*ŋ-greewʔ*	咬	yǎo	‘to bite’
*m-		-w	*m-gweeds*	袂	mèi	‘sleeve’
		-r	*m-graag*	貉	hé	‘raccoon dob’

Table 2.10 Sesqui-syllabic clusters with Crown *r- in OC (sC structure)

Crown	Head	Reconstructed OC (*)	Character	Pinyin	Gloss
*r-	-t	*r-tenʔ*	展	zhǎn	‘to turn’
	-d	*r-dooŋ*	撞	zhuàng	‘to strike’
	-l	*r-laa*	茶	chá	‘tea’
	-r	*r-raaŋʔ*	冷	lěng	‘cold’

Table 2.11 Sesqui-syllabic clusters with Crown stops (*t-, *k-)[22] in OC (sC structure)

Crown	Head	Reconstructed OC (*)	Character	Pinyin	Gloss
*t-	-k^{h}	*t-k^{h}aab*	鮇	qū	‘flounder’
	-k	*t-kuub*	答	dá	‘to reply’
	-n	*t-nobs*	笳	zhuì	‘a coach whip made of bamboo’
	-ŋ^{h}	*t- ŋ^{h}aans*	炭	tàn	‘charcoal’
*k-	-p^{h}	*k-p^{h}ob*	妥	fá	‘female looking’
	-m^{h}	*k-m^{h}ɯɯ*	恢	hūi	‘(ambition) great’
	-ŋ	*k-ŋoon*	冠	guān	‘hat’

[22] Sesqui-syllabic clusters with *p- are missing in sC structure.

Table 2.12 Sesqui-syllabic clusters with Crown stops (*p-, *t-, *k-) in OC (sCl structure)

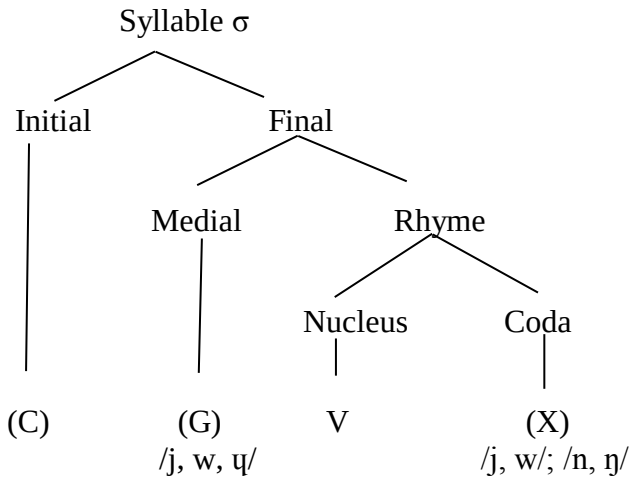
Crown	Head	Tail	Reconstructed OC (*)	Character	Pinyin	Gloss
*p-	-q^h-	-r	*p-q^hraaŋ*	烹	pēng	'to boil'
	-q-		*p-qrɯg*	皂	zào	'oak seed'
	-k-		*p-kren?*	覿	jiàn	'to see'
	-q^h-	-wr	*p-q^hwraa*	葩	pā	'flower'
*t-	-q^h-	-w	*t-q^hwaans*	唤	huàn	'to call out'
*k-	-m^h-	-l	*k-m^hlɯɯ*	詼	huī	'to tease'

In both Zhengzhang and Pan's reconstructions, six vowels (*u, *i, *a, *e, *o, *ɯ) are reconstructed in the vowel inventory (Pan, 2000; Zhengzhang, 2003). And there are no diphthongs in the two reconstructions, while the distinction between long vowels and short vowels is recognized. The six vowels all have their corresponding long forms: *uu, *ii, *aa, *ee, *oo, * ɯɯ.

2.2 Chinese syllable structure

The syllable structure of Chinese changed through its development from OC to Modern Chinese. Traditionally, the structure of Chinese is divided into the initial and the final. The final includes the medial and the rhyme. And the rhyme is further divided into two parts – the nucleus and the coda (Norman, 1988). Only single consonant is permitted in the initial. The medial is one of the three glides -j-, -w-, -ɥ-. The nucleus, the only obligatory segment in Modern Mandarin syllable, is also the tone-bearing unit. The coda can be either a glide (-j or -w) or a nasal (-n or -ŋ)[23]. The commonly adopted structure for Modern Mandarin is represented as below:

[23] Retroflex [ɻ] can sometimes be attached to the coda position when a syllable is affixed with a suffix - *er* /əɻ/ 儿 'to indicate the small size of the object named, the speakers' affection or intimacy toward it, or a euphemistic tone'. Since this suffixation is restricted to Modern Beijing Mandarin, and appears to be a fairly recent modern



(Based on the structure in van de Weijer and Zhang 2008)

Note: segments in brackets are optional.

Figure 2.1 Syllable tree structure of Modern Mandarin

The syllable structure does not change much from Middle Chinese to Modern Mandarin. In Middle Chinese, only single consonant is allowed in the initial too. Coda can be one of the three nasals *-n, -m, -ŋ, or one of the two glides *-j and -w, or one of the three plosives *-p, -t, -k, or consonant clusters like *-wŋ (one glide plus one nasal) and *-wk (one glide plus one plosive). There is much less agreement regarding the medial. According to Karlgren's reconstruction (Nie 1987), the medial in Middle Chinese includes contrasts between vocalic *-i and *-u and consonantal *-j and *-w.

In Middle Chinese and Modern Mandarin, the medial is not included in the rhyme, but is always combined with the rhyme. This constituent, which is called the final, is strongly supported by the tradition of fǎnqiē 反切. fǎnqiē (lit. "reverse-correspondence") is a traditional method to indicate the pronunciation of a character by using two other characters. Sun Yan (孫炎) is generally considered to be the first to adopt fǎnqiē in *Eryǎ Yīny ì*(爾雅音義, "Sounds and

development, I do not include it in the syllable template. For further discussion of segmental affixation in modern Chinese languages that yields syllables which do not conform to 2.1, see Lin (1989).

Meanings of Erya") during the period of Three Kingdoms (220–280 AD) (Casacchia, 2006; Zhu, 1990). In the original fǎnqiē, an unfamiliar character's pronunciation is represented by two familiar characters. The initial consonant is represented by that of the first of the two characters (上字 "upper word", as Chinese was written vertically); the final (including the medial glide, the nuclear vowel and the coda) and the tone are represented by those of the second of the two characters (下字, "lower word").

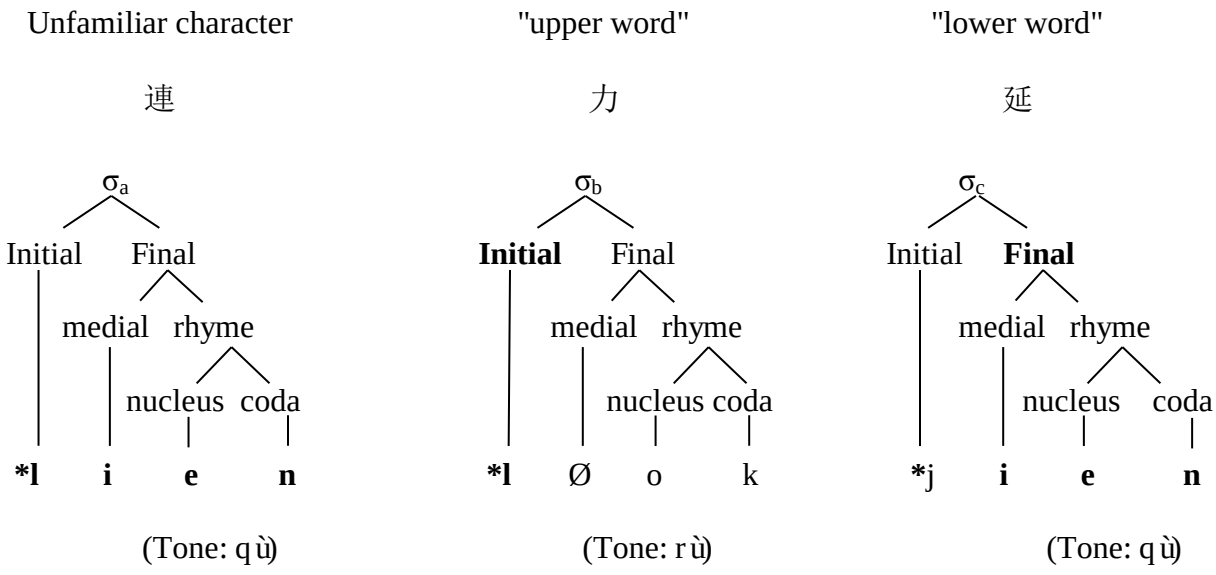


Figure 2.2 An example of fǎnqiē

Just as mentioned in section 2.1.1, *Qiēyùn* 切韻, the first Chinese rhyme dictionary using fǎnqiē was published in 601 AD during the Sui Dynasty. This tradition of combining the medial and the rhyme as the final has been preserved until now, and the syllable structure of Modern Chinese has not changed much from that of Middle Chinese.

However, when we discuss OC phonology, there will be a problem if we still adopt the concepts of the initial and the final. There were some consonant-cluster "initials" and some complex "medials" such as *-rj-, -lj-, -wj-, -wl- and -wlj- in OC phonology. The medial in OC is more closely related to the initial consonants. This feature of reconstructed OC was almost lost in

Middle Chinese. So some OC phonologists prefer grouping the medial into the "initial" instead of the "final" (Li, 1980; Zhengzhang, 2003). As argued in Sun, H. (2001), the medial in Sino-Tibetan languages mainly came from weakened Tails of the initial (H. Sun, 2001). On this basis, I hypothesize that the medial does not exist in OC. In this dissertation, I will divide the syllable structure of OC into the onset and the rhyme with the so-called medial of Middle Chinese grouped with the onset. The terms "initial" and "final" are only used to discuss Middle Chinese cases and to review the literature of Middle Chinese phonology. Based on Zhengzhang's reconstruction, the minimal and maximal syllable structures of OC are CV (e.g. *mu 牟 móu 'bellow') and CCCVCC[24] (e.g. *gwranʔ 睅 hàn 'to open one's eyes wide'), excluding the sesqui-syllabic Crown C. There are no diphthongs in nucleus position in OC and the main long vowel (e.g. *goo 侯 hóu in Pinyin 'shooting target') is considered as a single vowel in the following structure. The reconstructed OC syllable structure adopted in this study is represented in Figure 2.3:

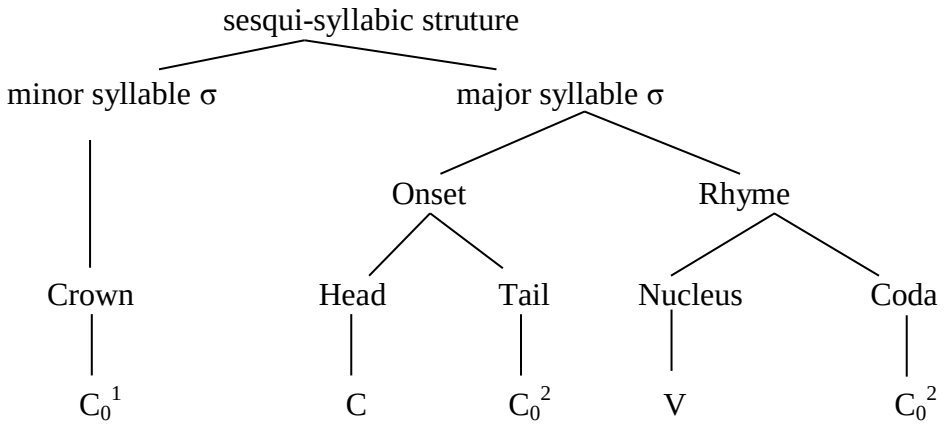


Figure 2.3 The syllable tree structure of OC

[24] Zhengzhang claimed that the Tail of onset may include three approximants */-wlj-/ at most. This could turn the maximal syllable into CCCCVCC if sesqui-syllabic Crown is not included. But in his reconstructed word list, no such examples are found.

CHAPTER 3: THE LIANMIAN WORD DATA

3.1 Preliminaries

Lianmian words 聯綿詞, refer to a category of disyllables, also named liánmiánz ì聯綿字 'connective characters' in *Fùgǔbiān* 復古編 by Zhang You 張有 of the Song dynasty (960 AD–1276 AD). As introduced in Section 1.2.2, a Lianmian word includes two syllables, which are phonologically related to each other. The two syllables of Lianmian words cannot be semantically decomposed. That is, Lianmian words are mono-morphemic disyllables, not compound words. They are roughly sub-categorized into Shuāngshēng and Diéyùn. According to Zhang You's description, the two syllables of Shuāngshēng share the same initial onset consonant. In Diéyùn, the two syllables share the same coda and nucleus with each other. Regular sound patterns of Shuāngshēng and Diéyùn that strictly conform to the description in both OC and Modern Mandarin are provided in Figure 3.1 and Figure 3.2:

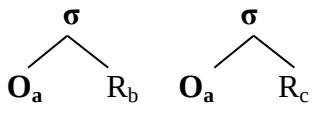
Example	**Sound pattern**
***r**eeŋ **r**ooŋ 玲瓏	
líng **l**óng 'exquisite'	

Figure 3.1 The stringent sound pattern of Shuāngshēng with an example

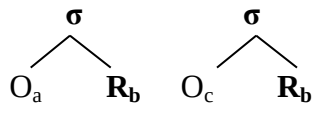
Example	**Sound pattern**
*k**ood** l^{h}**ood** 活莖	
h**uó**t**uō** 'bush or small arbor '	

Figure 3.2 The regular sound pattern of Diéyùn with an example

Lianmian words are the earliest disyllabic words recorded for Chinese, and appear to increase in number in EOC. With the rise of disyllabic compounds around LOC and EMC, Lianmian words began to lose their productivity. The identity and definition of Lianmian words thereupon became blurred. From *Maozhuanzhengjian* 毛轉鄭箋 (127 AD – 200AD) to *Hanshuzhu* 漢書註 written by Yan, Shigu (581 AD – 645 AD) and *Wujingzhengyi* 五經正義 by Kong, Yingda (574 AD – 648 AD), many scholars misunderstood Lianmian words as compounds, analyzing their two syllables into two independent morphemes. For example, the Lianmian word – *q^{h}*ljuʔ q^{h}*ljaʔ* 首鼠 shǒu shǔ 'to hesitate' was misunderstood to be a combination of two morphemes – *q^{h}*ljuʔ* 首 shǒu 'head' and *q^{h}*ljaʔ* 鼠 shǔ 'mouse'. Thus the word was misinterpreted as 'the head of a mouse turning back and forth' (Zhu, 1985). Not until the 19th century, did the Chinese scholars clarify the definition and features of Lianmian words. They brought up two principles that the two syllables or characters in Lianmian words cannot be semantically decomposed, and that there are no fixed characters for Lianmian words. The second principle explains why one Lianmian word often takes different written forms.

As has been mentioned in Footnote 6 of Chapter One, many modern scholars and linguists would like to define the phonological connections between the two syllables of Shuāngshēng and Diéyùn according to their Modern or Middle Chinese pronunciations, while ignoring their OC pronunciations. "Because of sound changes, many characters that sound different historically have become homophones (Lü, 1963, p. 440)". Thus, the slight OC sound differences between the two rhymes of a Diéyùn or the onsets of a Shuāngshēng word usually disappear in Middle and Modern Chinese. Some modern scholars then take it for granted that the two rhymes of Diéyùn should be the same and the two onsets of Shuāngshēng should be the same. But in the OC reconstructed system, this is not necessarily the case. Through data collection, this chapter

will investigate if in the improved OC reconstruction systems, the sound patterns of Shuāngshēng and Diéyùn are in conformity with their popular definitions. At the same time, the data will serve as the basis for a new analysis of the phonological origins of Lianmian words.

3.2 Data collection

For the purpose of this study, a corpus of Lianmian words in OC was compiled. It contains a total of 300 words, including 96 Shuāngshēng words (Appendix I), 172 Diéyùn words (Appendix II) and 32 Splitting-sound words (Appendix III).

The data were collected mainly from two sources: Yu, Suisheng and Guo, Li (1987) "Shuōwén jiězì de fùyīncí" 說文解字的復音詞[Disyllabic Words in *Shuōwén Jiězì*] and Sun Jingtao's dissertation "Reduplication in Old Chinese" (1999). The data of Splitting-sound words were mostly from Sun's collection. He was the first researcher who identified the Splitting-sound words from Lianmian words by investigating their OC reconstructions[25]. 166 out of 173 Diéyùn words and 90 out of 95 Shuāngshēng words were taken from "Shuowen Jiezi de fuyinci". 7 Diéyùn words and 5 Shuāngshēng words were then added to the data from Zhu, Guangqi (1985) "*Shījīng*-Shuangyin ci lungao" 詩經-雙音詞論稿 [*Shījīng:* Discussion of Disyllabic Words]. The original data were just handwritten in simplified characters without glosses or pronunciations. That is, the collectors, including Yu and Guo directly accumulated the Lianmian data from preserved old literature, such as *Shuōwén Jiězì* and *Shījīng* without reexamining their OC reconstructions. In this collection, every character in this database was looked up in the online Chinese dictionary – *Hàndiǎn* 漢典 at www.zdic.net. The simplified characters were transformed into traditional ones which show potential original semantic clues to the origins of Lianmian

[25] In Zhu (1985), Yu and Guo (1987)'s Lianmian collections, the Splitting-sound words had been classified into 'Non-Diéyùn-and-Shuāngshēng' category with other disyllabic loan words, since they did not identify the words' special phonological pattern.

words. Their glosses and Modern Chinese pronunciations in Pinyin were also included (See Appendices I, II and III). As for the reconstructed OC pronunciations, the data collection took advantage of the Eastling database[26] and selected four different OC reconstructions for comparison (including Zhengzhang, Pan, Li and Baxter's reconstructions).

3.3 Data composition

The corpus data can be classified in at least two ways: according to their parts of speech and according to their phonological patterns.

According to sound pattern, the data can be divided into 3 categories, which are: (A) Dié yùn; (B) Shuāngshēng; (C) Splitting-sound words. The breakdown of the data based on phonological pattern is displayed in Table 3.1.

Table 3.1 Classification of data by phonological pattern

	Shuāngshēng	Dié yùn	Splitting-sound words
n	96	172	32
%	32	57.33	10.67

In terms of part of speech, the data consist of three types: (A) noun; (B)verb; (C) adjective. Nouns are organized in semantic category: plants, animals and others. Nouns in animal category are further organized in size, from small to big animals: insects, birds, amphibians and big mammals.

The proportions of these three types of Lianmian words in the corpus are listed in Table 3.2. It can be seen that nouns and adjectives constitute the bulk of the data. Nouns make up a higher percentage in Splitting-sound words than in Dié yùn and Shuāngshēng. Adjectives make up a much higher percentage in Dié yùn than in Splitting-sound words.

[26] This online database (http://www.eastling.org/) is set up and sponsored by Comparative Linguistics of E-Graduate Center of Shanghai Universities.

Table 3.2 Classification of data by part of speech

	Shuāngshēng		Diéyùn		Splitting-sound words	
	N	%	n	%	n	%
Nouns	53	55.21	89	51.74	20	62.50
Verbs	9	9.37	9	5.23	4	12.50
Adjectives	34	35.42	74	43.03	8	25.00

3.4 Target structure: Sound patterns and onsets

There is always a debate over how to define the sound patterns of Diéyùn and Shuāngshēng among scholars and phonologists. Some modern scholars stick to the popular assumption that the two syllables in Shuāngshēng should have the same initial onset consonant, in terms of place of articulation and manner of articulation, and that the two syllables in Diéyùn should have the same main vowel and coda. But this strict principle rests on the basis of Modern and Middle Chinese phonology.

Wang, Li (1980) in his *Hànyǔ Shīgǎo* points out that when some Chinese scholars discuss the classification and definition of Lianmian words, the strict phonological rule for Shuāngshēng and Diéyùn are usually bent to some extent. The onset consonants of the two syllables in Shuāngshēng and the rhymes of the two syllables in Diéyùn do not have to be exactly the same. For example, the onsets *sr- and *s- can be considered similar enough to constitute the Shuāngshēng pattern; the rhymes *jij (an example from zhǐbù脂部) and *jɨj (an example from *wēibù*微部) can create a Diéyùn rhyme. Whether one uses stringent or loose phonological constraints on identity preserved Lianmian words should be analyzed in terms of OC reconstructions, and not with respect to Middle Chinese or Modern Chinese phonology, since OC is the language of their origin.

However, some OC reconstructions are established on the assumption that all rhyming words (including Diéyùn words) must share the same vowels and coda, and all Shuāngshēng words must have the same onsets. That is to say, the study of rhyme groups is considered as a primary factor of vowel reconstruction in these systems. Once the analysis of Lianmian words, especially Diéyùn words, rests on the basis of these reconstructions, circular arguments could be created.

This data collection consulted four OC reconstruction systems in consideration of difficulties and disagreements in OC reconstructions. The important point is that the reconstruction of rhymes, especially the vowels in these systems is not established on the basis of rhyme groups. The data are categorized and pre-analyzed according to the following two criteria:

a) The reconstructed OC pronunciation rather than Modern Chinese pronunciation is the criterion for Diéyùn and Shuāngshēng patterns.
b) A word can be identified as a Diéyùn or Shuāngshēng word, as long as one of its four reconstructions displays Diéyùn or Shuāngshēng pattern.

Although the modern Chinese sound of 嵯峨 cuó é does not conform to Diéyùn pattern, *zlaal ŋaal 嵯峨 cuó é 'high and stiff' could be included in Diéyùn data according to criterion a), because in OC reconstructions the two syllables of the word share the same nucleus and coda *aal. To explain the criterion b), Shuāngshēng word 棠棣 'Chinese bush cherry' is cited as an example. It has four different reconstructions: **daaŋ l'ɯɯds* (Zhengzhang) **g ɬaaŋ b ɬɯɯds* (Pan) **daŋ diədh* (Li) **daŋ ləts* (Baxter). Li's reconstruction justifies its Shuāngshēng identity.

Based on the above two criteria, this study makes several adjustments to the original data: four Shuāngshēng words from the original data (Yu & Guo, 1987) are eliminated from the current Shuāngshēng collection which are listed in Table 3.3; and three Diéyùn words are

eliminated from the current Di éyùn collection which are listed in Table 3.4. These are mono-morphemic disyllables, classified as Lianmian words in *Shuōwén*. However, they do not display Lianmian sound patterns according to the updated reconstructions.

Table 3.3 Four eliminated words from Shuāngshēng data

Character	OC reconstructions (*)				Pinyin	Gloss
	Zhengzhang	Pan	Li	Baxter		
蚣蝑	klooŋ sŋas	klooŋ sŋas	kuŋ sjiagh	koŋ sŋja	gōng xū	‘centipede’
閶闔	t^hjaŋ gab	t^hjaŋ gaab	t^hjaŋ gap	t^hjaŋ gap	chāng hé	‘the gate of Heaven’
蚿蚗	meenʔ kweed	meenʔ kweed	meenʔ kwet	meenʔ k^wiat	mi án ju é	‘a kind of cicada’
踧踖	l'ɯɯwg s^hag	g-liiwg sk^hag	diəkw ts^hjak	diwk ts^hjak	c ù j í	‘with mincing steps to show respect ’

Table 3.4 Three eliminated words from Di éyùn data

Character	OC reconstructions (*)				Pinyin	Gloss
	Zhengzhang	Pan	Li	Baxter		
咄欭	ʔl'ood q^hlɯd	k-lood q^hlɯd	tot hjət	tuat xjət	du ó ch ù	‘to reproach’
螭魅	r^hel mrɯds	p^h-rel mrɯds	hljar mjiədh	hrjej mrjəts	l ím èi	‘monster’
委虒	qrol sle	qol sle	ʔjiar stjig	ʔrjoj sje	wěi sì	‘legendary horned tiger’

In my Lianmian word data, there are many words that do not fit the stringent phonological patterns of Lianmian words (Shuāngshēng: $\mathbf{O_a}R_b\mathbf{O_a}R_c$ and Di éyùn: $O_a\mathbf{R_b}O_c\mathbf{R_b}$). In order to display to what extent their patterns could be bent, the ‘exceptions’ of Shuāngshēng and Diéyùn will be listed and analyzed respectively.

Table 3.5 Exceptions to the stringent sound pattern ($O_aR_bO_aR_c$) of Shuāngshēng

	Reconstructed OC (*)	Characters	Pinyin	Gloss
1.	**kr**aa **kw**ag	貑玃	jiā jué	'a large female ape'
2.	**k^h**id **k^{h}l**ud	蛣蝍	jié qū	'scorpion'
3.	**s^{h}l**uum **s^{h}r**aal	參差	cēn cī	'uneven, irregular'
4.	**kr**ee **k**ooʔ	薤荂	xiè hòu	'water chestnut'
5.	**p**iʔ **br**aa	枇杷	p íp á	'loquat'
6.	**s ɬ**al **l**o	虓蝓	sī yú	'snail'
7.	**mr**eeg **m**oog	霢霂	m ài m ù	'light rain'
8.	**g**aans **gr**uug	韃鶱	h àn xu é	'titmouse'
9.	**kr**aa **k^w**ag	貑玃	jiā jué	'a large female ape'
10.	**k**ɯʔ **kr**aa	麝貑	jiù jiā	'a kind of deer'
11.	**l^h**uun **n^h**aan	涒滩	tūn tān	'a way of numbering the years in the lunar calendar'
12.	**s**id **sr**uds	悉蟀	xī shuài	‘crickets’
13.	**g**rees **g**oost	邂逅	xi èh òu	'to meet without appoinment'
14.	**k^{h}l**ud **k^h**ɯɯds	忼慨	kāng kǎi	'emotionally aroused'
15.	**g**a **k^h**wed	渠蠽	q ú qu è	'a kind of bug'
16.	**ts**jəgw **dz**id	蝤蠐	qi ú q í	'lonicorn's larva'
17.	**d**ɯw **t**as	薵箸	ch óu zh ù	'hesitant'
18.	**p^h**aaʔ **b**u	溥浮	pǔ fú	'black finless porpoise'
19.	**d**i **t**a	荎藸	chí zhū	'the fruit of Chinese magnoliavine'
20.	**k**iig **ŋ**ed	鶟鱍	jié jiā	'wild ducks'
21.	**q^h**eem **g**o	黚姁	xiān xǔ	'reddish yellow'
22.	**k**ralʔ **q**oo	踦區	qí qū	‘uneven, rough’

The exceptions in Table 3.5 show that the stringent sound pattern of Shuāngshēng could be extended in three aspects.

Exceptions 1-14 show ONSET EQUIVALENCE 1:

Onset equivalence 1: Head C identity

Exceptions 15-19 show ONSET EQUIVALENCE 2:

Onset equivalence 2: Cs are equivalent despite differences in laryngeal features (voicing, aspiration).

Exceptions 20-22 show ONSET EQUIVALENCE 3:

Onset equivalence 3: Cs are equivalent despite differences in dorsal place (velar/uvular) and manner.

ONSET EQUIVALENCE 1, reflected in 1-14 exceptions, extends the pattern greatly: as long as the two syllables of a Lianmian word share the same Head consonant in onset, the word is identified as Shuāngshēng. In the case of *__kr__aa __kw__ag 貑玃 jiā jué ‘a large female ape’, the two syllables’ onsets differ in Tail (*__kr__aa has a *r liquid in the Tail of onset, while *__kw__ag has a w glide), but they have the same consonant *k in the Head of onset. The two syllables of *s ɬal lo 虓蝓 ‘snail’ have the same Head of onset *l, but the two onsets differ in the Crown position (the first syllable includes a *s- Crown, while the second does not). The distinctions of the Crown and Tail consonants in the two cases can be eliminated, since the Head consonant is put in priority in the identity of onset equivalence.

ONSET EQUIVALENCE 2 shows that when the two Head consonant of a Lianmian word differ only in the voicing and aspiration features, the word is still identified as Shuāngshēng. This extension of Shuāngshēng pattern could be revealed in the bold onset consonants of 15-19 exceptions.

Exceptions from 20-22 show us that the distinctions of the dorsal Head consonants in place and manner of articulation could be eliminated. The dorsal Head consonants include *q^h,*k, *g and *ŋ.

Table 3.6 Exceptions to the stringent sound pattern ($O_aR_bO_cR_b$) of Dì éyùn

	Reconstructed OC (*)	Character	Pinyin	Gloss
23.	qlɯɯ nɯɯ**s**	埃皚	**á**i n**à**i	‘foggy and dark’
24.	ruul**ʔ** tuul	磊竣	lěi duì	‘description of piling up’
25.	ql**oo**ŋ ʔ sl**o**ŋ	蝓蜙	wēng zōng	‘a kind of gad-fly’
26.	t**a**ŋ gw**aa**ŋ	粻餭	zhāng huáng	‘provisions’
27.	t**o** m**u**	竈蟊	zhū máo	‘a kind of fly’
28.	gr**uu**ŋ sr**oo**ŋ	桻雙	xiáng shuāng	‘a sail made of thin bamboo stripes’
29.	t**oo** m**u**	兜鍪	dōu móu	‘helmet used in the battle’
30.	ʔ sr**u** ŋ**oo**	齺齵	zōu yú	‘uneven (teeth)’
31.	t**ɯɯ**ws ʔ ɯmr**ee**w	窱窅	dāo yáo	‘deep, distant’
32.	ʔ ɯ**ee**wʔ ʔ ɬ'**ɯɯ**wʔ	杳窱	yǎo tiǎo	‘deep, dark, profound’

The exceptions in Table 3.6 show that the stringent sound pattern of Dì éyùn could be extended in three aspects too.

Exceptions 23-24 show RHYME EQUIVALENCE 1:

Rhyme equivalence where the final Coda consonants are equivalent

Exceptions 25-26 show RHYME EQUIVALENCE 2:

Rhyme equivalence where V is equivalent to V:

Exceptions 27-32 show RHYME EQUIVALENCE 3:

Rhyme equivalence where back/round vowels and front/central unround Vs are equivalent

RHYME EQUIVALENCE 1, reflected in exceptions 23-24, tells us that when two rhymes of a mono-morphemic word differ only in the final coda, the word is still grouped in Diéyùn. As mentioned in Chapter 2, 11 single consonants and 10 consonant clusters may appear in coda of an OC syllable. If one of the two rhymes of a Diéyùn word has a complex coda and the other has a single-consonant coda, the final coda consonant of the two rhymes may differ (e.g.*ruulʔ tuul* 磊諄 lěi duì 'description of piling up'). If one of the two rhymes of a word has a single-consonant coda and the other has no coda at all, the word is also identified as a Diéyùn word as long as the two syllables have the same vowel (e.g. **qlɯɯ nɯɯ**s 埃瞖 ái nài 'foggy and dark').

RHYME EQUIVALENCE 2 shows that the length of the vowels of the two rhymes could be different in Diéyùn pattern. In some Cantonese folk songs and operas, the syllables with the same vowel but with different vowel length are also classified into the same rhyme group. A similar rhyme pattern allowing different vowel length could also be found in the ballads of Yao, an unrelated Hmong-Mien language (Zhengzhang, 2003, p. 186-187; Purnel, 1988).

RHYME EQUIVALENCE 3, reflected in exceptions 27-32, extends the pattern of Diéyùn to a loose-fitting one. The vowel distinctions between /u/ and /o/, /ɯ/ and /e/ are ignored in the rhymes of Diéyùn.

We could contribute these 'exceptions' to the limitations of OC reconstructions. But the more reasonable conclusion we could come to from the data is that the stringent sound pattern which the widely-accepted definition of Shuāngshēng and Diéyùn follows should perhaps be reconsidered. The result of this data collection proves wrong the popular notion that Diéyùn adheres to a strict pattern of rhyme identity, while Shuāngshēng adheres to a strict pattern of onset identity. Instead, when the two onsets of Shuāngshēng contain complex consonants, it is allowed in the loose-fitting pattern that only the Head consonants are the same. Rhyme group in

Diéyùn is not equal to the conception of rhyme vowel, although in Chinese both words have the character -韻 yùn 'rhyme'. The two rhymes of a Diéyùn do not have to be exactly the same.

The other phonological structure to be analyzed in this research is the onsets of each syllable in Diéyùn and Splitting-sound words. The analysis mainly refers to Zhengzhang's reconstruction, unless his reconstruction is at odd with the Diéyùn pattern or the Splitting-sound pattern. Diéyùn and Splitting-sound words will be further categorized according to the types of their syllable onsets. The types will be discussed in detail in Chapters 4 and 5. And the results of the analysis will be adopted to support the hypothesis that the simplification of complex onsets in OC was a central motivating factor for the rise of the earliest disyllabic forms – Splitting-sound words and Diéyùn words.

CHAPTER 4 THE RISE OF SPLITTING-SOUND WORDS

4.1 Introduction

The study of disyllabicity in the history of Chinese must start from the exploration of Lianmian words, the first Old Chinese disyllables. Among Lianmian words, focus should be on Splitting-sound words which display some noteworthy and interesting phonological features. In this chapter, the phonological motivation for the rise of Splitting-sound words is examined on the basis of the analysis of their sound patterns. I hypothesize that Splitting-sound words are the earliest disyllabic words in Early Old Chinese, and detail their evolution from earlier monosyllables.

Traditional scholars tend to compare Splitting-sound words with Diéyùn and Shuāngshēng words, overlooking the special sound pattern of Splitting-sound words. They would like to categorize those words into a non-Diéyùn-or-Shuāngshēng group. Recently, many researchers have started to examine the sound pattern of Splitting-sound words more carefully (Zhao, 1979; Li, 1991; Sun, 1999). After analyzing the Splitting-sound words preserved in Ancient Chinese literature and those in existing Chinese dialects, they conclude that Splitting-sound words are the products of the splitting process of one-syllable words into two syllables. This identity-preserving process has guided some phonologists into reduplication theory, considering Splitting-sound words to be products of reduplication, an interface between phonology and morphology (Sun, 1999). However, reduplication theory cannot explain why this 'reduplication' process soon lost its productivity as a word formation tool around EMC.

In order to solve the mystery of how Splitting-sound words arose in EOC and why they began to lose productivity around EMC, this chapter is going to analyze the reconstructed phonological system and to reveal the phonological motivation for their rise. The study of the

origin of Splitting-sound words will contribute to the study of disyllabicity, though most linguists exclude Splitting-sound words from the database of disyllables they are studying.

This chapter is organized as follows. In 4.2, the concept of Splitting-sound words is introduced, and the Splitting-sound word data are pre-analyzed; in Section 4.3, previous research is reviewed, including Sun's Reduplication theory, Serruys' (1959) diachronic explanation and Pan's (date) historical phonological hypothesis. My hypothesis is illuminated in Section 4.4; Section 4.5 analyzes the language-internal evidence, and in Section 4.6 evidence from Mon-Khmer languages is discussed. Section 4.7 sums up the account for the rise of Splitting-sound words.

4.2 What are Splitting-sound words?

A Splitting-sound word is also called a cutting-foot word 切腳語 by Hong Mai 洪邁 (1123AD - 1202AD). Starting from the 20th century, these kind of words have been called 'inserting-l' words 嵌 *l* 詞 (Zhao, 1979). Sun calls the process fission reduplication when discussing reduplication in OC in his dissertation (Sun, 1999). Serruys gave it another name, calling it "dimidiation binom" in his analysis of the dialects of the Han time (Serruys, 1959). In this study, I would like to adopt the popular name 'Splitting-sound word' which is a word-to-word translation of the Chinese term 分音詞, referring to a disyllable that is split from a one-syllable word: the onset of the monosyllabic word is reflected in the first syllable onset of the Splitting-sound word, and the rhyme of the monosyllabic word is preserved in the second syllable onset. Hong Mai 洪邁 is the first person who brought forward the concept of cutting-foot word and described the sound pattern of his seventeen examples in his *Rǒngzhāi Suǐbǐ* 容齋隨筆. Some examples selected from his data are shown in (1):

(1) Hong's examples of Splitting-sound words in Middle Chinese[27] (see Appendix IV):

	Cutting-foot word (Splitting-sound forms)	Monosyllabic forms	Gloss
a.	*kut lua 骨露 gū lù	*kua 锢 gù	'run metal into cracks'
b.	*k^{h}yt lɥan 屈孿 qū luán	*k^{h}ɥan 圈 quān	'circle'
c.	*tɬut lwan 突欒 tū luán	*tɬuan 團 tuán	'knot'
d.	*pɬua laŋ 步廊 bùláng	*pɬaŋ 旁 páng	'vast, boundless'
e.	*tɬut lak 突落 tū luò	*tɬak 鐸 duó	'kind of bell'

According to Hong, the disyllabic forms above all result from division of the monosyllabic forms, with which the disyllables in each case share the same meaning (Sun, 1999).

This study collected 32 Splitting-sound words in OC (see Appendix III), among which only 6 words have their monosyllabic forms:

[27] The reconstructed Middle Chinese forms are 鐸 adapted from Sun, Jingtao's data (7) (1999, p. 133). Five of Hong's seventeen examples are not included in Appendix IV, because these five disyllables are disyllabic phrases which are phonetically split from five monosyllables, such as 不可 'not possible, not permissible' (**puu khaal? > *pyot k^{h}a*) from 叵 **pk^{h}aal?* (Sun, 1999, p. 133).

(2) Examples of Splitting-sound words in OC:

	Monosyllabic forms	Gloss	Source	Splitting-sound forms	Glosses	Source
a.	*ziil?/zliil? 薺 qí	'tribulusterre-stris'	*Shuōwén*	*zid riil 蒺藜 jílí	'tribuluster-restris'	*Yìjīng*
b.	*q^{h}at 歇 xiē	'stop'	*Zuǒzhuàn*	*q^{h}al lats 戲泄 xìxiè	'to rest'	*Fāngyán*
c.	*pew 猋 biāo	'whirlwind'	*Shījīng*	*pa lews 扶搖 fúyáo	'whirlwind'	*Zhuāngzǐ*
d.	*doo 頭 tóu	'head'	*Lǐjì*	*doog roo 髑髏 dúlóu	'skull'	*Zhuāngzǐ*
e.	*kog 挶 jú	'to work with crooked arms'	*Shuōwén*	*koo roog 鞠録 qúlù	'description of working with crooked arms'	*Xúnzǐ*
f.	*ɢwaag 鑊[28] huò	'large cooking pot, cauldron'	*Oracle-bone inscriptions*	*ɢwaa raag 瓠落 hùluò	'spacious'	*Zhuāngzǐ*

In the data, the monosyllabic forms of 24 Splitting-sound words cannot be identified or traced back. After examinations of the sound patterns, the Splitting-sound process is illustrated in Figure 4.1:

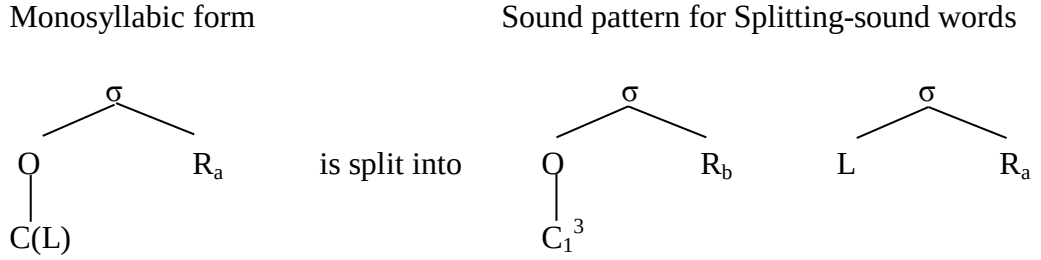


Figure 4.1 Sound-splitting process

[28] It is still controversial that **ɢwaag* 鑊 should be considered as the monosyllabic form of **ɢwaa raag* 瓠落. This point will be discussed in detail in 4.3.

There are sound variants of the onsets of monosyllabic form and of the first syllable of the Splitting-sound form: it could be a single consonant or a CC consonant cluster (e.g. (2a) and (2f)). There are two cases in the data where the first syllable onset of Splitting-sound form can consist of three consonants (e.g. *q^h*wraa reels* 華離 hu á lí ‘crooked’).

Unlike Shuāngshēng and Di éy ùn words, Splitting-sound words are still used in colloquial speech in some dialects of Modern Chinese languages. A few examples from some dialects of Modern Jin [cjy] and North Min [mnp] are given in (3) to (7). Data in (3) – (6) present the examples from dialects of Jin [cjy]; examples in (7) are from North Min [mnp].

(3) Examples of Splitting-sound words in Shunping dialect[29] (see Appendix V):

	Splitting-sound words	Monosyllabic words	Gloss
a.	tɕi$_{55}$liŋ$_0$	tɕiŋ$_{55}$ 惊	‘to be startled’
b.	k^{h}ə$_{21}$4ləw$_0$	k^{h}əw$_{214}$ 口	‘mouth of something’
c.	t^{h}u$_{55}$lwo$_0$	t^{h}wo$_{55}$ 脱	‘to hang loose’

(Based on Sun, 1999)

(4) Examples of Splitting-sound words in Yimeng dialect:

	Splitting-sound words	Monosyllabic words	Gloss
a.	t^{h}wəʔ$_{21}$luŋ$_{44}$	t^{h}uŋ$_{44}$ 臀	‘hip’
b.	p^{h}əʔ$_{21}$ lan$_{44}$	p^{h}an$_{44}$ 盘	‘dish’

(Based on Hou 1989)

(5) Examples of Splitting-sound words in Pingyao dialect:

	Splitting-sound words	Monosyllabic words	Gloss
a.	pʌʔ$_{54}$læ$_{53}$	pæ$_{53}$ 摆	‘to swing’
b.	kʌʔ$_{54}$ləŋ$_{53}$	ləŋ$_{53}$ 埂	‘slopes between terraced fields’

(Based on Li, 1991)

(6) Examples of Splitting-sound words in Taiyuan dialect:

[29] The subscript number following each syllable indicates the tone of that syllable.

	Splitting-sound words	Monosyllabic words	Gloss
a.	$kwaʔ_{54}la_{2}$	$kwaʔ_{54}$ 刮	‘to scrape’
b.	$məʔ_{54}ləŋ_{11}$	$məŋ_{11}$ 蒙	‘to cover’
c.	$kwəʔ_{54}lə_{53}$	$kwɤ_{53}$ 裹	‘to wrap’

(Based on Zhao, 1979)

(7) Examples of Splitting-sound words in Fuzhou dialect:

	Splitting-sound words	Monosyllabic words	Gloss
a.	$k^{h}je_{11}ljɛu_{213}$	$k^{h}jɛu_{213}$ 翹	‘to be placed upside down’
b.	$so_{11}lɔʔ_{23}$	$sɔʔ_{23}$ □[30]	‘to bind round tightly’
c.	$a_{31}lau_{44}$	au_{44} 凹	‘concave, dented’

(Based on Li, 1982; Zheng, 1983)

Comparing the data from (1) to (7), we may conclude that Splitting-sound words in OC display some features that differ from those in Middle Chinese and Modern Chinese dialects. First, in OC Splitting-sound words, both the lateral liquid *l- and the dental liquid *r- appear in the onset position of the second syllable, unlike the modern and MC cases, where only *l- is used (Sun, 1999). Second, in most of OC Splitting-sound words, the first rhyme differs from the second one only in coda (except 2c), while the analogy in the modern and Middle Chinese cases is that rhyme patterns vary in different dialects.

These features suggest a distinctive phonological change from OC to Middle and Modern Chinese. According to Jin (2002) and Zhengzhang (2003), the Head *r- in Old Chinese was changed into *l- in Middle Chinese. The coda distinction of the two rhymes of a Splitting-sound word in OC that is not evident in Middle and Modern Chinese can be explained by the historical origin of tones. Before the rise of tones, some codas played a morphological role. For example, the *-s coda, as a derivational suffix, has two morphological functions: deriving a noun from a

[30] ‘□’ here means there are no characters available for this meaning in formal language.

verb and deriving an export-oriented verb from an import-oriented verb (Mei, 1980; Sagart, 1999). Examples given by Mei are shown below (Zhengzhang's reconstruction):

(8) Examples of the morphological functions of *-s

a. Deriving a noun from a verb:

*nuubs 内 nèi 'the inside'　　*njubs 入 rù 'to enter'

*reds　例 lì 'patten, class'　　*red　列 liè 'to break up'

b. Deriving an export-oriented verb from an import-oriented verb:

*mɯns　問 wèn 'to ask'　　*mɯn　聞 wén 'to hear'

*l^{h}ɯɯgs 貸 dài 'to lend'　　*l^{h}ɯɯg 貣 tè/dài 'to borrow'

Tone arose between 500 BC and 500 AD as a result of the simplification of coda clusters in OC (Sagart, 1999). Thereupon the morphological functions that these coda endings or suffixes carried with them faded away.

Table 4.1 Correspondence between coda consonants in OC and four tones in Middle Chinese

	píng	shǎng	qù	rù
Post-coda	ø	-ʔ	-s	
Nasal coda	-m, -n, -ŋ	-mʔ, -nʔ, -ŋʔ	-ms, -ns, -ŋs	
Plosive coda			-bs, -ds, -gs	-b, -d, -g
Phonetic description of tones[31]	33	35	31	3

(Adapted from Zhengzhang, 1987)

Unlike the Splitting-sound words in OC, the written forms of the Splitting-sound words in Modern Chinese dialects are missing. This strongly suggests that Splitting-sound words in the dialects are only relics of older forms, preserved and used in colloquial speech.

[31] Chao (1930) introduced a five-level system for the phonetic description of tones. On this scale, 5 represents the highest tone level, and 1, the lowest, and each tone is represented by a starting pitch and a final pitch, and sometimes a mid pitch. Zhengzhang adopted Chao's digital system in his reconstruction of the tones in Middle Chinese.

The interesting feature that all the data in OC, Middle Chinese and Modern Chinese have is the Liquid in the second syllable onset. The choice of Liquids appears to have nothing to do with the monosyllabic form. Then two questions arise naturally relative to the occurrence of L: Why do only liquids appear in this position? What is the relation between the occurrence of liquids and the origin of Splitting-sound words? Chinese linguists have attempted to explain this point from different angles. In the next section, these different views will be introduced and discussed in detail.

4.3 Previous studies of Splitting-sound words

In Sun's dissertation (1999), he discusses the production of Lianmian words including Splitting-sound words as a reduplication process --- an interface between morphology and phonology. Splitting-sound words are interpreted as a product of a special reduplication: fission reduplication, during which the base syllable is split into two pieces corresponding to the onset and rhyme of the base (Sun, 1999, p. 123). From a morphological point of view, the semantic derivation from a monosyllabic base to a disyllabic fission reduplication form can be characterized as SPECIALIZATION (Sun, 1999, p. 139). He further summarizes the phonological process into two principles and three steps.

One Syllable/One Meaning Principle (OOP):

Chinese morphemes are monosyllabic in nature; that is to say, one meaning should be denoted by just one syllable and one syllable should always be a bearer of meaning.

Sonority Sequencing Principle (SSP):

Sonority rises through the onset, reaches a peak at the syllable peak, and falls through the coda.

The two principles are applied to the reduplication process in three steps: Total copying of the base syllable first results in two identical syllables. Under the constraints of OOP and SSP, the onset of the second syllable is replaced with the Liquid *l-, whose sonority is higher than other [+cons.] segments. Sun thinks that the occurrence of *l- prevents the sonority of the second syllable onset from falling down, creating the impression that the two syllables sound like one syllable. The third step involves the modification of the first syllable coda. Under the constraint that the first rhyme must be minimally modified, the first rhyme is modified in accordance with the closeness of Yīn, Yáng and Rù rhyme categories. If no suitable forms exist in the corresponding OC rhyme inventory, the rhyme *aŋ will be surfaced as the default form.

Sun's fission reduplication view seems reasonable at first sight, but after further investigation, several issues are found in this explanation that need addressing. The first issue is that if we interpret Splitting-sound words as a kind of reduplication, the uniqueness of its existence in Chinese becomes uninterpretable. Compared with other reduplication in Lianmian words, e.g. total reduplication, Shuāngshēng (double-onset disyllables), and Diéyùn (double-rhyme disyllables), the sound pattern of Splitting-sound words displayed in reduplication is rarely found in other languages. Total reduplication can be found in many languages; Shuāngshēng and Diéyùn can also be noted in some neighboring languages. For example, in Malay, Khmer and Vietnamese, parallel processes are termed Chiming and Rhyming. In rhyming reduplication of Malay, one of the base syllables, either the initial syllable (together with the following consonant) or the final syllable is copied onto the reduplicant. In Chiming, only the consonants are repeated, while the vowels undergo phonetic modifications (Zaharani, 2005, p. 14).

(9) Chiming examples from Malay [msa]:

gopoh ‘hasty’ gopoh-gapah ‘to do things hastily’

tanãh ‘soil’ tanãh-taneh ‘various kinds of soils’

The following data is from Khmer, showing Rhyming and Chiming reduplication respectively (Jacob, 1979):

(10) Rhyming and Chiming examples from Khmer [khm]:

A. Rhyming (with reduplication of vowel nucleus and final consonant):

/mɛːn/ ‘true’ /mɛːn – tɛːn/ ‘really’

/sruːəl/ ‘comfortable’ /sruːəl – buːəl/ ‘with ease’

B. Chiming[32] (with reduplication of initial consonants)

/praə/ ‘to use’ /praə - prah/ ‘to use generally’

/liːəy/ ‘mixed’ /liːəy - lɔm/ ‘jumbled up’

Rhyming and Chiming reduplications (Di éyùn and Shuāngshēng) are also found abundantly in Vietnamese.

(11) Chiming and Rhyming examples from Vietnamese [vie] (Nh àn, Ng ô Thanh, 1984):

A. Chiming

/đau/ ‘hurt’ / đau điếng / ‘hurt horribly’

/ mạnh/ ‘strong’ /mạnh mẽ / ‘very strong’

B. Rhyming

/loảng xoảng/ ‘blunt sound of glasses breaking into pieces’

/lục đục/ ‘the sound of hard, blunt objects hitting against each other’

[32] Jacob coined this type as ‘Alliterative’ to be distinctive from another category with reduplication of initial and final consonant.

However, the sound pattern of Splitting-sound words and the splitting process are rarely found in other languages, even in neighboring languages. Therefore, we have the reason to throw doubt on the Fission reduplication interpretation of the formation of Splitting-sound words because of its uniqueness.

The second issue is that we cannot interpret why the base forms of abundant Splitting-sound words in OC are missing. Actually in Sun's data, there are only 8 examples that have base forms, while 19 cases do not have corresponding mono-syllabic forms. Usually the base form of the reduplicant should be easily identified and preserved, because Sun argues that the original monosyllable (base) and the disyllable (reduplicant) are different with respect to their semantic significance. If the reduplicant had the semantic function of specialization, according to Sun's analysis, the base should have been preserved as it carries the original meaning. However, the case is that most of the 'base' forms of the sound-splitting words cannot be traced back while the disyllables are preserved in OC.

The third issue concerns Sun's analysis of semantic functions of Fission reduplication. Although the meanings of reduplication differ greatly, 'as pointed out by a number of linguists, the particular meanings associated with reduplication strikingly reoccur across languages' (Moravcsik, 1978, p. 316). There are universally four common semantic properties of reduplication: increase of quantity (Augmentation), decrease of quantity (Diminution), increase of degree (Intensification), and decrease of degree (Attenuation). Sun attempted to analyze his Splitting-sound word data in terms of the four properties. Nonetheless, his analysis did not come to a good result to support the theory. In half of his 8 cases, the meanings expressed by 'reduplication' are exactly the same as those of their bases (see details in (2a, c, d, e)). That is to say, 'reduplication' fails to serve any semantic function in these cases. Sun adopted 2f as

evidence to support his view that a monosyllabic base undergoes fission into two syllables, with a simultaneous specialization in meaning taking place. However, there exist several problems with this example, which further undermine his argument. Based on its phonological pattern and semantic properties, he assumed **Gwaa raag* 瓠落 hùluò as a fission reduplication word. Then, with the aid of its sound pattern, he reconstructed a sound *Gwaag, the corresponding character of which he found was **Gwaag* 鑊 huò. Since he found the monosyllable **Gwaag* 鑊 'big pot' and the disyllable **Gwaa raag* 瓠落 hùluò 'spacious' have some connections in their meaning, Sun made a bold assumption that **Gwaag* 鑊 huò was the base form of **Gwaa raag* 瓠落 hùluò. Eventually a circular argument was created, when he adopted the loose semantic connection between the so-called 'base form' and the 'fission reduplication word' as support for his reduplication theory that the fission process from a monosyllable into a disyllable is a morphological process rather than a phonological one. Unlike the other splitting-sound words, the first syllable of **Gwaa raag* 瓠落 carries a meaning and could be used as a free morpheme or a monosyllabic word by itself in OC.

In order to investigate the origin of **Gwaa raag* 瓠落 and whether its first syllable - **Gwaa* 瓠 is a free or bound morpheme, let us reexamine the story quoted by Sun from *Zhuāngzǐ ·Xiāoyáoyóu* (Sun, 1999, p. 151):

> 魏王貽我大瓠之種，我樹之成而實五石，以盛水漿，其堅不能自舉也。剖之以為瓢，則瓠落無所容 'The king of Wei gave me the seed of big bottle gourd. I planted it until it ripened into a bottle gourd with a volume of five *shi*. If used to hold water or sour drink, it is not firm enough to bear such weight; if it is cut into two gourd ladles, then, either is so

huge that there is nothing proper which could be contained in it'.

The character 瓠 corresponding to the first syllable of **Gwaa raag* 瓠落 also appears in the first line of the story where it means 'bottle gourd'. Although **Gwaa raag* 瓠落 hùluò could also be written as 護落 hùluò with the same pronunciation when it means 'being frustrated, coming down in the world', it must originate from the noun 瓠'bottle gourd' rather than **Gwaag* 鑊 'big pots', because **Gwaa raag* 瓠落 is often found in the poems about bottle gourds. For example, in Liu, Congyi's (1179 AD - 1222 AD) poem – *Ridicule of the Third Brother's Planting Gourds and his poem*《三弟手植瓢材且有詩予亦戲作》:

早知瓠落終無用，只合江湖養不才 If I had known my capacity being too huge to be used just as the bottle gourd, I would have lived in seclusion far from the royal court.

I argue that no matter whether the word **Gwaa raag* 瓠落 hùluò should be considered as a Lianmian word or not (because the first syllable could be used as a free morpheme), its sound pattern evolves from the monosyllable **Gwaa* 瓠 hù 'bottle gourd' rather than **Gwaag* 鑊 huò 'big pot'. In Zhengzhang's reconstruction, 瓠 hù 'bottle gourd' is restored as **gwlaa* (*gwaa in Pan's reconstruction, **gwag* in Li's and *wa in Baxter's). I assume that the disyllable **Gwaa raag* 瓠落 hùluò is a product of the Splitting process from the monosyllable **Gwaag*. The analysis of this case proves wrong Sun's claim that **Gwaa raag* 瓠落 hùluò 'spacious' is the product of fission reduplication of **Gwaag* 鑊 huò 'big pot'. And all this proves conclusively that the semantic method on which Sun's fission reduplication theory is built cannot be relied on to identify the monosyllabic form of Splitting-sound words. Also, the 'gourd' case makes it more

convincing that Splitting-sound words should be accounted for as a product of splitting process rather than that of reduplication.

To sum up, Sun is the first one who classifies Splitting-sound words as a type of reduplication and presents an insightful analysis of their sound patterns. However, his reduplication view of Splitting-sound words gives rise to several problems that are hard to resolve: synchronically, Fission reduplication cannot explain the uniqueness of Splitting-sound pattern and process; diachronically, Fission reduplication cannot explain why the base forms in OC are missing, and why in EMC this reduplication lost its productivity.

Serruys (1959) provides a diachronic explanation for the creation of the Splitting-sound words (or "dimidiation binoms" in Serruys' terms). He reveals that in the sound simplification process, the loss of some of the clusters caused these words to be less easily identified or separated from others, thus the monosyllabic words were enlarged by means of various devices, including reduplication, compounding and this Splitting-sound device to prevent the ensuing homophony. Although mentioning that the phonological development in OC affected the formation of the Splitting-sound words, Serruys thinks of this phonological factor as an indirect motivation. The primary cause was functional -- homophony avoidance. Also he fails to analyze in detail how the phonological simplification process gave rise to Splitting-sound words.

With the development of OC reconstruction, more and more historical phonologists have supported the idea of the existence of consonant clusters in OC (Baxter, 1992; Li, 1980; Mei, 1989; Zhengzhang, 2003). And more and more scholars have begun to connect the creation of Splitting-sound words with the development of OC phonology. Among them, Pan, Wuyun (1999) is the first one who attempted to clarify this point with proofs. However, he assumes that the Splitting-sound words developed from sesqui-syllables and sesqui-syllabic clusters (C.C) rather

than consonant clusters of regular syllables in OC (CC). He argues that sesqui-syllables have three developmental directions: Splitting-sound words, reduced sesqui-syllables and mono-syllables. He cited Modern Cantonese [yue] examples from Chen (1984) to support his hypothesis:

(12) Cantonese examples with sesqui-syllables (adapted from Chen, 1984)

Character	Monosyllables	Splitting-sound words	Reduced form	OC (Zhengzhang system)	Gloss
角	kɔ:k	kɔ:k lɔ:k	kᵊlɔ:k	*kroog	'corner'
結	ki:t	ki:t li:t	kʰᵊli:t	*kiid	'knot'
筆	pɐt	pɐt lɐt	pᵊlɐt	*prud	'pen'

Pan argues that all three forms in Modern Cantonese shown in (12) developed from sesqui-syllables in OC. In summary, Pan claims that the Splitting-sound words found in Modern Chinese dialects, including those examples in (3), (4), (5), (6) and (7), are not retentions of Splitting-sound words in OC or Middle Chinese, but the products of the development of the sesqui-syllables in OC.

Pan's reconstruction system and his diachronic explanation of Splitting-sound words blew a fresh breeze to the studies on Splitting-sound words, but we may find many problems in his assumptions. First, his hypothesis on three developmental directions of sesqui-syllables awaits further investigation. The evidence in Modern Chinese dialects does not necessarily prove the existence of sesqui-syllables in Proto-Chinese or OC. They could be variants of complete disyllables or retentions of the Splitting-sound words in OC and Middle Chinese. Pan does not provide a clear motivation for the change from sesqui-syllables to Splitting-sound words. Nor does he give a route for this change. Second, if we accept his conclusion that Splitting-sound words reflect the existence of sesqui-syllables, we still cannot clarify why only liquids appear at

the second syllable onset of the Splitting-sound words, because there is no such a limit that only liquids could appear at the major syllable onset of sesqui-syllables. What if an obstruent appears at the major syllable onset position? Pan does not point out that only sesqui-syllable that have liquids at the major syllable onset position could develop into Splitting-sound words.

In Pan's discussion, Splitting-sound words in Modern Chinese dialects are adopted as evidence to support his 'sesqui-syllable' hypothesis. He fails to advance a detailed analysis of Splitting-sound words in OC. In the following discussion, my hypothesis on the rise of Splitting-sound words and its significance for the rise of other disyllables in OC will be proposed.

4.4 My hypothesis

Based on Sun's phonological descriptions of Splitting-sound words in OC, Pan and Serruys' diachronic views about their creation, and on Zhengzhang (2003)'s reconstructed OC system, I propose that the production of Splitting-sound words was motivated by phonological developments in OC. When consonant-cluster onsets were undergoing simplification towards single-consonant onsets around MOC and LOC, Splitting-sound words were created to preserve the identity of each of the consonants in the cluster.

4.5 Internal evidence

4.5.1 Statistics to show the simplification process of syllable onsets

According to the syllable structures of OC, Middle Chinese and Modern Mandarin introduced in Section 2.2, syllable onsets undergo huge changes, especially simplification over time. Most historical linguists attribute this to a series of regular sound changes, such as palatalization (Zhengzhang, 2003; Zheng, 2010) and labialization (Zhou, 2003).

Table 4.2 Some regular sound changes in syllable onsets from OC to Middle Chinese

Simplification of onset clusters		Examples	Gloss
palatalization	*mlj- > *ɲ-	*mljig > *ɲig 日	'sun'
	*glj- > *ʣ-	*gljag > *ʣek 石	'stone'
labialization[33]	*Kw- > *p-	*gwan > *p^hɯn 衯	'big and loose (of clothes)'

This simplification process is visible not only at the level of single lexical items, but also in terms of the form of the most stable lexemes in the language. Stable lexemes can be quantified by comparing basic vocabulary items across languages. A Swadesh list is one of several lists of vocabulary with basic meanings, developed by Morris Swadesh in the 1940–50s, which is used in lexicostatistics (quantitative language relatedness assessment) and glottochronology (language divergence dating). Wichman & the ASJP Consortium (2009) present statistics of this kind showing that of the Swadesh 100-item list, it is possible to rank items on a stability scale, where stable items are those least likely to be borrowed. Assuming, then, that these are the words mostly likely to reflect regular sound changes from OC to Middle Chinese to the Modern Chinese languages, I selected 50 items of highest stabilities across languages from the results of Wichman et al. (2009) for this research which are displayed in the following table:

[33] K stands for k, k^h and g.

Table 4.3 Results for stabilities of Swadesh list items

Rank	# In S. list	Meaning	Rank	# In S. list	Meaning
1	22	louse	26	30	blood
2	12	two	27	34	horn
3	75	water	28	18	person
4	39	ear	29	47	knee
5	61	die	30	11	one
6	1	I	31	41	nose
7	53	liver	32	95	full
8	40	eye	33	66	come
9	48	hand	34	74	star
10	58	hear	35	86	mountain
11	23	tree	36	82	fire
12	19	fish	37	3	we
13	100	name	38	54	drink
14	77	stone	39	57	see
15	43	tooth	40	27	bark
16	51	breasts	41	96	new
17	2	you	42	21	dog
18	85	path	43	72	sun
19	31	bone	44	64	fly
20	44	tongue	45	32	grease
21	28	skin	46	73	moon
22	92	night	47	70	give
23	25	leaf	48	52	heart
24	76	rain	49	36	feather
25	62	kill	50	90	white

(Adapted from Wichmann & the ASJP Consortium, 2009)

In fact, there are totally 55 words in my data, including 5 synonyms for 'I', 'tree', 'tooth', 'path' and 'full' respectively (see details in Appendix VI). I compared their pronunciations in reconstructed OC, reconstructed Middle Chinese[34] and Modern Mandarin, focusing on the change of their onsets. The distribution of two types of onsets (single consonant and consonant cluster) is illustrated below in Figure 4.2:

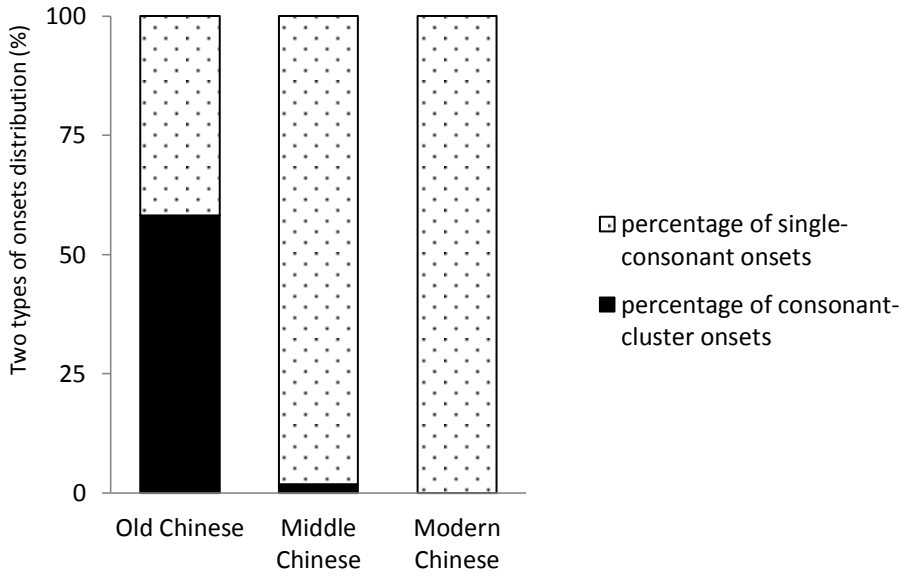


Figure 4.2 Percentage of onset types in 55 mono-syllabic words for three groups: OC, Middle Chinese and Modern Chinese

The above Figure 4.2 shows clearly the dramatic decrease of the consonant-cluster onsets from OC to Middle Chinese and Modern Chinese. In OC, there are 32 onsets that have consonant-cluster structure out of 55 onsets. In other words, around 58% of the basic vocabulary in OC includes complex onsets. In Middle Chinese, there remains just one onset of consonant-cluster structure *br-. *Cj- and *Cw- do not count as a consonant-cluster onsets in Middle Chinese, because glides in Middle Chinese are grouped as medial, a part of rhyme (Pan, 2000).

[34] Both the reconstructed OC and Middle Chinese forms are looked up in Pan's online database (www.eastling.org) and refer to his reconstruction system.

In Modern Chinese, all the consonant clusters have been simplified into single consonants. Obviously the most dramatic simplification process occurred from LOC to EMC.

As exemplified in Table 4.2, Head-and-Tail onset clusters with glides *j- and *w- underwent the sound changes of palatalization and labialization. The palatalization of *C(L)j- clusters may also be considered as fusion change. That is to say, the two segments *Cj- usually merged into one, bearing features of both the original sounds. For example, the onset structure of an unaspirated voiceless stop plus a glide *j merged into a single onset *tɕ- (Zhengzhang, 2003): the place feature of *j palatalized the stops, and the frication of the glide *j and the plosive features of the stops were also preserved in the output affricative onset *tɕ-. And the onset structure of a nasal plus a *j merged into a palatal nasal *ɲ- (Zhengzhang, 2003, p. 127). All the changes of *Cj- or *C(L)j- onsets involving palatalization are laid out in the following Table 4.4:

Table 4.4 The simplifications of *Cj-/C(L)j- onsets involving palatalization

The simplified onsets in Middle Chinese	Head-and-Tail onset clusters with Tail glide *j- in OC
*tɕ-	*tj-, *pj-, *plj-, *kj-, *klj-, *qj-, *qlj-, *qwj-
*tɕ^{h}-	*t^{h}j-, *k^{h}j-, *k^{h}lj-, *p^{h}j-, *p^{h}lj-, *m^{h}j-, *n^{h}j-, *hj-, *l^{h}j-, *r^{h}j-
*dʑ-	*dj-, *bj-, *blj-, *gj-, *glj-
*ɕ-	*q^{h}j-, *q^{h}lj-, *q^{h}wj-, *h.ŋj-, *h.nj, *h.mj-
*ʑ-	*Gj-, *Gwj-, *Glj-, *ɦ.lj-, *ɦ.blj-, *ɦ.glj-
*ɲ-	*nj-, *ŋj-, *mlj-, *ŋlj-

The palatalization of *C(L)j- onsets mentioned in Table 4.4 need further explanation. The palatalization was not completed in one step. As Zheng stated, the palatalization first occurred to the liquid *l and then, spread to the Head consonant.

(13) The palatalization of *C(L)j- clusters

*mlj->*ml^{j}->*mj->*m^{j}->*ɲ-
*glj->*gl^{j}->*gj->*g^{j}->*ʤ-

The glide *w in OC only occurs after velar consonants *k, *k^{h}, *g, *ŋ and glottal consonants *ʔ, *h, *ɦ (the later forms of *q, *q^{h}, *G). For the development of the glide *w, there are two changes that took place from OC to Middle Chinese. The first is Fusion change or labialization in *Kw- clusters at an earlier time. Zhou (2003) found there were many synonymous words with initial alternatives *Kw-/*p- in OC. He certified that they were cognates created by the fusion change - *Kw- >*K^{w}- > p-, which occurred around 1100 BC. For example, **gweŋ* 罽, 嬛 > **q^{h}wan* 嬛, 翾 > **p^{h}an* 翩 (Zhou, 2003, p. 62). For most cases of *w change that occurred around LOC and EMC time, *w was replaced with the medial vowel *u without frication. During the change into /u/, the low vowel after it is sometimes raised to a mid-vowel. For example, **gwa* > **guo* 狐‘fox’, 弧 ‘bow, arc’. And in some cases, after influencing the height of the low vowel after it, the vowel *u is centralized into *ɨ (**kwa* > **kɨo* 瞿‘frightened’, 衢‘intersection’) (Zhengzhang, 2003).

As shown in the data of Appendix VI, another sound change involved in the simplification of complex onsets around LOC and EMC is deletion. Onset structures undergoing deletion and their relevance to the rise of Splitting-sound words are discussed in the following sections.

4.5.2 Two relevant reconstructed onset structures in OC

As mentioned in 2.1.2, Zhengzhang is the first phonologist who analyzed the OC consonant onset structure in detail among the proponents of ‘consonant-cluster’ view. In Zhengzhang’s reconstruction, there are two types of onset clusters: sesqui-syllabic clusters (Crown + Head -

C.C-) and consonant clusters (Head + Tail - CC(C)-). The liquids could appear either in the Head position or the Tail. Therefore, there are two kinds of clusters involving liquids: C.L- and CL-. Given this, we can ask which of the two kinds of clusters is related to the rise of Splitting-sound words. Assuming that the origin of the liquids in Splitting-sound words is associated with C.L- structure, we cannot explain why other consonants that appear in the Head position are not preserved in the second syllable onset of Splitting-sound words, and why only liquids survive. There are only liquid and glide candidates in the Tail position. So even before looking in more detail at the changing paths of the two structures C.L- and CL-, we can hypothesize that CL is more likely to be related to the creation of the Splitting-sound words than C.L.

There is another noteworthy point about the reconstructed onsets in OC. While there are voicing and aspiration distinctions reconstructed for OC stop consonants these are not freely distributed. The distinctions are found in the Head position of the onset stem, while they are absent in Crown position. Only unaspirated voiceless stops*/p-, t-, k-/ are found in Crown position. This reconstructed feature corresponds with the research on esqui-syllables in many languages of Southeast Asia. In some Southeast Asian languages, such as Proto-Thai, the stops in the minor syllable onset of sesqui-syllables are also limited to unaspirated voiceless stops when they are followed by the liquids in the major syllable onset (Matisoff, 1973; Pittayaporn, 2009). What is more, from this feature we can deduce that the origin of Splitting-sound words is more likely be in *CL- than *C.L-: recall that we have already found voiced stops in the first syllable onset of **ɢʷaa raag* 瓠落 in (2f) and **doog roo* 髑髏 in (2d). If these disyllables are split versions of monosyllables, the original complex onset would contain a Dr cluster, D a voiced stop, and so, would constitute a CL and not a C.L sequence.

4.5.3 Clues from variants of Splitting-sound words

According to the definition of Splitting-sound words, the onset of the monosyllabic form that a Splitting-sound word is derived from is usually *C-. As shown in Figure 4.1, *CL- could also appear in the onset. This variant is reflected in the following example:

(14) A variant of the monosyllabic form of Splitting-sound pattern

Monosyllabic form	Gloss	Splitting-sound forms	Gloss
*ziilʔ/*zliilʔ荠	'Tribulusterrestris'	*zid riil 蒺藜	'Tribulusterrestris'

What's more, the onset of the first syllable of a Splitting-sound word could be a three-consonant cluster *C(G)L-.

(15) Variants of the disyllabic form of Splitting-sound pattern

a. 伯勞 'shrike' *praag raw (Zhengzhang's system)

b. 鈐鐪 'a kind of digging tool' *gram loolʔ (Pan's system)

c. 臾曳 'to tow' *gruds leds (Zhengzhang's system)

d. 熠燿 'sparkling' *ɢrub lewɢs (Zhengzhang's system)

e. 華離 'crooked' *q^{h}wraa reels (Zhengzhang's system)

These variants of the Splitting-sound pattern reflected in the OC data cannot be interpreted under a reduplication approach. If we seek a diachronic interpretation, it will be easier to find the answer: the inconsistent onset forms suggest that the complex onset was in an unstable state in LOC. It is hypothesized that this change might be associated with the rise of Splitting-sound words. According to the reconstructed Middle Chinese, the first syllable *C(G)L- onset of all the

Splitting-sound words in (15) are simplified into single member *C- onsets in Middle Chinese. For example, (15a) **praag* 伯 > **pak*; (14b) **gram* 鈐 > **giam*.

4.5.4 Different Change routes

*CL- and *C.L- have different sound change routes, though both underwent simplification from OC to Middle Chinese (Pan, 1999). The liquids in *CL- were either lost or became semi-vowels, while the plosives proceeding them were preserved (Xu & Pan, 1994). The following examples show the simplification route of *CL-> *C-:

(16) Examples of the simplification from *CL- to *C-:

O.C.	Middle Chinese	characters	glosses
*k^hwlas	*k^huo	褲	'pants'
*glu	*gou	球	'ball'
*blids	*bi	鼻	'nose'
*bral	*bie	皮	'hide'

In the above data, Middle Chinese sound forms are established ones, with many clues found in Middle Chinese dictionaries, like *Qiēyùn* and *Guǎngyùn*. Xu and Pan further support the reconstruction of consonant cluster onsets with the aid of loans in neighboring Austroasiatic languages from OC. For example, 'pants' in Va [wbm] is *k^hlab?*, and 'ball' in Wuming dialect of Zhuang [zyb] is *klau*. The assumption is that these loans date back to OC and that the loss of the liquids from OC to Middle Chinese is evident in forms like those in (16).

As to the change of *C.L-, things get a little complicated. Pan (1999) claimed that during the simplification of the sesqui-syllabic cluster C. L in LOC, the minor syllable consonant was lost, while the liquid onset of the major syllable was preserved in Middle Chinese. Unlike the C

in CL clusters whose intensity is higher than that of the liquid, the syllabic consonants in the minor syllables are more easily lost, with lower articulatory intensity than that of the following liquid. The simplification route of C.L structure is exemplified in (17):

(17) Examples of the simplification of C.L onset from OC to Middle Chinese

	OC	Middle Chinese	Characters	Pinyin	Glosses
a.	*k·raŋ	ljaŋ	量	liáng	‘quantity’
b.	*k·raŋ	ljaŋ	涼	liáng	‘cool’
c.	*p.rel	lie	籬	lí	‘bamboo fence’
d.	*k.raam	lam	藍	lán	‘blue’
e.	*k.roo	lu	楼	lóu	‘building’

Pan (1999) reconstructed the OC forms in (17) based on cognates from other Sino-Tibetan languages and on loans from or in neighboring genetically unrelated languages. For the reconstruction of the words in (17a-17c), he cites the cognates from Tibetan [bod] in (18) as evidence.

(18) Cognates from Tibetan suggesting OC complex onsets

OC	Characters	Middle Chinese	Glosses (Chinese)	Tibetan cognates	Glosses (Tibetan)
*k·raŋ	量	ljaŋ	‘quantity’	graŋs	‘number’
*k·raŋ	涼	ljaŋ	‘cool’	graŋ	‘cold’
*p ɾel	籬	lie	‘bamboo fence’	bres	‘hedge’

Pan also found evidence from Chinese loans in Thai [tha] and Vietnamese [vie], and Uyghur [uig] loans in Chinese. For instance, 藍 ‘a blue plant’ – a loan word from OC is pronounced as *gram*

in Thai. So we can confirm that the original form must have had a plosive before the liquid ‘r’, otherwise the loan form should be adapted as *ram*, since in Thai /r/ is a distinctive segment that can be used as a single onset. Lóulán 楼兰 was borrowed from Uyghur *Krorayina* into OC, referring to a famous kingdom in Central Asia, a relic in Xinjiang Uyghur autonomous region. Although its Middle Chinese pronunciation is **lulan*, we conclude that in OC the adapted form for the first syllable 楼 should be **k.roo* with a plosive before the liquid *r.

Similar change routes are also found in Tai-Kadai languages with sesqui-syllables (Kosaka, 2007). Pittayaporn (2009) gives a clear analysis of the development of sesqui-syllabic clusters with medial liquids in Proto-Tai. When the initial consonant C is a voiceless plosive, the Proto-Tai sesqui-syllabic *C̥.L- has two developmental directions:

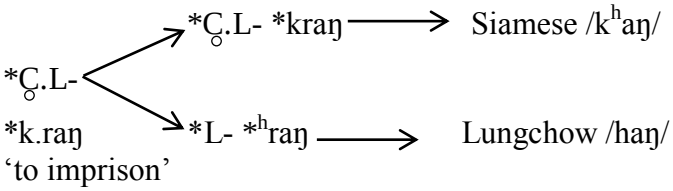


Figure 4.3 The development of sesqui-syllable clusters in Proto-Tai

In Proto-Tai, a sesqui-syllable cluster with a voiceless plosive initial consonant C̥ could keep both consonants but became a tauto-syllabic cluster (Pittayaporn, 2009). However, in some dialects, the Proto-Tai sesqui-syllabic *C̥.L- underwent simplification and became *ʰL- before the mono-syllabification took place. This change route in Proto-Tai may be seen as an parallel development to the simplification of sesqui-syllabic clusters in OC.

4.5.5 Synthesis

Based on the analysis of the reconstructed onset types in OC and of the special sound features of Splitting-sound words, I suspect *C.L- and *CL- structures are likely to be related to

the creation of Splitting-sound words. In order to make it clear which structure is the source of Splitting-sound words, we need to seek out the motivation for their creation. The statistics in Figure 4.3 show that onsets underwent simplification from OC to Middle Chinese. And this change occurred throughout the whole OC period, though in EOC it was very slow. It is the simplification of the onset in OC that furnishes the motivation for the creation of Splitting-sound words. *C.L- and *CL- onsets have different change routes in the simplification process: the former changed into L, with the consonants in the minor syllable lost; the latter turned into C, with the liquids lost. This examination leads to the deduction that *CL- is more likely to be the initiator of Splitting-sound words than the *C.L- structure. It was the loss of liquids in the simplification process that motivated the creation of Splitting-sound words where the liquids are preserved as the second syllable onsets. On the contrary, liquids in *C.L- were preserved in the simplification, so there was no need to "preserve" them in Splitting-sound words. Someone may argue that the loss of C in the simplification of *C.L- structure could lead to the creation of Splitting-sound words. But since the lost C existed in the minor syllable, and reduction or loss of minor syllables is the natural simplification process in sesqui-syllabic languages under a drift of mono-syllabification, the simplification of *C.L- cannot be considered as the motivation of Splitting-sound words. Furthermore, there are lots of reduced sesqui-syllable combinations besides *C.L- (e.g. *C.N-, *C.T[35]-, according to Zhengzhang's reconstructed consonant structure in Table 2.3). Since all of them underwent simplification, it is hard to explain why only *C.L- should motivate the creation of Splitting-sound words. Finally, recall that some splitting-sound words reflect voiced obstruent initials, and therefore cannot be reconstructed as initial elements in sesqui-syllables, since these elements lacked specified laryngeal features.

[35] N here stands for nasals, and T stands for obstruents.

Though splitting-sound words lost their productivity long ago, sporadic retentions of splitting-sound words can still be found in some Modern Chinese dialects. These examples support the hypothesis of the origins of Splitting-sound words. An example from the Wenzhou dialect of Wu is identified to illustrate the hypothesis.

Table 4.5 An example from Wenzhou dialect

Monosyllabic form		Splitting-sound word	Gloss
OC	Middle Chinese		
*grooŋs 巷	*ɦuŋ	巷弄 /haŋ loŋ/	'lane'

The Splitting-sound word /haŋ loŋ/ 巷弄'lane' in this dialect is a product of splitting process from the monosyllable 巷 /haŋ/ 'lane'. The word 巷 'lane' underwent great sound changes from OC to Modern Chinese: **grooŋs* (OC) > **ɦuŋ* (Middle Chinese) > *haŋ* (Modern Wenzhou). The liquid /r/ was lost in the onset simplification process, generating the disyllable **ɦuŋ luŋ* 巷弄. The Splitting-sound word /haŋ loŋ/ found in Wenzhou is the retention of the Splitting-sound word in LOC and EMC. The character 弄 (**rooŋs* > **luŋ*) in OC meant 'to play', which had nothing to do with the meaning 'lane'. But since sharing the same rhyme with 巷 'lane' and had a liquid onset, it was adopted to form a Splitting-sound word with 巷 'lane' to preserve the lost liquid in the onset. The same-rhyme feature of the Splitting-sound word had been preserved until the vowel of 巷 *ɦuŋ* was changed into /a/ in Modern Chinese varieties. Since the Splitting-sound form has the same meaning as the original monosyllabic form, 巷弄 *haŋ loŋ* can be seen as a counter-example to the Fission reduplication argument. Moreover, this case provides a clear picture of how and why Splitting-sound words came into being.

Based on the range of internal evidence, I conclude that the rise of Splitting-sound words is motivated by the simplification of consonant-cluster onsets. More specifically, the loss of liquids in complex *CL- onsets involved simplification and neutralization of CL > C. Awareness of this neutralization provided a phonological motivation to preserve the lost liquids in the cluster and so, to preserve the contrast between *CL- and *C- onsets.

4.6 External evidence

Due to the scarcity of evidence from reconstructed Old Chinese and Proto-Chinese, and because of the argument as to whether there existed consonant-cluster onsets (esp. the sesqui-syllable structure) in Proto-Chinese, it is necessary to resort to external evidence to see if there is any clue from neighboring languages to the hypothesis that the pattern of Splitting-sound words should be related to the consonant-cluster onsets.

The Yao language [ium], also called Mien in Vietnam, belongs to the Hmong-Mien family. Early linguists classified the Hmong-Mien languages into the Sino-Tibetan family, but the current consensus among Western linguists is that they constitute a family of their own. The family is said to have had its origin in Western-southern China and gradually spread to neighboring Southeast Asian countries, as the speakers migrated to Thailand, Burma, Vietnam and Laos within the past 300-400 years (Blench, 2004). Many Mien dialects still preserve complex onsets in their phonology. Zhao, B. & Zhu (1998) find an interesting corresponding relation between the Splitting-sound words in the dialects of Jin and the monosyllabic cognates in Yao language:

Table 4.6 The onset correspondence between the Splitting-sound words in Jin and the monosyllables in Yao

a. Taiyuan (Jin)			Quanzhou dialect of Yao	Gloss
Splitting-sound form		Monosyllabic form	Monosyllabic cognates	
薄拉	pəʔ$_8$la$_1$	爬 pa$_1$	plan$_2$	'to crawl'
圪老	kəʔ$_8$lɔ$_3$	攪 kiɔ$_3$	klwo$_2$	'to stir'
扑涝	pəʔ$_8$lau$_3$	跑 p^{h}au$_3$	plau$_5$	'to run'
圪佢	kəʔ$_8$lɔ$_1$	角 kiɔ$_1$	klɔ$_1$	'corner'
骨拢	kuəʔ$_8$luŋ$_3$	滾 kuŋ$_3$	klin$_5$	'to roll'
泼棱	p^{h}əʔ$_8$ləŋ$_1$	蓬 p^{h}əŋ$_1$	plɔŋ$_2$	'classifier'
窟窿	k^{h}əʔ$_8$luŋ$_1$	孔 k^{h}uŋ$_1$	k^{h}lɔŋ$_1$	'hole'
圪离	kəʔ$_8$li$_1$	-	kli$_1$	'armpit'
刻撩	k^{h}əʔ$_8$liau$_1$	翹 k^{h}iau$_1$	kləi$_2$	'to arch'
骨聯	kuəʔ$_8$lyɛ$_7$	卷 kyɛ$_7$	gli$_1$	'to curl'

(Adapted from Zhao &Zhu, 1998)

Based on the data in Table 4.6 provided by Zhao & Zhu (1998), Deng (2007) compared them with the dialects of North Min and OC reconstructions, presenting a striking result that all the complex CL onsets in OC monosyllables and the dialect of Mien are reflected in the onsets of the Splitting-sound words of the dialects of Jin and North Min.

Table 4.7 The onset correspondence between the Splitting-sound words in Modern Chinese varieties and the monosyllables in Quanzhou dialect of Yao and OC

Taiyuan Dialect (Jin)			Gaizhu Dialect (North Min)				Quan-zhou dialect of Yao	OC	Gloss
Splitting-sound form		Mono-syllabic form	Splitting-sound form		Mono-syllabic form		Mono-syllabic cognates		
薄拉	pəʔ$_8$la$_1$	爬 pa$_1$	爬□	pa$_2$la$_2$	爬	pa$_2$	plan$_2$	*braa 爬	‘to crawl’
圪老	kəʔ$_8$lɔ$_3$	攪 kiɔ$_3$	攪□	ko$_3$lo$_3$	攪	ko$_3$	klwo$_2$	*kruuʔ攪	‘to stir’
扑涝	pəʔ$_8$lau$_3$	跑 p^{h}au$_3$	抛□	p^{h}o$_1$lo$_1$	抛	p^{h}o$_1$	plau$_5$	*bruu 跑	‘to run’
圪佢	kəʔ$_8$lɔ$_1$	角 kiɔ$_1$	角□	ku$_7$ lu$_7$	角	ku$_7$	klɔ$_1$	*kroog 角	‘corner’
骨拢	kuəʔ$_8$luŋ$_3$	滾 kuŋ$_3$	滾□	kuaŋ$_3$luaŋ$_3$	滾	kuaŋ$_3$	klin$_5$	*kluunʔ滾	‘to roll’
泼棱	p^{h}əʔ$_8$ləŋ$_1$	蓬 p^{h}əŋ$_1$	蓬□	p^{h}aŋ$_1$laŋ$_1$	-		plɔŋ$_2$	*blooŋ 蓬	‘classifier’
窟窿	k^{h}əʔ$_8$luŋ$_1$	孔 k^{h}uŋ$_1$	空窿	k^{h}aŋ$_1$laŋ$_1$	-		k^{h}lɔŋ$_1$	*k^{h}looŋʔ孔	‘hole’
圪离	kəʔ$_8$li$_1$	-	胳肢	ku$_5$li$_5$	-		kli$_1$	*klaag 胳	‘armpit’
刻撩	k^{h}əʔ$_8$liau$_1$	翹 k^{h}iau$_1$	翹□	k^{h}iau$_5$liau$_5$	翹	k^{h}iau$_5$	kləi$_2$	*kloŋʔ 翹	‘to arch’
骨聯	kuəʔ$_8$lyɛ$_7$	卷 kyɛ$_7$	□□	ki$_7$li$_7$	-		gli$_1$	*glonʔ 卷	‘to curl’

Table 4.7 shows that there are five Splitting-sound words whose mono-syllabic forms could not be identified in Taiyuan and Gaizhu dialects. Among the OC mono-syllabic words, only the Splitting-sound form of the word **kroog* 角 ‘corner’ could be found in OC literature as **qoog*

roos 屋漏 ‘the west corner of a room’. Although both of them fall into the Chinese language family, Jin [cjy] 晉語 and North Min [mnp] are not dialects of Mandarin and differ much in phonology and lexicon. They are also geographically far from each other, so similarities cannot be due to borrowing. So the correspondence set between these two languages is not a coincidence or a result of borrowing, but reflects the Splitting-sound trait of their ancestor language – Middle Chinese and OC. In this sense, the words in 4.7 could be considered retentions of Splitting-sound words in Middle Chinese or OC. The striking correspondence among the complex CL onsets in the OC mono-syllables and the Mien mono-syllables, and the single onsets of the split disyllables in Modern colloquial Chinese varieties confirms that the motivation of the rise of Splitting-sound words in OC is closely related to the simplification of the complex onsets. Furthermore, according to Middle Chinese reconstruction, almost all the complex *CL onsets of the ten OC mono-syllabic forms in Table 4.7 were simplified into *C onsets in Middle Chinese words with the loss of liquids *r, l.

(19) The Middle Chinese sound of the ten OC monosyllables:

OC	Middle Chinese	Gloss
*braa 爬	*ba	‘to crawl’
*kruuʔ攪	*kau	‘to stir’
*bruu 跑	*bauk	‘to run’
*kroog 角	*luk	‘corner’
*kluunʔ 滾	*kuon	‘to roll’
*blooŋ 蓬	*bun	‘classifier’
*k^{h}looŋʔ 孔	*k^{h}un	‘hole’
*klaag 胳	*kak	‘armpit’
*kloŋʔ 翹	*gjeu	‘to arch’
*glonʔ 卷	*kiuen	‘to curl’

Although the data in (19) include an exception --- 'corner', in which the *kr- onset changed into *l- rather than *k- in Middle Chinese (standard Middle Chinese), some Chinese varieties display different change routes: in Cantonese, 'corner' could be pronounced as either kok_3 or luk_6; in Wu language, the sound development is *$kroog$ > *kok > $ko?_4$. This also indicates that the loss of Liquid in the simplification of *CL- onset is a common phenomenon in Middle Chinese varieties. Thus, we have reason to argue that the rise of Splitting-sound words took place in order to preserve the liquid being lost in the simplification process of complex *CL- onsets around LOC and EMC.

This study also attempted to find evidence in Austroasiatic languages. I take advantage of the online Mon-Khmer etymological dictionary[36] of the SEAlang Mon-Khmer Languages Project[37], based on the assumption that Austroasiatic languages may have had some early connections with Proto-Sino-Tibetan[38]. The phonological forms of the Splitting-sound words were looked up in the dictionary with the aid of their glosses for comparison of their onsets.

[36] It provides an on-line hierarchical reference that puts the data in context. It is based on the late H.L. Shorto's Mon-Khmer Comparative Dictionary (2006).

[37] This project is directed by Paul Sidwell and Doug Cooper, with the support of the National Endowment for the Humanities, 2007-2009 and 2009-2011.

[38] Most scholars think Sino-Tibetan and Austroasiatic languages belong to separate families because phonologically most words in Austroasiatic languages are disyllabic morphemes or a sesqui-syllable followed by a major syllable, while Sino-Tibetan features in its mono-syllabic morphemes. However, some linguists have brought it up that Proto-Chinese or OC has strong lexical, phonological and grammatical connections with the Tibeto-Berman languages to the west and with the Southeast Asian languages to the south, especially Tai-Kadai, Hmong-Mien and Mon-Khmer (Norman & Mei, 1976; Sagart, 1995, 1999; DeLancey, 2011). DeLancey (2011) proposes an explicit model of the contact, suggesting that Proto-Chinese could have been widely used as a lingua franca by non-Chinese outside the Chinese state at that time until the widespread lingua franca version replaced the original everywhere ultimately.

Table 4.8 Onset correspondence between Splitting-sound words in OC and potential cognates in Mon-Khmer languages

OC Splitting-sound words	Glosses	Potential Cognates	Mon-Khmer Language Source	Glosses
*gaa l^{h}eeb	'butterfly'	kla kạr-lā	Rumai, in Palaungic Branch [rbb] Palaung, in Palaungic Branch [pce]	'butterfly'
*p^{h}ug l^{h}uu	'larval locust'	p^{h}luŋ	Mal, in Khmunic Branch [mlf]	'locust'
*kw ə́ŋ^{j} r ə̀ŋ^{j}	'a kind of mouse'	krɛŋ kooŋ rajah	Sedang, in Bahnaric Branch [sed] Katu, in Katuic Branch [kuf]	'mouse deer' 'Mouse, rat'
*qoog roos	'the west corner of a room'	cruŋ c^{h}ruŋ *cʔuk	Khmer, in Khmeric Branch [khm] Pear, in Pearic Branch [pcb] Proto Mon- Khmer	'corner'
*gaa renʔ	'a kind of sacrificial vessel'	*kraʔ *crɒŋ	Proto Mon-Kmer Proto Wa-Lawa	'kind of jar' 'cup, bowl'
*g^{w}a log	'mynah'	kəloiŋ	Mon[39] [mnw]	'mynah'

So far there is not any semantic clue for massive lexical borrowing in either direction, although the languages have been in contact with Chinese for more than 2,000 years (DeLancey, 2011). Because they are from different language families, and no regular sound correspondences

[39] Shorto, 2006:C:757-2 via **[k]rliiŋ*

between sets of words could be established, I consider their sound similarities as clues to potential cognates.

From Table 4.8, we can see that the single-consonant onsets of the two syllables in Splitting-sound words are reflected either in the consonant-cluster of monosyllabic 'potential cognates', or in the two onsets of disyllabic 'potential cognates' in neighboring languages. Therefore, I infer that the relatively conservative Mon-Khmer languages still preserve their consonant clusters and disyllable structures even after mono-syllabification, while OC underwent the change from disyllables to sesqui-syllables and to monosyllables, and the consonant-cluster onsets of monosyllables also underwent simplification. Although in Modern Chinese, monosyllables with simple consonant onsets are preserved as the common mono-morphemic forms, we still see traces of evidence for the earlier existence of sesqui-syllables and consonant-cluster onsets in OC, especially in EOC, with the comparison of the sound patterns of Splitting-sound words and the 'potential cognates' in Mon-Khmer languages. With this evidence, we propose that the rise of Splitting-sound words occurs to preserve an older form: the lost liquids of the consonant-cluster onsets.

4.7 Summary

To sum up both internal and external evidence, the rise of Splitting-sound words is the result of a reactive process, with an attempt to highlight or preserve the segment which has disappeared or is disappearing in mono-syllabic words. The segment here refers to the second consonant – liquids *l or *r, in the consonant-cluster onset which was lost in the onset simplification process. This explains why the onset of the second syllable in Splitting-sound words in OC data is always the liquid *l or *r.

However, there are some Splitting-sound words in the data that do not strictly obey the sound pattern prescribed in the definition: the onset of the first syllable is a consonant cluster CGL, including a liquid in the Tail position. An example from (15f) is rewritten in the following (20).

(20) A Splitting-sound word with the first-syllable onset cluster CGL

OC	Character	Gloss
*q^hwraa reels	華離	'crooked'

According to my hypothesis, a Splitting-sound word results from the simplification of the complex onset. Therefore, the lost liquid is not expected to appear in the second-syllable onset of the Splitting-sound word. In an evolutionary framework, this 'odd' case, like (20) is well studied and given a good interpretation. A pattern of sound change may show exceptions if the sound change is gradual and observed 'in progress' (Blevins & Wedel, 2009). The exceptions reflected in (15) just confirm the hypothesis that the sound change of CL > C was gradual and in progress in OC. The evolutionary approach can not only offer explanations for the exceptions to the Splitting-sound pattern, but also provide a likely scenario giving rise to Splitting-sound words: During the gradual sound change, older speakers, with CL representation attempt to maintain CL by (i) writing CL-words with two characters, - one to represent the initial C, another to represent the L; and by (ii) pronouncing them slowly, in a fortis way, so that the Head C is expanded to a full syllable, and the rest of the original syllable with the L and the rhyme is preserved. In addition to the preservation of the liquid, the paradigmatic contrast between *CL- and *C- in OC is also maintained.

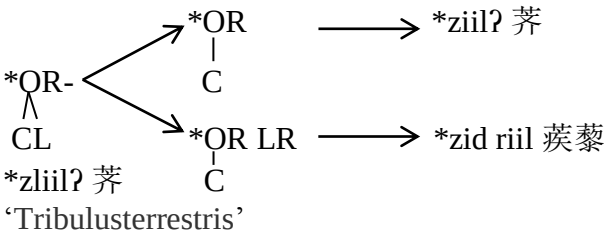


Figure 4.4 An exemplar to show the results of CL > C sound change

In order to preserve the phonological contrast between *zl- and *z- being lost when the CL > C sound change was in progress, the Splitting-sound form was created by older speakers to maintain and accentuate the paradigmatic contrast. This evolutionary explanation also gives an account for the myth of why there are so many mono-syllabic forms corresponding to the Splitting-sound words in the data that are missing. In the example of Figure 4.4, the mono-syllabic product of the change eliminated the contrast between *zl- and *z-, which had already existed in the OC onset structures. Compared with the mono-syllabic form, the disyllabic Splitting-sound word preserves more of the original segments in a more stable situation. Therefore, the Splitting-sound form won out, while the category of mono-syllables was gradually deserted.

The reason Splitting-sound words associated with this precise kind of sound change are not found in Tibeto-Burman and Mon-Khmer languages, is because the ancient writing systems of both families WRITES the CL cluster. Because the conservative pronunciation is visible in the written language, it can always be recovered, - there is no need to hyper-articulate or to devise a new written form of words to express the CL cluster that is disappearing in the vernacular. In contrast, as discussed in Chapter 2, Chinese is not a language in which subsyllabic onset structure is typically represented. For this reason, a CL > C sound change could not be recovered from a single written character.

The onset simplification process came to an end around LOC and EMC. This was just the period when Splitting-sound words lost their productivity. Without phonological motivation, they lost their "raison d'etre". However, the influence of their sound pattern was lasting, as discussed in the next two chapters. As the earliest type of mono-morphemic disyllable, splitting sound words created a seed of disyllabicity which continued to grow in the course of history of the Chinese language.

CHAPTER 5

THE RISE OF DIÉYÙN WORDS IN LATE OLD CHINESE:

A DISYLLABIC TEMPLATE

5.1 Introduction

Lianmian words enjoy a unique status in Chinese history because of their internal sound correlations and their mono-morphemic status. What's more, they are the earliest disyllabic forms in Chinese that are evidenced in classical literatures. The study of Lianmian words may help reveal the formation of the disyllabic template that governs Modern Chinese prosody.

The formation of disyllabic compounds that appeared around LOC and EMC needs to be connected with Lianmian word formation. Previous studies of Lianmian words ignored this connection and their influence on the formation of Chinese prosody. Researchers preferred analyzing Lianmian words synchronically as products of reduplication. Feng (1998) and Acordia (2007), who studied the formation of disyllables, focused primarily on the disyllabic compounds that still preserve their productivity in Modern Chinese, excluding discussion of Lianmian words. Thus, the story of the rise and the development of disyllables in Chinese has never been complete.

In Chapter Four, I discussed the rise of Splitting-sound words, which are considered a special kind of Lianmian word. I claim in this chapter that Lianmian words are the source of the long Chinese disyllabicity 'river'. The rise of Splitting-sound words is the root of the 'source'; that is to say, the sound pattern of Splitting-sound words had been generalized into a disyllabic template when it gave rise to Diéyùn words - a typical kind of Lianmian words.

5.2 What are Lianmian words and Diéyùn?

Let us review the main characteristics of Lianmian words. In a broad sense[40], Lianmian words, or Lianmianzi ('connective characters') are defined as mono-morphemic disyllables, where there is normally no semantic connection between the two single syllables (characters) and the whole word (Sun, 1999). In other words, neither of the two single syllables (characters) could be used as an independent free morpheme. Lianmian words were possibly created during the Late Western Zhou dynasty (771 BC), since they were not found in unearthed oracles of the Shang dynasty (16^{th} - 11^{th} century BC), but are found in the inscriptions on the unearthed bronze vessels made in Western Zhou. And they gradually lost their productivity after Qin and Han (206 BC to 200 AD) (Zhu, 1988; Xu, 2003).

Traditional morphologists simply classify Lianmian words into reduplicated words and non-reduplicated words. Reduplication here only refers to total reduplication. Recently, linguists seek to examine them more closely from a phonological view. Here I adopt a phonology-oriented classification and describe each category in detail.

Lianmian words can be classified into two groups according to their phonological forms and origins: the first group contains the words in which there are some sound correlations between the two syllables; the second group consists of the words in which there are no sound correlations between the two syllables. Most words of this group are loanwords borrowed from Old Persian roots (Feng, 2004; Miao, 2005).

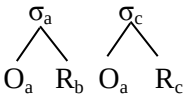
[40] There is a narrow definition that only includes words where there are some sound connections between their two syllables (Wang, 1944).

Table 5.1 OC loans from Old Persian

OC	Character	Pinyin	Persian Romanization	Gloss
*mood sug	苜蓿	mùxù	*buksuk	‘alfalfa’
*baa dagw	蒲桃	pútáo	budawa / buda	‘grape’
*q^{h}laaʔ p^{h}raag	琥珀	hǔ pò	keherba	‘amber’
*koo k^{h}ɯʔ	枸杞	gǒ qǐ	gojeh	‘wolfberry’

The first group of Lianmian words is further classified into three types: total reduplicated words (doubling), partial reduplicated words and Splitting-sound words 分音詞. There are two subtypes of partially reduplicated words: Shuāngshēng 雙聲 ‘paired initial’ (a disyllable consisting of two syllables with the same onset) and Diéyùn 疊韻 ‘duplicated rhyme’ (a disyllable consisting of two syllables with the same rhyme). The phonological patterns are displayed in the following charts:

a. Shuāngshēng



b. Diéyùn

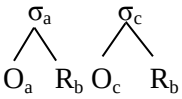


Figure 5.1The regular phonological patterns of Shuāngshēng and Diéyùn

These two subtypes are found in the majority of Lianmian words. In *Shījīng* 詩經‘Classic of Poetry’, the earliest existing collection of over 300 poems, written between early Western Zhou to the Spring and Autumn Period (11th century BC – 476 BC), there are 73 Shuāngshēng and Duéyùn words out of 99 Lianmian words (Xiang, 1987). In (1) and (2) below, just a few examples are given:

(1) Shuāngshēng examples:

A. **se so* 斯须 sī xū ‘a little while’

From *Lǐjì.Yuèjǐ*: 禮 樂 不 可 斯須 去 身.

Etiquette music not able a.little.while leave body

‘Noble people cannot live for a little while without etiquette and music.’

B. **deg dog* 踟躕/蹢躅 zhí zhú ‘to pace up and down’

From *Shījīng*: 愛 而 不 見， 搔 首 踟躕

Love and not see scratch head pace.up.and.down

‘Loving and not seeing her, I scratch my head, and am in perplexity.’

C. **k^{h}eeds k^{h}eed* 契闊 qì kuò ‘to be separated far away’

From *Shījīng*: 死 生 契闊， 與 子 成說

death live to.be.separated.far.away with you make promise

‘For life or for death, however separated, to our wives we pledged our word.’

(2) Di éyùn examples:

A. **sk^{h}aaŋ kraaŋ* 倉庚 cāng gēng, ‘oriole’

From *Shījīng*: 春 日 載 陽， 有 名 倉庚

Spring day become warm have sing orioles

‘Orioles are singing in the spring sunlight.’

B. **qaŋ tjaŋ* 鞅掌 yāng zhǎng ‘busy and tired’

From *Shījīng*: 或 王事 鞅掌

some business assigned by King tired

‘Some are busy and exhausted.’

C. **sew lew* 逍遙 xiāo yáo ‘free and unfettered’

From *Zhuāngzǐ*: 逍遙 游

free travel

‘free journey.’

5.3 My analysis of the origin of Diéyùn words

This study focuses on the discussion of Diéyùn because these words display a similar sound pattern to that of Splitting-sound words. I propose that the origin of Diéyùn words should reflect the formation of a disyllabic template (σσ). This proposal is built on two previous studies of Lianmian words which are in the framework of synchronic and diachronic theories respectively.

5.3.1 Reduplication theory

Sun (1999), in his dissertation, gives abundant data and a clear phonological description which is useful to the current study. He found that the reduplicated forms of the Diéyùn pattern exhibit a fixed liquid *r- or *l- in the onset position of the second syllable. This feature, also found in Splitting-sound words, shows the possibility of a connection between the rise of Splitting-sound words and that of Diéyùn.

In Sun’s data, there are 46 Diéyùn words displaying the $O_aR_bLR_b$ sound pattern. The occurrence of Liquids in the onset position of the second syllable is independent of the consonant feature of the onset of the first syllable (O_a), and free of the semantic implication of the reduplicated forms as well. In addition, the occurrence of Liquids in the onset position of the second syllable has nothing to do with the part of speech of the reduplicated forms either. Here I cite some examples from Sun’s data (1999, p. 59 - 60) to show this feature:

(3) Di éyùn examples with $O_aR_bLR_b$ sound pattern:

OC (Pulleyblank System)	Character	Pinyin	Gloss	Source
*ʔw ài l ài	委蛇	wěi yí	'winding, roundabout'	*Shījīng*
*màŋs làŋ	望洋	w àng y áng	'description for looking up'	*Zhuāngzǐ*
*gàŋ ràŋ	强梁	qi áng li áng	'fierce, intrepid'	*Lǎozǐ, Zhuāngzǐ*
*ts^{h}àŋ^ɥ làŋ^ɥ	從容	cōng róng	'at leisure, casually'	*Shàngshū, Lǐjì*
*s àw l àw	逍遙	xiāo yáo	'free and unfettered'	*Zhuāngzǐ*

Sun summarizes that the $O_aR_bLR_b$ sound pattern is the major phonological pattern of Di éyùn words, while some other variants are minor patterns. In his Di éyùn data, there are 16 words that display the $O_aR_bNR_b$ pattern (exemplified in (4), where N stands for *n, m and ŋ), 13 words where the second-syllable onset is glottal *ʔ- or *x- (represented in (5)), and 8 words where the second-syllable onset is a C_1^2r- cluster *sr-, *sŋr-, *tsr- (shown in (6)).

(4) Di éyùn with $O_aR_bNR_b$ sound pattern (Sun, 1999, (22), (24)):

OC	Character	Pinyin	Gloss	Source
*t àkᶣ ŋàkᶣ	鸅鴜	zh ú y ù	‘kind of aquatic bird’	*Shījīng*
*gáŋ máŋ	鸿蒙	h óng m éng	‘chaos’	*Zhuāngzǐ*
*ts àɣs n àɣs	沮洳	j ù r ù	‘moist, damp’	*Shījīng*
*tsàw ŋjáw[41]	僬僥	jiāo yáo	‘a pygmy country’	*Guóyǔ*
*ʔ àlʔ n álʔ	猗儺	yǐ nuǒ	‘soft and graceful’	*Shījīng*
*t àɥ n àɥ	侏儒	zhū rú	‘drawf’	*Zuǒzhuàn*
*təjʔ nəjʔ	蓫荽	dǐ nǐ	‘kind of grass’	*Eryǎ*
*ʔj àwʔ nj àwʔ	蕘繞	yǎo rǎo	‘the root of the marrow leaved polygala’	*Eryǎ*
*gwət mət	崛岉	ju é w ù	‘towering’	*Lúlíngguān gdi ànf ù*
*p^{h}alʔ ŋálʔ	駊騀	pǒ ě	‘description of horse’s shaking its head’	*Shījīng*

(5) Di éyùn, in which the second onset is *ʔ- and *x- (from (25) and (27) in Sun, 1999):

OC	Character	Pinyin	Gloss	*Source*
*t^h ák ʔw ák	尺蠖	chǐ huò	‘inchworm’	*Yìjīng*
*b àɣ ʔ àɣ	蚹蜎	fū yú	‘common house centipede’	*Fāngyán*
*t^h àkʷ ʔj àkʷ	綽約	chuō yuē	‘graceful’	*Sh àngl ńf ù*
*x àɥʔ ʔ àɥʔ	煦嫗	xǔ yǔ	‘description of warmness’	*Lǐjì*
*b ánʔ ʔw án	伴奂	b àn hu àn	‘beautiful’	*Shījīng*

(6) Di éyùn, in which the second onset is C_1^2r- cluster (from (29) in Sun, 1999):

[41] Recall that in Pulleyblank’s reconstruction system *j- is considered a medial rather than a part of the onset, so rimes are not technically identical.

OC	Character	Pinyin	Gloss	*Source*
*ʔwˤànˀ tsrwˤànˀ	蜿蟤	wān zhuān	‘a kind of snake’	*Jiusi*
*gˤàɣ srˤàɣ	蘧蔬	qú shū	‘a kind of fungus’	*Eryǎ*
*bˤàɣ sŋrˤàɣ	扶疏	fú shū	‘of branches and leaves’	*Hánfēizǐ*
*gəw srəw	蝤螋	qiú sōu	‘a kind of worm’	*Guǎngyǎ*

Although based on reconstruction systems (Zhengzhang, Pan, Li and Baxter) different from Pulleyblank’s system that Sun relies on, similar sound patterns of Diéyùn are also found in the data collected in this study: 59 out of 172 Diéyùn words display the $O_aR_bLR_b$ sound pattern; 24 Diéyùn words the $O_aR_bNR_b$ pattern; 31 words display the $O_aR_bTR_b$ pattern (T for obstruents - consonants except *l, and *r, such as *p^h, p, b, t^h, t, d, q^h, q, k, g, G, s^h, s, z); 29 words display the $O_aR_bC_1{}^2LR_b$ pattern, in which the second onset includes a Crown consonant and a Head consonant plus a liquid *l- or *r-; the last category is the $O_aR_bCGR_b$ pattern (a consonant plus a Glide *j- or *w-), with 27 Diéyùn words. Besides the above five categories, there are two cases where O_2 are/is *ʔs- and *ŋg-. The structures of these categories are illustrated in the following diagrams:

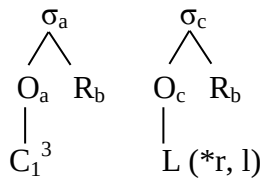


Figure 5.2 The structure of Diéyùn words of $O_aR_bLR_b$ pattern

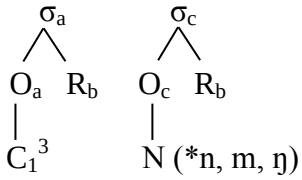


Figure 5.3 The structure of Diéyùn words of $O_aR_bNR_b$ pattern

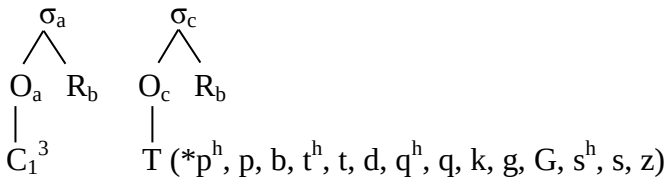


Figure 5.4 The structure of Di éyùn words of $O_aR_bTR_b$ pattern

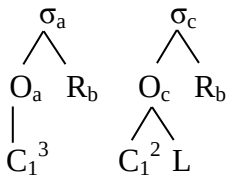


Figure 5.5 The structure of Di éyùn words of $O_aR_bC_1^2LR_b$ pattern

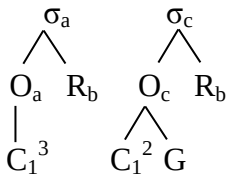


Figure 5.6 The structure of Di éyùn words of $O_aR_bC_1^2GR_b$ pattern

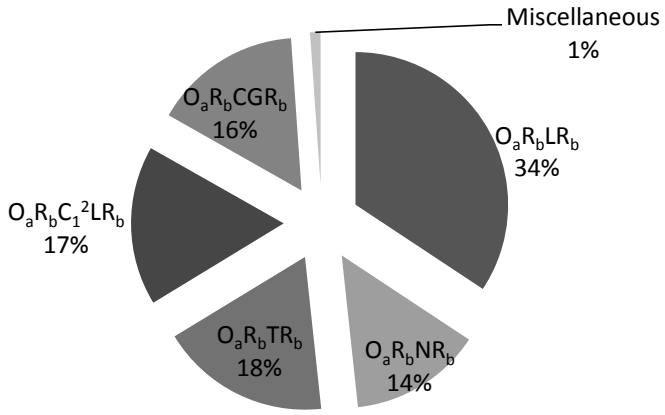


Figure 5.7 Percentage of different sound patterns of Di éyùn words

Similar to the explanation of Splitting-sound pattern described in Chapter Four, Sun does not provide a justified explanation for the Di éyùn pattern either, especially for the occurrence of Liquids on the second syllable onset and the relationship between the Liquids and other second

syllable onset variants. According to the Reduplication theory, Di éyùn is a partial reduplication, developed from total reduplication. The rhyme of the reduplicant could be explained as a total copy of the base rhyme, but the reduplicant onset cannot be interpreted as a product of reduplication. Sun (1999) attempted to use the One Syllable/One Meaning Principle (OOP) and the Sonority Sequencing Principle (SSP) to account for the occurrence of Liquids in the second syllable onset. He argues that the liquids are the most sonorous consonants, able to raise the sonority level of the second syllable onset. The rise of the sonority level between the sonority peaks of the two rhymes obscures the distinction between the two syllables. The two syllables will sound more like one syllable. The OOP constraint won't be violated. However, OOP and SSP could only be applied to explain the $O_aR_bLR_b$ pattern. Within this reduplication framework, other Di éyùn patterns, such as $O_aR_bNR_b$, $O_aR_bTR_b$, $O_aR_bC_1{}^2LR_b$ and $O_aR_bC_1{}^2GR_b$, cannot be given a reasonable explanation. For example, in the $O_aR_bTR_b$ sound pattern, the occurrence of obstruents cannot raise the sonority level of the second syllable onset, since the sonority level of obstruents is the lowest among all the segments.

Since Sun's synchronic reduplication approach fails to provide a reasonable account for Di éyùn variation, I present an alternative diachronic proposal. I suggest that the rise of Di éyùn be the product of both partial reduplication and the development of a disyllabic template ($\sigma_a\sigma_b$) that is generated from an earlier Splitting-sound pattern. The formation of the template is summarized in figure 5.8:

$O_aR_bLR_c$	$O_aR_bLR_b$	$O_aR_b\{L,N\}R_b$	$O_aR_bO_c\{L, N, T\}\ R_b$
Splitting-sound pattern	1st stage	2nd stage	3rd stage

Figure 5.8 The evolution of the disyllabic template ($\sigma_a\sigma_b$)

Figure 5.8 shows the evolution of the template across time, due to different generalizations by language learners. A conclusion can be drawn after a close examination of Di éyùn patterns that the disyllabic $\sigma_a\sigma_b$ template was not formed in one step. The whole process starts with the pattern of Splitting-sound words, undergoing three stages. In the first stage, the pattern - $O_aR_bLR_c$ was integrated with a reduplication pattern. That is to say, the Splitting-sound pattern was interrupted by the plain reduplication process. In this course, the sound pattern was changed into $O_aR_bLR_b$, with the first rhyme reduplicated in the second syllable. After the first stage, the pattern experienced a series of generalizations until it was transformed into a stable template: First, the feature specifications of the second syllable onset extended from [+cons, +son, -nasal] to [+cons, +son], so the consonants for the second onset were extended to all liquids and nasals *l-, *r-, *n-, *m-, *ŋ-. The output of this generalization is displayed in the second stage in the above Figure 5.3; the final step was the greatest generalization - the second onset consonants extended from [+cons, +son] to [+cons], arriving at the third stage $O_aR_bO_c\{L, N, T\}R_b$. The third stage symbolizes the formation of the template $\sigma_a\sigma_b$, because the consonants of the second syllable onset in this stage cover all the consonants in the OC consonant inventory. The occurrence of consonant clusters on the second syllable onset in the patterns of $O_aR_bC_1^2LR_b$ and $O_aR_bC_1^2GR_b$ symbolizes the completion of formation of the $\sigma_a\sigma_b$ template. The $O_aR_bC_1^2LR_b$ and $O_aR_bC_1^2GR_b$ patterns, just as the Splitting-sound variations, also reflect that the sound change of CL > C was gradual and in progress. As the simplification of CL clusters came to an end in EMC, all the complex onsets in Di éyùn words were simplified into single-consonant onsets so that the Di éyùn pattern was fossilized into the disyllabic template: $O_aR_bO_cR_b$.

5.3.2 Origin and developmental direction of $\sigma_a\sigma_b$

Pan (1999) is the first historical phonologist who related the origin of Lianmian words directly to historical phonological developments. Pan concluded that the simplification of disyllables into sesqui-syllables is the origin of Dié yùn, and that the Dié yùn pattern generates the pattern of Splitting-sound words. This conclusion is built on the assumptions that Proto-Chinese is also a language with many disyllables (mono-free-morphemic disyllables) just like neighboring languages, and that the disyllables had developed into sesqui-syllables before they changed into monosyllables. He cited a Splitting-sound example - *talau* from the Modern Fuzhou dialect of North Min: the reconstructed form of *talau* should be **taulau*, a Dié yùn word (Pan, 1999, p. 143). He claimed that the Dié yùn **taulau* had developed from a disyllable through vowel neutralization, and then developed into *talau* after the first syllable vowel 'au' underwent monophthongization (Pan, 1999, p. 143). Pan also argued that only this developmental pathway can explain why the vowel in the first syllable of the Splitting-sound words always corresponds to the main vowel of the second syllable.

My hypothesis departs from Pan's discussion in two aspects: in the origin of the disyllabic template and in its developmental direction. The origin of disyllabicity relies on the analysis of the following two issues.

First, I argue that Splitting-sound words, not Dié yùn, are the earliest disyllabic words in OC. That is to say, disyllabicity starts from the rise of Splitting-sound words. Pan hypothesizes that Proto-Chinese is full of disyllables. Therefore, in his theory, it was mono-syllabicity rather than disyllabicity that featured the lexical development from Proto-Chinese to OC. This bold hypothesis is controversial because of a scarcity of written documents.

Second, according to Pan, the motivation for the transition from a disyllable to a sesqui-syllable is the stress shift from the first syllable to the second one. If so, the stress shift should also be reflected in the Diéyùn pattern. The pattern contains no evidence for this argument. A counter-example may be found in a special Diéyùn word: *t^h*uuw ruw* 绸缪 chóu móu 'to twine' (from *Shījīng*). In this word, the first syllable has a long vowel while the second has a short vowel, although their main vowels and codas are the same. The longer vowel duration in the first syllable suggests that it is not an unstressed syllable.

It is also hard to accept Pan's argument that Diéyùn could be a product of neutralization for a disyllable in Proto-Chinese. He assumed the change **taulau* < *tVVlau was a neutralization. Pan did not point out that the reconstructed **taulau* was an OC or Middle Chinese form. As mentioned in Section 2.2, according to both Zhengzhang and Pan's reconstructions, no diphthongs exist in OC (Pan, 2000; Zhengzhang, 2003). If it is an OC sound, **taulau* is in contradiction with his own reconstruction. If **taulau* is a Middle Chinese sound, it is impossible to explain why the Proto-Chinese form *tVVlau could be well preserved throughout the OC period until Middle Chinese. I argue that it is better to interpret this double-rhyme phenomenon in Diéyùn words as a reduplication process rather than one of neutralization.

The origin of Diéyùn cannot be traced back to so-called 'disyllabic' Proto-Chinese. To solve the problem of the developmental direction, it is necessary to consider whether Diéyùn generates Splitting-sound patterns or vice versa. Although $O_aR_bO_cR_b$ is a widely-accepted Diéyùn pattern among native speakers, the second onset of most Diéyùn is *L- (Sun, 1999). That is to say, the most common Diéyùn pattern is $O_aR_bLR_b$, similar to the Splitting-sound pattern OR_aLR_b. The important difference between Diéyùn and Splitting-sound patterns lies in their rhymes rather than their onsets.

To investigate this argument, let's examine the rationality of Pan's assumption that after the first rhyme went through monophthongization, Dié yùn turned into Splitting-sound words. Pan considered the Splitting-sound word *talau* in Modern Fuzhou dialect to be developed from Dié yùn word **taulau* after the first syllable went through monophthongization. According to his reconstruction and diachronic presumption, Dié yùn words are transitional forms in the evolution from Proto-Chinese disyllables to Splitting-sound words. All Dié yùn words in this dialect should have evolved into Splitting-sound words. However, this is not the case. In Fuzhou dialect, both Dié yùn and Splitting-sound words are still found. In addition, Pan's assumption is based solely on Modern Fuzhou dialect and cannot cover the Splitting-sound examples in other dialects, nor in Middle Chinese or OC. If we accept this developmental route, it would be impossible to explain why there were so many Splitting-sound words in OC in which diphthongs are still preserved in the first rhymes. Pan's explanation fails to interpret how the coda in the first syllable of Dié yùn dissimilated from the one in the second syllable in the transition from Dié yùn into Splitting-sound words, while the vowels did not change at all. My analysis of Splitting-sound patterns shows that the first rhyme of most Splitting-sound words only differs from the second in their coda part, and not in their vowels.

I propose that the developmental direction is from Splitting-sound pattern to Dié yùn pattern. The disyllabic template $\sigma_a\sigma_b$ generated by Splitting-sound pattern was firstly interrupted by the reduplication process. On this stage, $O_aR_bLR_b$ emerged as a result. The $O_aR_bLR_b$ pattern is apparently simpler than Splitting-sound pattern OR_aLR_b. It is more reasonable to believe that $O_aR_bLR_b$ is a late form developing from OR_aLR_b. Moreover, the transition on the later stage from $O_aR_bLR_b$ to $O_aR_bO_cR_b$ could naturally be interpreted as a generalizing process, for the 'disyllabic template' did not come into being until this point. Once the 'disyllabic template' was built, the

words in this template could be more easily preserved. This may explain why Modern Chinese vocabulary preserves more Diéyùn words than Splitting-sound words, though both of them had lost productivity around LOC.

A morphological analysis of Diéyùn can also help us understand the direction problem. In Diéyùn forms, we find some words that could hardly be called mono-morphemic, because at least one of the two Diéyùn syllables can be a bound morpheme. They are termed Diéyùn compounds, and I discuss them in detail in the next Chapter. In contrast, all Splitting-sound words are mono-morphemic. The developmental history of Disyllabicity tells us that compounds were later developed disyllables in Chinese. Therefore, we have reason to think of Diéyùn as a later form than Splitting-sound words.

Many Diéyùn words from *Shījīng* show that their semantic references are more subtle and richer compared with Splitting-sound words which have more basic meanings. As shown in Table 3.2, 43.03% of Diéyùn words in the data are adjectives, while only 25% of Splitting-sound words are adjectives. 62.50% of Splitting-sound words are nouns, used to describe animals or plants in nature. This also suggests that a Splitting-sound word was created slightly earlier than Diéyùn, as an original disyllabic form. In contrast, as Diéyùn became productive, it was used for other semantic purposes, e.g. as a descriptive or poetic device, including possible sound symbolism.

5.4 Summary

Built on previous synchronic and diachronic studies on Diéyùn words (Sun, 1999; Pan, 1999, 2003), my tentative conclusion is that phonological factors did not directly lead to the creation of the Diéyùn pattern. They are, instead, generated from the Splitting-sound pattern, and may well be the result of the interaction between morphology and the disyllabic $\sigma_a\sigma_b$ template.

The most frequent Diéyùn pattern, $O_aR_bLR_b$, inherited the Splitting-sound pattern OR_aLR_b, which preserved the second onset Liquid. After the two steps of generalization, the pattern was finally fossilized into a stable disyllabic template $O_aR_bO_cR_b$, which contributed to the preservation of most Diéyùn words throughout the long history of Chinese. Based on their doubled rhymes and later-developed doubled tones, Diéyùn words have been better preserved in Chinese poetic literatures, compared with Splitting-sound words. However, it was the generalization of the Splitting-sound pattern that resulted in the obscurity of the Liquid identity that had been lost in the simplification of CL onset cluster.

Although phonological factors influencing the rise of Splitting-sound words have faded away in the formation of the template, this disyllabic $\sigma_a\sigma_b$ template continues to play an important role in the development of disyllabicity. When the simplification of consonant-cluster onsets ended around LOC and EMC, the phonological motivations for Splitting-sound words disappeared. This did not deter the development of disyllabicity because of the formation of this stable template. The template became a new word formation tool, broke the morpheme boundary, and began to produce disyllabic compounds. It is this template that also determines that disyllabicity, rather than tri-syllabicity or multi-syllabicity should be the direction of Chinese lexical development.

CHAPTER 6

EXTENSION OF THE DISYLLABIC TEMPLATE:

FROM MONO-MORPHEMIC TO BI-MORPHEMIC WORDS

6.1 The characteristics and development of disyllabic compounds

The productivity of EOC Lianmian words, including Splitting-sound words and Di éyùn is reflected in some classical Chinese literary works, such as *Shījīng* 詩經 and *Chǔcí* 楚辭. After MOC, those disyllables gradually decreased in number in literature. Disyllabic compounds began to increase in LOC, while mono-morphemic disyllables (Lianmian words) were gradually losing their productivity (Xu, 2003). Before analyzing the creation of disyllabic compounds, it is necessary to clarify their characteristics and their developmental route.

6.1.1 Lianmian words in transition

Unlike phonologically strict Lianmian words, disyllabic compounds in (3) were made up of two morphemes. However, the transition from a phonologically strict word-formation process to a compounding process did not happen in one step. Cheng (1992), for example, states that there appeared to be a type of compound that was created through both reduplication and compounding. Examples of this type of word are given in (1), (2), and (3), based on Cheng (1992, p. 85-88).

(1) Total reduplicates in OC

Reconstructed OC	Character	Gloss	Meaning	Source
*kaaw kaaw	高高	high-high	'high'	*Shījīng, Zhōusòng*
*s^{h}uumʔ s^{h}uumʔ	慘慘	miserable-miserable	'miserable'	*Shījīng, Dàyǎ*
*s ɬk^{h}eeŋ s·k^{h}eeŋ	青青	green-green	'fresh'	*Shījīng, Wèifēng*

(2) Disyllabic Compounds with the same or similar onsets

Reconstructed OC	Character	Gloss	Meaning	Source
*loŋʔ lewg	踴躍	jump-up fast	‘jump up’	*Shījīng*
*p^hral p^hɯd	披拂	spread-touch lightly	‘wind-blowing’	*Zhuāngzǐ*
*kaaŋ kɯʔ	綱紀	Principle-law	‘law’	*Shījīng*

(3) Disyllabic Compounds with the same or similar rhymes

Reconstructed OC	Character	Gloss	Meaning	Source
*kwaaŋ mraŋ	光明	light-bright	‘bright’	*Shījīng*
*p^hraan ɢwans	攀援	draw-pull by hand	‘to climb’	*Shījīng, Sānguózhì*

In this type of Lianmian compound, there are some phonological connections between the two disyllables. However, they do not obey the reduplication patterns strictly. For instance, in (2), the two syllables of *p^hral p^hɯd* 披拂 do not have the same onsets: the first syllable has a consonant-cluster onset, while the second one just copies the stem of its onset. In (3), the two codas of * *phraan ɢwans* 攀援 are not exactly the same: one is *-n, the other is *-ns instead. This difference might reflect the fact that the two syllables had different tones[42]. Thus, we may assume that the tone in Dì éyùn of this period might not possibly be duplicated.

The most important criterion that is employed to determine whether a Lianmian word is a Dì éyùn compound or not, lies in its ability of being decomposed semantically. As shown in (2) and (3), most examples are not mono-morphemic any more: One or two syllables in these disyllables have semantic connections with that of the whole word; that is to say, at least one syllable could be used as an independent free morpheme. For instance, *mraŋ 明 ‘bright’ in

[42] Linguists who study the historical origin of tones (tonogenesis) in OC tell us that *píng* ‘level’ tone arose in OC in open syllables and syllables ending in sonorants (glides and nasals), and *shǎng* ‘rising’ tone in syllables ending in a glottal stop (Mei,1970). Some instances of *qù* ‘departing’ tone arose from a final *-s (Haudricourt, 1954). Of course, there is still a debate on whether OC is a tone language. Pulleyblank (1962) thinks that OC did not distinguish syllables for tone (see details in Table 4.1 of Chapter 4).

**kwaaŋ mraŋ* 光明 'bright' and **p^{h}raan* 攀 'to draw' in **p^{h}raan ɢwans* 攀援 'to climb' can both be used as free morphemes. Although these Diéyùn words still preserve the phonological connection between their two syllables, they are virtually Diéyùn compounds rather than Lianmian words. Compared with the mono-morphemic Diéyùn words discussed in Chapter Five, these Diéyùn compounds are better preserved in Modern Chinese, because they are composed of two morphemes, one of which, at least, has independent meaning, able to constitute other compounds. For example, **p^{h}raan* > **p^{h}an* 攀 'to draw' in Middle Chinese could combine in such compounds as **p^{h}an toŋ* 攀登 (to draw + to ascend) 'to climb' **p^{h}an ba* 攀爬 (to draw + to crawl) 'to climb'.

6.1.2 Types of compounds

The typical compounds in OC have one of three types of structures (Packard, 2000): parallel structure or coordinating structure — in which neither member dominates the other; non-parallel structure or subordinating structure — in which one member modifies and is subordinate to the other; and a "one-sided" compound — in which the entire compound takes on the meaning of only one of its members, effectively eliminating the meaning of the other member (Packard, 2000, p. 12; Cheng, 1992). The following tables (4) (5) and (6) illustrate the three types of compounds with examples based on Cheng (1981, p. 72). Translations are based on Packard (2000). Reconstructions are modified for consistency. Zhengzhang's system is used here.

(4) Parallel compounds

	Reconstructed OC	Compound	Gloss	Meaning
a.	*luuʔ gˤraags	道路	way-path	'road'
b.	*kuud njug	骨肉	bone-flesh	'kin'
c.	*gloolʔ puɢ	禍福	calamity-blessing	'fate'

(5) Subordinating compounds

	Reconstructed OC	Subordination relation	Compound	Gloss	Meaning
a.	*q^{h}l'iin sɯʔ	modifier-head	天子	heaven-son	‘emperor’
b.	*ʔsaŋ kun	verb-object	將軍	lead-troops	‘general’
c.	*slɯ noom	verb-object	司南	control-south	‘compass’

(6) “One-sided” compounds

	Reconstructed OC	Compound	Gloss	Meaning
a.	*kwɯɯg kraa	國家	(king or lord’s) feud- (high official’s) feud, home	‘state’
b.	*m ɾads gaads	厲害	benefit-harm	‘harm’

6.1.3 The characteristics of compounds

Two pertinent morphological characteristics of compounds in OC were selected from Cheng (1992)’s results.

First, these compounds fixed through syntactic structures had been derived from syntactic phrases. Cheng cites 32 compounds of parallel structures whose two syllables (AB) can be ordered either A+B or B+A, to indicate that compounds were derived from phrases and had been in an unstable state before being frozen into regular compounds (Cheng, 1992, p. 97-99). What’s more, all the syntactic structures of compounds can be found in those of phrases.

The other morphological feature pertains to the internal syntactic structure of compounds. The two common syntactic structures — coordinative structure and subordinate structure with modifier-head, are considered to be the earliest and the most popular structures of disyllabic compounds in OC (e.g. 4a and 5a). With the development of disyllabicity, the compounds of other structures became more and more popular after OC. In the wake of the increase of

disyllabic compounds and the enrichment of their structures, mono-morphemic Lianmian disyllables gradually decreased.

6.1.4 Extension of the disyllabic template

As onset simplification ran its course, Lianmian words, including Splitting-sound words and Di éyùn, lost their productivity before the disyllabic $\sigma_a\sigma_b$ template became the primary word-formation tool. Therefore, monosyllabic free morphemes had not been replaced by disyllabic Lianmian morphemes and were still in overwhelming majority in LOC. The preserved disyllabic template was gradually fed by free monosyllabic morphemes with independent meaning. During this transitional phase from mono-morphemic disyllables to disyllabic compounds, the $\sigma_{a+}\sigma_b$ template still retained its reduplicated-rhyme pattern, while containing a clear morpheme boundary. That is to say, the template produced Lianmian compounds — two syllables of which had some loose phonological connections. This diachronic analysis addresses the question raised by Packard (2000, p. 15): how it is possible for a word to appear to be at once both a compound and a reduplication? During that period, speakers of OC did indeed form words productively by joining free words that shared both semantic and phonological properties (Packard, 2000, p. 15).

Based on this change, the $\sigma_a\sigma_b$ template finally broke through the phonological constraints and changed into a free template producing bimorphemic compounds. This developmental route is illustrated in Figure 6.1:

Figure 6.1 Extension of disyllabic template

The final $\sigma_{a+}\sigma_b$ template was fed with free mono-syllabic morphemes, attracting disyllabic phrases. If we say the formation of the disyllabic $\sigma_a\sigma_b$ template reflected the operation of phonological processes upon disyllabicity, then the extension of the template from $\sigma_a\sigma_b$ to $\sigma_{a+}\sigma_b$ is the product of the disyllabic $\sigma_a\sigma_b$ template and lexicalization through compounding. This extension route is reflected in the examples of Table 6.1:

Table 6.1 Examples to show extension of the disyllabic template

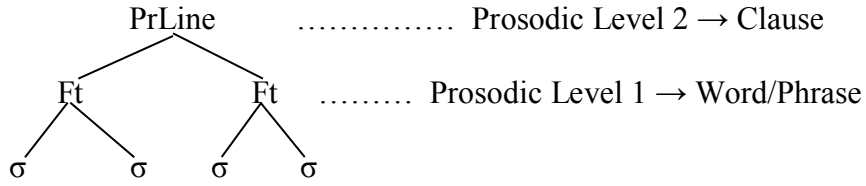
	Stage 1 (mono-morphemic Lianmian words)	Stage 2 (bi-morphemic Lianmian compounds)	Stage 3 (disyllabic bi-morphemic compounds)
Examples	*ʔiwʔ-liwʔ窈窕 'elegant' (Di éyùn) *ts^{h}rjum-ts^{h}rjaj 參差 'uneven' (Shuāngshēng)	*kwaaŋ mraŋ 光明 'bright' (Di éyùn compound) *p^{h}ral p^{h}ɯd 披拂 'swing in the wind' (Shuāngshēng compound)	*bood djeb 跋涉 'trudge' *ʔsaŋ kun 將軍 'general'
Semantic connections between the two σs	No semantic connections between them	Some semantic connections between them (parallel relation)	More rich semantic connections between them (parallel relation & subordinate relation)
Sound connections between the two σs	They have the same rhyme *-iwʔ or onset *ts^{h}r-	They have similar rhymes *-aaŋ vs. *-aŋ or onsets *p^{h}r- vs. *p^h-.	No sound connections

The syntactic phrases under the control of the disyllabic template gradually froze into disyllabic compounds. As a result, the disyllabic compounds became the major lexical forms and greatly enriched the vocabulary. Just as Anderson (1985), who explores ways in which compounds reflect the properties of phrasal constructions, says, compounding bridges morphology and

syntax. This function of compounding is very well reflected in Chinese history of lexical development. Syntactically created phrases were easily absorbed into the disyllabic $\sigma_{a+}\sigma_b$ template to form disyllabic compounds especially in LOC and EMC when free morphemes were mono-syllabic. This led to the two most salient structural properties of Chinese compounds: compounds are structurally very similar to syntactically-created phrases; and compounds and phrases often have the same sequence of formatives (Anderson, 1985).

In LOC and EMC the template seemed to be an important lexical formation tool producing myriad disyllabic compounds. Meanwhile, its prosodic function began to surface in poetry. In *Shījīng*, every line is composed of four syllables, which was usually divided into two disyllabic units – each equated with a foot. A prosodic line is usually a clause and the "foot" is always a word (either a mono-morphemic disyllable or a bi-morphemic compound) or phrase.

(7) Verses from *Shījīng*:



**quɯds* **njɯ* **pɯ* **keens*,
愛 而 不 见，
hide and not seen
'Loving and not seeing her'

**suu* **hljuʔ* **deg* **dog*
搔 首 蜘 躅
scratch head walk.back.and.forth
'I scratch my head and am in perplexity.'
(Phrase) (Lianmian Word) (from *J ǹgn ü* in Chapter *Bèifēng*)

**daads *pa* **bood *djeb*

大 夫 跋 涉

official trudge

'An official has gone through all the rivers,'

(Compound word) (Compound word)

**ŋaalʔ* **slɯm* **skɯɯg* **qu*

我 心 則 憂

my heart so worry

'My heart is full of sorrow.'

(Phrase) (Phrase) (from *Zǎichí* in Chapter *Yōngfēng*)

The above analysis of verses from *Shījīng* shows that there is no restriction on syntactic structures within this disyllabic unit. The unit could contain syllable boundaries, morpheme boundaries, or even word boundaries. $\sigma_{a+}\sigma_b$ is not simply a disyllabic template any longer, but a prosodic unit, leading to the incongruence between syntactic structure and prosodic structure. The poetic form of four-syllable-lines was only in use for a short period of time (around LOC), because the four-syllable-line constraint controlled by two prosodic levels was later broken by the increase of disyllabic compounds. Nevertheless, its use reflects the prosodic structure and great prosodic influence of the disyllabic template.

6.2 The influence on Modern Chinese

6.2.1 Disyllabicity in Modern Chinese

Liberated from phonological constraints, the template's great productivity led to the first increase of disyllables in LOC and EMC. This developmental feature was reflected in some statistics mentioned earlier in the Introduction: disyllables increased to 40%-44% in Middle Chinese literatures (Li, 2007) from 16%-28% in OC literatures (Qiu, 2010). Duanmu & Zhang's

(2010) statistics show that in the Modern Chinese lexicon, disyllables take up 72%. Therefore, it is possible to conclude that there was another dramatic rise in disyllabic compounds from Middle Chinese to Modern Chinese. Starting from Early Modern Chinese (Old Mandarin), compounding through the disyllabic prosodic template has become a very productive new-word formation tool. In addition, a large number of new disyllabic compounds formed through this template reinforced its lexicalization function.

Wedel (2004, 2007) provides models of how the lexicon changes over time, which could be applied to the study of the dramatic increase of disyllables from Middle Chinese to Modern Chinese. With the aid of computational simulations, Wedel proposes that the model of pattern change should be built on a simple loop between production and perception, and that feedback loops can operate in a rich-memory system to produce regularity (Wedel, 2007, p. 9). The outputs that the disyllabic template $\sigma_{a+}\sigma_b$ created reinforced learners' perception. And these previously experienced disyllabic compounds were favored by the speakers' biases, promoting the usage of the template as a word-formation tool. Free from system-external factors, the self-feeding cycle created by the perception and production loop resulted in the dominance of the disyllabic template in the Chinese lexicon.

The influence of the $\sigma_{a+}\sigma_b$ template in Modern Chinese may also be reflected in its control of loanword adaptation. For example, the loanword fǎlánxī 法蘭西 *falanɛi* was adapted phonologically from 'France' under the constraint of Chinese CV(N) syllable structure. However, the output had three syllables since the original mono-syllable input included one consonant-cluster onset /fʁ/ and one consonant-cluster coda /ns/, which was adapted into two syllables in Chinese through vowel insertion. This was the original phonological adaption. Then a new disyllabic compound fǎguó 法國 *fakuo* 'France' was created by combining the first syllable fǎ 法

of the phonologically-adapted form and a redundant syllable guó國 for ‘country’. This disyllabic compound is preferred in Modern Chinese, conforming to the general preference for disyllabic words. Many place names whose phonologically-adapted forms consist of more than two syllables are nativized in this way to fit the disyllabic template. This template corresponds to the disyllabic foot and accounts for the dramatic increase of disyllables in LMC.

The disyllabic template initiated from Splitting-sound patterns in OC determines the trend of disyllabicity in the lexicon. Once the disyllabic template freezes, the word-formation tool - compounding develops in the direction of disyllabicity rather than trisyllabicity or other multi-syllabicity. As mentioned in Chapter One, the long history of disyllabicity from OC to Modern Chinese went through two periods of rapid development: from LOC to EMC, and from LMC to Modern Chinese. I argue that the rise of disyllables in OC is motivated by attempts to preserve complex onsets in OC that were being simplified at that time, but the rapid development of disyllables from Middle Chinese to Modern Chinese is greatly influenced by functional considerations. Although the simplification of complex onsets from OC to Middle Chinese gave rise to Lianmian words, this sound change did not promote many homophonous syllables because the rich coda and vowel systems were still preserved in Middle Chinese. However, the dramatic sound mergers preceding Modern Chinese resulted in a large number of homophones. The sound simplification processes that happened around LMC and Old Mandarin are listed in (8) (Tang, 2011):

(8) LMC/Modern Mandarin sound change

a.	The simplification of Finals	the loss of *-p, -t, -k; the merge of final *-m and *-n into -n
b.	The simplification of Initials	the loss of voicing distinction in obstruents (all voiced obstruents became voiceless); the loss of *ŋ- and *ʔ-; the merge of apical sibilants and velars into alveolo-palatals before close front vowels (*k-, *ts- > tɕ-; *k^h-, *ts^h-, *g-, *dz- > $tɕ^h$ -; *ɦ-, *s- > ɕ - before j, ɥ).

Lü (1963, p. 440) once discussed the increase of disyllables in Modern Chinese, "Why is there a strong tendency for disyllabic words in Modern Chinese? The large number of homophones should be an important factor". Examples to show the relationship of sound mergers to homophones in Modern Mandarin are displayed in Table 6.2 (The Middle Chinese reconstruction refers to Baxter, 1992):

Table 6.2 Homophony caused by phonological mergers from Middle Chinese to Modern Mandarin

Middle Chinese	Homophones in Modern Mandarin	Phonological factors
*san 山 'mountain' *sam 衫 'gown'	ʂan 山 shān ʂan 衫 shān	the merge of final *-m and *-n into -n
*bit 弼 'to assist' *bik 愎 'obstinate'	pi 弼 bì pi 愎 bì	the loss of *-p, -t, -k
*gjwaŋ 狂 'crazy' *kjwaŋ誑 'to cheat' *bæj 敗 'to defeat' *pɛj 拜 'obeisance'	kʰwaŋ 狂 kuáng kʰwaŋ 诳 kuáng paj 敗 bài paj 拜 bài	the loss of voicing distinction in obstruents
*ʔan 案 'long, wooden table' *ŋan 岸 'bank'	an 案 àn an 岸 àn	the loss of *ŋ- and *ʔ-
*gjiew 翹 'to elevate' *dzjew 誚 'to ridicule'	tɕʰjaw 翹 qiào tɕʰjaw 誚 qiào	the merge of apical sibilants and velars into alveolo-palatals before close front vowels

According to Duanmu's statistics (1999b, p. 10), there are over 3000 distinct syllables (including tonal contrasts) in Middle Chinese, but Modern Chinese has just over 1300. The great reduction of syllabic contrast has led to a great deal of homophony: on average, each syllable in Modern Chinese corresponds to 5.4 morphemes (Duanmu, 1999b). Therefore, in order to avoid ambiguity, Chinese speakers make use of a large number of disyllabic compounds under the constraint of the disyllabic template.

Here I need to clarify my position on the influences of the disyllabic template and homophone-avoidance on disyllabicity and their relationship. Homophone-avoidance has never been a primary factor leading to disyllabicity. Disyllabicity has always been a consequence of

the disyllabic template. The disyllabic template originated in the earliest Lianmian forms as a means of emphasizing complex onsets that were being lost; later it acted as a system-internal attractor for disyllabic compounds, (Blust, 2007; Wedel, 2007) especially in the period of LMC and Old Mandarin. Homophone-avoidance functioned as an external boost accelerating this lexical change during the period following the sound change mergers shown in (8). There is no evidence of using disyllabic compounding as a tool to eliminate the ambiguity caused by homophones. The coinage of new disyllabic compounds in Modern Chinese truncated from phrases by the disyllabic prosodic template does not result in the reduction of homophones, but to some extent aggravates homophony in the Modern Mandarin lexicon instead. Chinese speakers usually derive the meanings from collocation and context. Examples are shown below in (9):

(9) A homophone example in Modern Mandarin lexicon

Disyllabic Compounds (truncated)	代孕 'surrogate (mother)'	待運 'empty (taxi)'
Pronunciation (in Pinyin)	dài yùn	dài yùn
Original phrases	代替 怀孕	等待 運載
	dài tì huái yùn	děng dài yùn zǎi
	substitute pregnancy	wait delivery
	'to surrogate somebody to be pregnant'	'to wait for delivery'

6.2.2 The influence on Modern Chinese morphology and prosody

The disyllabicity instigated by phonological factors greatly influenced Modern Chinese morphology and prosody. Three language-particular features are generated from this template.

First, some of the disyllabic compounds are not completely frozen in the Modern Chinese lexicon, so the boundary between disyllabic phrases and disyllabic compounds becomes less

distinct. The compounding outputs have conventionalized meanings that are no longer compositionally derived. From this point of view, these forms seem to satisfy the requirement of grammatical words. However, the forms fail to meet other conventional criteria of grammatical wordhood as terminal nodes in the syntax, such as non-interruptibility and ordering constraints, since they are diachronically derived from phrases. Examples from (10) and (11) illustrate these two features of Modern Chinese words:

(10) Interruptibility of Modern Chinese words:

dān-xīn 擔心‘ to worry’ vs. *dān shénme xīn* 擔什麼心‘don’t worry’

(*dān* 擔 ‘to undertake’ + *xīn* 心 ‘heart’; *shénme* 什麼 ‘what’)

(11) Variable order in Modern Chinese words:

qiān-guà 牽挂 vs. *guà-qiān* 挂牽 ‘to miss’ (*qiān* 牽 ‘to pull’ + *guà* 挂 ‘to hang’)

bēi-shāng 悲傷 vs. *shāng-bēi* 傷悲‘sorrow’ (*bēi* 悲 ‘grief’ + *shāng* 傷‘wound’)

In (10), *dān-xīn* is a disyllabic compound, with a verb-object structure. It is considered a compound because it can be used as a transitive verb, followed by an object. For example, *wǒ dān-xīn tā* 我擔心他‘I worry about him’ or *wǒ dān-xīn zhè jiàn shì* 我擔心這件事 ‘I worry about this thing.’ Sometimes, however, progressive or perfective aspect markers (*zhe* 著, *guò*過) can be inserted between the two syllables *dān* and *xīn*. In this case, *dān* 擔 could be considered a verb. In (11), a WH-word *shénme* 什麼 ‘what’ is inserted into the compound to express a negative meaning. Since *xīn* 心 is a free morpheme, and WH-word *shénme* 什麼 is often found inside [V N] phrases, the insertion of WH-word *shénme* 什麼 changes *xīn* 心 into an object of *dān* 擔. Thus, the interruptibility feature that complicates the syntactic wordhood in Modern Chinese reflects the special lexicalization route of Chinese: disyllabic phrases under the control

of disyllabic template were frozen into disyllabic compounds which sometimes still function as phrases.

Standard definitions of grammatical words include a constraint against free morpheme permutation (Bickel & Nichols, 2007; Dixon & Aikhenvald, 2002). The example in (11) indicates that in many parallel compounds, the two syllables can be permuted. This kind of permutation does not happen arbitrarily but purposefully to meet the rhyming requirement. The two lexical phenomena are more clearly understood when considered diachronically. These compounds are not the result of compounding but of disyllabicity. In other words, what allows this flexibility in the Modern Chinese lexicon is that these words were created under the disyllabic prosodic template, not motivated by any phonological factor directly, and that monosyllabic morphemes within the compounds are still semantically well-preserved.

The second language-particular feature brought by the disyllabic template to the Chinese lexicon is the Dual-vocabulary phenomenon (Duanmu & Zhang, 2010). According to Duanmu and Zhang, 40% of the modern Chinese words have both a long (disyllabic) and short (monosyllabic) form. For example, 'country' in Modern Chinese has two forms — a long disyllabic one: *guó-jiā* 國家, and a short monosyllabic one: *guó*國. *guó jiā* 國家 is a compound or pseudo-compound, and *guó*國 is a bound morpheme. However, when they are used in a phrase, the adopted form depends on the syllable numbers of the phrase element close to them. There is a strong preference for the number of syllables to conform to the disyllabic template.

(12) Dual-vocabulary phenomenon[43]:

[43] The asterisk signifies ungrammaticality; the question mark means limited or awkward use.

<xiǎo **guó**>	? [xiǎo <**guó-jia**>]	‘small/weak country’
小 國	?小 國家	
small country	small country	
[<ruò-xiǎo> <**guó-jia**>]	×[<ruò-xiǎo> **guó**]	‘weak country’
弱小 國家	× 弱小 國	
weak small country	weak small country	
<**guó** nàn>	× [<**guó-jiā**> nàn]	‘national calamity’
國 難	× 國家 難	
country calamity	country calamity	
[<**guó-jiā**> <zāi-nàn>]	× [**guó** <zāi-nàn>]	‘national calamity’
國家 災難	× 國 災難	
country calamity	country calamity	

From the above examples, we notice that the disyllabic template determines the choice between the long and short forms. In Duanmu’s words, it is Foot Binarity that controls the choice (A foot has two syllables) (Duanmu, 1999a, p. 34). A disyllabic compound is a preferred form when combined with another disyllable to create a prosodic phrase. A monosyllabic form is also preferred when building a disyllabic foot with another monosyllabic form. From the diachronic view, the incompleteness of disyllabicity results in the Dual-vocabulary phenomenon: many disyllabic compounds have been created, while the old mono-syllabic morpheme did not die out. They remain as bound morphemes to satisfy the need of disyllabic prosody and word-formation.

The third feature generated by the disyllabic template regards the influence it has on Modern Chinese prosody. In Modern Mandarin, the disyllabic unit is a primary prosodic unit. Since any monosyllabic morpheme can be fed into this disyllabic ‘machine’, the disyllabic template became a prosodic unit right after the morphemic boundary was broken around LOC

and EMC. Although the debate among Chinese linguists on whether this template should be considered as a disyllabic foot (Duanmu, 1999a; Shih, 1997) or a disyllabic phonological word (Feng, 2001; Yip, 1992) in prosodic structure is ongoing, it is agreed that this disyllabic unit can override syntactic junctures and can be applied to any level above syllables in prosody.

Some of the strongest evidence for this disyllabic constituent comes from tonal phonology. Shih (1997) adopted Third Tone Sandhi in Modern Mandarin to test the function of this disyllabic unit in a prosodic structure. Mandarin is a tone language, where each lexical word has its tone. The tonal inventory of Mandarin consists of four lexical tones, which are summarized in different representations in the following tone table:

Table 6.3 Four tones in Modern Mandarin

Tone	Pitch contour	Tone Markings in Pinyin	Chao's digits[44]	IPA Diacritic
Tone 1	High	–	55	/
Tone 2	Rising	/	35	v
Tone 3	Low	v	214	\
Tone 4	Falling	\	51	^

When two or more syllables occur next to each other, their tones tend to undergo change or Tone Sandhi (TS). 'Third Tone Sandhi' refers to the most frequent TS in Mandarin – Tone 3 Sandhi (T3S). The rule stipulates that a third tone T3 becomes a second tone T2, when it precedes another T3. Prosody plays a more important role in the domain formation, since T3S is very sensitive to prosodic structures. The following example shows that morphemes with different syntactic structures have the same T3S domains.

[44] Chao (1930) introduced a five-level system for the phonetic description of tones. On this scale, 5 represents the highest tone level, and 1, the lowest, and each tone is represented by a starting pitch and a final pitch, and sometimes a mid pitch. Except for the High tone which is level, the other three tones have changes in pitch and hence are referred to as contour tones. Mandarin thus makes use of a contour tone system, where tones are distinguished by their pitch trajectories, i.e., level, rising, falling.

(13) An example to show the insensitivity of T3S to syntactic structures

Word mǎ yǐ 螞蟻 T3 T3 → [T2] T3

'ants'

Compound mǐ-jiǔ 米酒 T3 T3 → [T2] T3

rice-wine

'alcohol made from rice'

Phrase nǐ hǎo 你好 T3 T3 → [T2] T3

you good

'How are you?'

Although the T3S domains shown in (13) have different syntactic structures, the domains are all disyllabic. In the analysis of T3S domain, Shih (1986) is the first one to develop a Mandarin Foot theory. The units of prosodic structure in Mandarin, proposed by Shih (1986) are disyllabic feet, super feet and phrases.

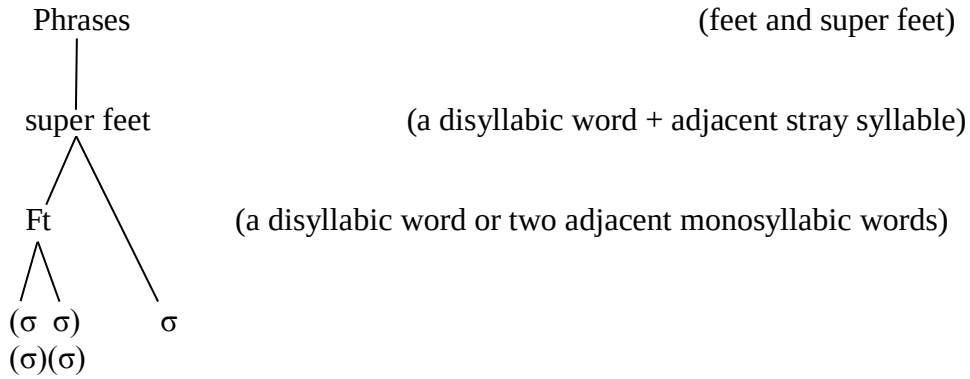


Figure 6.3 Prosodic structure of Modern Mandarin proposed by Shih (1986)

Within one disyllabic foot, the T3S applies obligatorily. When two T3S are grouped into separate prosodic units, a prosodic boundary occurs. And the T3S rule is optionally applied. Chen (2000) and Duanmu (1999a) later adapted this structure greatly based on the analysis of TS of other

Chinese dialects or languages (such as Shanghai dialect of Wu [wuu]), but the disyllabic unit remains the core in their accounts.

To summarize, I argue that with the completion of onset simplification in EMC, the phonological identity between the two rhymes of a Dié yùn word was not put in priority in the disyllabic template. Lianmian compounds were produced in the transition from mono-morphemic disyllables to bi-morphemic disyllables. When the template $\sigma_a\sigma_b$ extended to $\sigma_{a+}\sigma_b$, a more powerful word-formation tool was created. The template $\sigma_{a+}\sigma_b$ attracted disyllabic phrases or absorbed free monosyllabic morphemes directly from the EMC lexicon to coin new disyllabic compounds. Both system-internal and external factors contribute to the dramatic increase of disyllabic compounds from LMC to Modern Chinese. The disyllables that accounted for more than 40% of the MC lexicon (Li, 2007) created positive feedback, which served as a source of regularity over many cycles of production and perception (Wedel, 2007, p. 1). This system-internal attractor led to the steady and vigorous growth of disyllabic compounds in Modern Chinese, from Old Mandarin to Modern Mandarin. I argue that the disyllabic foot and homophone-avoidance functioned as two main system-external factors, reinforcing disyllabic compounding as the primary word-formation process.

This disyllabic template, as a product of the phonological change in OC, has played an increasing role in both Chinese lexicalization and Modern Chinese prosody. I argue that the template is responsible for ttwo language-specific features in Modern Chinese morphology: dual-vocabulary phenomenon (Duanmu & Zhang, 2010) and the fuzzy boundary between morpheme and word (Hoosain, 1992; Packard, 2000). The role of the disyllabic template in Modern Chinese prosody is reflected in the sensitivity of tone sandhi rules to the disyllabic foot domains across Modern Chinese languages and dialects.

In sum, the study of the rise of the disyllables in Old Chinese, from the extension of the disyllabic template from the earliest disyllabic pattern to the template's development throughout the history of disyllabicity, plays a vital role in better understanding Modern Chinese morphology and prosody.

CHAPTER 7

ALTERNATIVE ACCOUNTS FOR THE RISE OF DISYLLABLES IN CHINESE

The general view in Chinese linguistics is that the change from monosyllabic to disyllabic words has been a strong trend throughout Chinese history (Wang, 1980). 'Disyllabicity', as an interface between morphology and phonology, has aroused the interests of many researchers. Chinese morphologists and phonologists have sought to discover the origin of disyllabicity and the motivations for the tendency toward disyllables.

Regarding these issues, various theories have been proposed in the literature. In this chapter, I will review the two major types of proposals and compare them with my approach detailed in the previous chapters.

7.1 Functional approaches

In studies of disyllabicity, a conventional proposal is that disyllables were innovated to meet communication requirements caused by social development. Following this functional explanation, many linguists have attempted to explore disyllabicity more deeply from the pragmatic viewpoint (Chen, 1978; Cheng, 1992; Pi, 1992; Packard, 2000) as well as the phonological perspective (Lin, 2001; Shi, 2002; Wang, 1998).

7.1.1 Pragmatic view

Proponents of the pragmatic view assume that in the transition from a slave-based society to a feudal one (between 1000 BC and 300 BC), with the development of Chinese metallurgy and culture[45], there was a pragmatic need to create new words for new referents, which included more complicated tool names, refined descriptions of nature, and more subtle feelings (Cheng, 1992, p. 44; Packard, 2000, p. 265-266). If this need had only depended on the creation of

[45] This period corresponds to Zhou Dynasty (1046 BC – 256 BC), when feudalism and Confucian bureaucracy started to rise and the origins of native Chinese philosophy developed (e.g. Confucianism, Taoism, Mohism and Legalism) (Gernet, 1996).

monosyllabic words, large amounts of homophones would have entered the vocabulary and caused ambiguity in speech (Chen, 1978). Therefore, disyllables were created to disambiguate homophones.

Although it might look plausible a-priori, this conventional explanation has its drawbacks. Just as Duanmu argues, "there is no proof of any disyllabic word that has been introduced and then lexicalized because of homonym-avoidance" (Duanmu, 1999b, p. 11). As introduced in Chapter 2, around EOC and MOC when Lianmian words began to rise, OC was rich in consonant inventory and full of onset clusters and coda clusters. That is to say, OC of that period had plenty of distinctive syllables to avoid homophones even if the social development cried out for the creation of large amounts of new words. In this context before the great reduction of syllable types, homophone-avoidance cannot be the reason for the rise of disyllables in OC, because the phonological basis for the creation of homophones simply did not exist.

Furthermore, this explanation cannot answer the questions raised in the introduction: why doesn't homophony-avoidance lead to trisyllabicity or polysyllabicity, the innovation and productive use of affixes, or other word-formation processes? How can the unique sound patterns of Lianmian words (esp. Splitting-sound words and Di éyùn) be explained on the basis of vocabulary expansion alone, without taking into account phonological factors?

This view is also unable to explain why disyllabic compounds did not increase dramatically when social transition presumably caused the most homophone pressure for vocabulary innovation in the OC period. The greatest increase in disyllables occured after the cataclysmic social changes, in LMC and Modern Chinese, a fact which is unexpected in the functional pragmatic model.

7.1.2 Phonological view

Guo (1938), Wang (1944), Karlgren (1949), L ü (1963), Li and Thompson (1981) have acknowledged that the simplification of OC phonology and disyllabicity are two remarkable features in OC development. If there is a causal relation between the two features, in which direction did it occur? (Packard, 2000: 6) Proponents of the phonological view claim that the simplification of the phonological system initiated, rather than resulted from, the development of disyllabic words (Karlgren, 1923, p. 22-23; Li & Thompson, 1981, p. 14; Norman, 1988, p. 86-87; Wang, 1980, p. 342). Syllable structure simplification occurred as a natural phonetic weakening process. The simplification caused syllables which had been phonologically distinct to become similar or homophonous. Disyllabic words emerge consequently within the language (via reduplication or compounding processes) as a means of overcoming problems in communication caused by this proliferation of homophonous mono-syllabic words (Packard, 2000).

This phonological view provides a phonological explanation for the homophony pressure, drawing researchers' attention to historical sound change in OC phonology. Unlike the previous pragmatic view of disyllabicity as the result of the pressure brought by social changes to OC language, this view claims that language-internal phonetic changes and accompanying neutralization of former contrasts are primary motivations for the rise of disyllabic words in the language.

However, this 'phonological view' still includes the functional principle of homophony-avoidance and has its drawbacks. First, proponents of the 'phonological view' do not provide a detailed analysis for how syllable structure simplification caused the rise of disyllables, which part of syllable structure underwent what kind of simplification, and how this led to homophony. Without these phonological details, this view ultimately falls into the homophone-avoidance

explanation again, because without the analysis of syllable structure simplification and disyllables' phonological patterns, the hypothesis cannot relate phonological change closely to the rise of disyllables. This view also fails to demonstrate homophones in the OC lexicon that were caused by the simplification of OC phonology. Actually, the phonological system did not change very much in OC. At least the simplification of OC phonology was not powerful enough to cause many homophones in the OC lexicon. The big simplification process that caused syllables to become homophones started from EMC, while the rise of disyllables happened in EOC. Therefore, the homophone-avoidance component of the 'phonological view' cannot be justified.

Secondly, proponents of this phonological view, including Kalgren (1923), Li & Thompson (1981) and Wang (1980), fail to provide a reasonable explanation for why mono-morphemic disyllables (Lianmian words) arose first in EOC, while disyllables formed through compounding processes emerged later in MOC when Lianmian words began to lose productivity. If both reduplication and compounding processes are considered simply as a means of avoiding homophones, why did the two processes not occur at the same period of time and maintain their developmental momentum throughout the entire history of disyllabicity?

These problems lead one to search for other explanations for the rise of disyllables in the history of Chinese.

7.2 Prosodic approach

A second approach to disyllabicity is brought up by Feng, Shengli (1998, 2001), one of the first linguists to propose a detailed purely phonological motivation for disyllabicity and to arrive at a prosodic explanation. Feng only discusses disyllabic compounds which are considered the products of prosody and morpho-syntax, and does not include Lianmian words (mono-

morphemic disyllables). Feng's proposal has attracted the attention of morphologists like Packard (1998, 2000) and Acordia (2007).

Feng's proposal is built on the analysis of syllable structure simplification in OC. He uses the results of research done by Ting (1979, p. 717-736) and Yu, N. (1985, p. 290) to show how syllable structure became simplified in the history of Chinese. Based on Li's reconstruction, Feng shows the maximal and minimal syllable structures in OC, Middle Chinese and Modern Chinese in the following table:

Table 7.1 Syllable structure simplification in the history of Chinese

	Maximal syllable structure	Minimal syllable structure
OC	CCCMVCCC	CVC[46]
Middle Chinese	{C, S} V {C, S}	CV
Modern Chinese	{C, S} V{N, S}	V

In Li's reconstruction, the minimal syllable was CVC and the maximal one was CCCMVCCC, with both complex onsets and codas. In Middle Chinese, no consonant clusters were allowed in the coda; only two types of final consonants, nasals (*n, m, ŋ) and stops (*p, t, k) can appear in the coda. The final stop consonants did not survive in Old Mandarin (Acordia, 2007, p. 84; Feng, 1998, p. 224). In Modern Mandarin, the minimal syllable can be a vowel, which may be preceded by a semivowel and/or a single consonant; no consonant clusters are allowed, and the only possible consonant codas are /n/ and /ŋ/.

[46] Feng adopts Li's (1980, p. 33) reconstruction which assumes that there are no open syllables in OC, although this assumption is still controversial. Some counter-examples are cited in the discussion that follows.

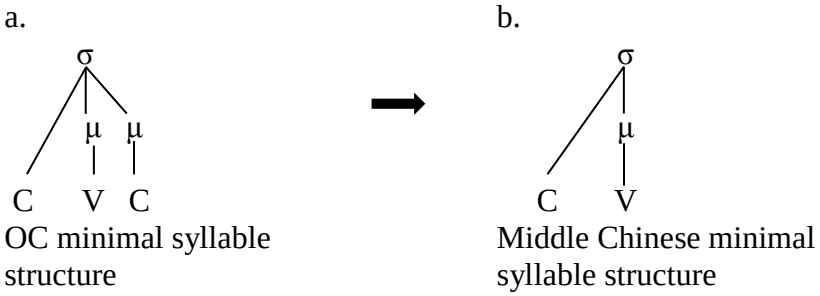


Figure 7.1 The change of minimal syllable structure from OC to Middle Chinese (Feng, 1998, p. 229-235)

Feng (2001) argues that in Chinese, prosodic word is the minimal independent unit of prosody, realized by the foot. It is given as a rule that a foot must be binary — either under syllabic or moraic analysis (Acordia, 2007, p. 84; Feng, 1998, p. 227; McCarthy and Prince, 1993, p. 43). In the transition from OC to Middle Chinese, the minimal syllable structure becomes CV, having only one mora, and the bimoraic foot was not always satisfied by a single syllable. Words containing only one minimal syllable could not constitute a foot any longer, and an additional syllable or mora needed to be added to make a new disyllabic foot (Feng, 1998, p. 228, quoted in Acordia, 2007, p. 85).

This new disyllabic foot coincides with what Feng calls a "two-word prosodic combination"; foot and phrase eventually merge. If the two elements in a disyllabic prosodic word or a phrase at this stage are frequently used, their relative order may become frozen, becoming an 'idiomatized Prosodic Word' (Feng, 1998, p. 236); the idiomatized Prosodic Word may in turn evolve into a compound through lexicalization. This process is summarized by Feng as:

phonological change → disyllabic feet → disyllabic phrases → idiomatized Prosodic Words → compounds

Feng's prosodic approach connects phonological change with the lexicalization of disyllables, and addresses the problem of why and how in Middle Chinese and Modern Chinese disyllabicity speeds up. Moreover, the prosodic structures preserved in Modern Chinese support a prosodic explanation. However, there are some serious problems with Feng's discussion of the prosodic changes generated by phonological simplifications. The big change in phonology from OC to Middle Chinese that Feng uses to motivate the evolution of disyllables does not occur until EMC. This, then, cannot be adopted to explain the rise of disyllables in OC. Through the whole OC period, syllable structure — especially the coda, does not change greatly. Even in EMC, many syllables still preserve plosive codas. The minimal syllable structure did not change much either. In addition, according to Baxter and Pan's updated reconstructions, CV or CGV open syllables did exist in OC content words (e.g. **ma* 摸 'to stroke' and **ŋa/ŋja* 鱼 'fish'). These are counter-examples to the OC minimal syllable structure that Feng proposes in Figure 7.1 (a). Thus the coda simplification that Feng mentions from OC to Middle Chinese is not substantial enough to generate a major change in the metrical structure of syllables.

According to Feng's analysis of metrical change in Chinese, the origin of disyllabicity, especially the formation of disyllabic compounds, could not happen in OC because its condition, the new disyllabic 'foot', had not been formed at that time. His prosodic explanation, therefore, has a serious chronological inconsistency. In addition, Feng's failure to consider Lianmian words strips his prosodic approach of a solid phonological support. If his approach did not exclude Lianmian words from the discussion of disyllabicity, the special Lianmian sound patterns might lead him to a more plausible phonological motivation for the lexical revolution upending monosyllables.

In summary, neither the functional approach nor the prosodic approach offers a comprehensive language-internal analysis of the rise of the two distinct types of disyllables and their relative chronology. The functional approach merely provides language-external explanations for the rise of disyllables, which are not well supported by the evidence in OC phonology, and which offer no insight into the chronology of distinct types of disyllalbes; the prosodic approach only studies the rise of disyllabic compounds, excluding the oldest disyllables, Lianmian words, from discussion. In contrast to these two approaches, my study not only analyzes the relation between syllable structure simplification in OC phonology and the rise of Lianmian words, but also discusses diachronic connections between the development of Lianmian words and the subsequent rise of disyllabic compounds.

CHAPTER 8

CONCLUSION

8.1 Summary of my approach

The study of the rise of disyllables in the history of Chinese involves many inter-related subfields: phonology, prosody, morphology and syntax. However, some researchers tend to analyze the issue from one subfield, because of their insistence on the strict segregation of synchronic and historical description. This strict segregation of synchrony and diachrony is not always productive. As emphasized in work within Evolutionary Phonology "this one size fits all methodology invariably fails to explain – and often fails even to describe accurately – many of the sound patterns that recur in the world's language" (Blevins, 2004, p. 3), In order to make a comparison and a summary of both the earlier hypotheses and new ones, I highlight the questions related to the rise and the development of disyllables and the answers offered by earlier approaches in Table 8.1.

Table 8.1 The issues that previous approaches address

Previous approaches / Questions to be answered	Sun (1999)'s reduplication approach	Duanmu (1999b)'s metrical approach	Feng (1998)'s prosodic approach	Functional approach (pragmatic view)
What is the motivation for the rise of disyllables	Not applicable	Due to an increase in new vocabulary, most of which consists of polysyllabic borrowings, polymorphemic translations, and polymorphemic creations.	Due to the prosodic change caused by the simplification of syllable coda from OC to Middle Chinese	Due to homophone-avoidance
When do disyllables start to increase? What is the earliest disyllabic form?	Not applicable	Modern Chinese Disyllabic compounds	Early Middle Chinese Disyllabic compounds	Late OC Lianmian words, and disyllabic compounds
How were Lianmian words created?	Through reduplication process	Not addressed	Not applicable	Through reduplication process
Why did Lianmian words lose their productivity around Late OC and Early MC?	Not addressed	Not addressed	Not applicable	Not addressed
What is the relation between the rise of disyllables and the creation of Lianmian words?	Not applicable	Not addressed	Not applicable	Not addressed
What are the influences of the rise and development of disyllables in Chinese lexicon?	Not applicable	The use of disyllables is determined by metrical structure. The development of disyllables has no impact on metrical structure.	Not addressed	Not addressed

Table 8.1 shows that previous approaches, including synchronic reduplication and metrical approaches, and historical functional and prosodic approaches, fail to explain why Lianmian

words lost their productivity around LOC and EMC, what the relation between the rise of disyllables and the creation of Lianmian words is, and what the influences of the rise and development of disyllables in Chinese lexicon is. Failure to answer these questions means that the rise of disyllabic compounds in LOC, and their increase in Middle and Modern Chinese remain mysterious. Although having made efforts to explore the metrical explanations for the rapid development of disyllabic compounds in Modern Chinese lexicon, Duanmu failed to seek the historical motivation for the rise of disyllables or to give a diachronic account for the formation of the disyllabic template. Although Feng attempted to use both prosodic considerations and a diachronic approach to solve the mystery of the rise of bi-morphemic disyllables in OC, there is a serious chronological blemish in his phonological analysis, which undermines his argument. His account does not cover the motivation for the rise of mono-morphemic disyllables in OC.

My study is built on both previous morphological and phonological research on disyllabicity and OC phonology, investigating the relation between OC phonology and the origin of disyllabicity diachronically. Monosyllabic words prevailed in OC (L. Wang, 1980), though from EOC, disyllables began to increase in written literature. Cheng, Xiangqing (1992) concludes that the development of disyllabic structures in EOC underwent three phases: the phase when mono-morphemic disyllables were formed through phonological patterns, the phase when disyllabic compounds were formed through phonological patterns, and the phase when disyllabic compounds were formed through morph-syntax (Cheng, 1992, p. 110). Based on these conclusions, the current study of disyllabicity aims to put Lianmian words and disyllabic compounds together under investigation and to bridge the development from one to the other.

The approach that I propose attempts to discover the diachronic phonological factors for internal stages of disyllabicity.

8.1.1 Considerations of syllable structures

The core of this approach lies in the analysis of the changes of OC segmental phonology and the sound patterns of Lianmian words. After examining Zhengzhang's reconstructed OC and Middle Chinese phonologies, I found that the biggest change in phonology is the simplification of onsets within a syllable. This simplification process progressed through the entire period of OC. Associated with Pan's hypothesis that in Proto-Chinese, there had been lots of sesqui-syllables, just like neighboring languages in which sesqui-syllables once abounded, I assume that the complex onsets in EOC are the products of reduction from sesqui-syllables. And they continued to undergo simplification until EMC.

What is the relation between this syllabic structure change and the rise of disyllables? The special sound pattern of Splitting-sound words ($O_aR_bLR_c$) provides us with a clue: the occurrence of Liquids in the second syllable onset is not accidental at all, rather reflecting the Liquids (*l or *r) in the consonant-cluster onset (CL-) which were in the process of being lost in the onset simplification process. As detailed in chapter 4, Splitting-sound words can be viewed almost as hyper-articulated structures, attempts by conservative speakers to preserve the complex onsets being lost in the vernacular speech surrounding them.

To connect Splitting-sound words with other Lianmian words, it is necessary to closely examine their phonological patterns. In Diéyùn data, Sun (1999) found a similar pattern: $O_aR_bLR_b$. Although there are some other variations of the second onset and the generalized pattern of Diéyùn is always $O_aR_bO_cR_b$, this similarity shows how the pattern develops and is

generalized into a more stable disyllabic template. Figure 5.8 in Chapter 5 is repeated here to illustrate the formation of the disyllabic template generated by the Splitting-sound pattern:

$O_aR_bLR_c$	$O_aR_bLR_b$	$O_aR_b\{L,N\}R_b$	$O_aR_bO_c\{L, N, T\}\ R_b$
Splitting-sound pattern	1st stage	2nd stage	3rd stage

Figure 8.1 Disyllabic Template formation process

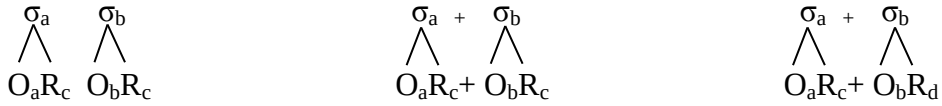
The extension from Splitting-sound pattern ($O_aR_bLR_c$) to the first stage of Di éyùn ($O_aR_bLR_b$) still follows the "faithfulness" principle, preserving the onset identity which was under pressure of loss in onset simplification. The following two steps toward the final generalized template are analogical extensions of the disyllabic pattern.

Phonological factors motivated the disyllabic template, which determined that disyllabicity instead of tri-syllabicity or multi-syllabicity would be the future direction of Chinese word development. Phonological motivation disappeared around LOC and EMC as the simplification of onsets ended. As a result, Lianmian words lost their productivity, but the template they defined was preserved, wielding more powerful influence on word formation in later Chinese development.

8.1.2 Prosodic and morphological considerations

Once Lianmian words lost their productivity, the disyllabic template was fed with monosyllabic independent morphemes. In the transitional phase, however, there are still some weak phonological connections between the two monosyllables including fully reduplicated syllables, those sharing similar onsets, or those sharing similar rhymes. Lianmian words in this period are called 'Lianmian compounds' (examples found in 6.1.1 (1-3)). They are considered a special kind of compound, mainly because at least one of the morphemes is free, and each of the morphemes is semantically associated with the compound as a whole. The two morphemes are

syntactically parallel, neither of which dominates the other. This syntactic feature of Lianmian words only agrees with that of the earliest and the most common type of compounds — parallel compounds (examples are shown in 6.1.2). This accordance confirms my assumption that Lianmian compounds are transitional forms which bridge over phonological and syntactic lexicalization under the disyllabic template:



Lianmian words	Lianmian compounds	Compounds
phonoligcal lexicalization	transitional phase	syntactic lexicalization

The beginning of syntactic lexicalization symbolized the formation of the prosodic template, when the disyllabic template began to absorb disyllabic phrases and idomatize them into prosodic words before some of them were finally lexicalized into compounds (Feng, 1998). In Middle and Modern Chinese, syntactic lexicalization through this disyllabic template became a primary new word formation tool. Because some of the disyllabic phrases did not become fully lexicalized, the incompleteness of disyllabicity led to the Dual-vocabulary phenomenon and the confusion between disyllabic phrases and disyllabic compounds in Modern Chinese. With this, the 'fluidity' of the word-morpheme boundary became a language-specific feature of Modern Chinese morphology.

In summary, my approach offers a plausible diachronic explanation for the rise and the development of disyllabicity in the history of Chinese, covering both mono-morphemic disyllables (Lianmian) and disyllabic compounds.

8.2 Contributions

The current study, in a diachronic research framework, contributes to disyllabicity theory, OC phonology, sound change, language contact and research on Chinese lexical change.

First, the data collection is the first large one of Lianmian words with updated OC reconstructions. It is the investigation of this database that leads to insights about the relation between disyllabicity and the rise of Lianmian words. This study places Lianmian words in an equal position with disyllabic compounds in the research on disyllabicity. Researchers who have focused on disyllabicity, especially Chinese phonologists, are inclined to exclude Lianmian words from their discussion, because disyllabic compounds are so frequent in Modern Chinese, while Lianmian words lost their productivity totally in EMC. In fact, the puzzle of the origin of disyllabicity can never be solved without considering the historical role of Lianmian words.

No previous analysis has offered a diachronic account for the $O_aR_bLR_c$ sound pattern of Splitting-sound words. Inspired by the unique pattern of Lianmian words, I explored the historical phonological reasons for their origins — which are supported by both language-internal and external evidence. Furthermore, this study shows that historical phonological motivations for the origins of Lianmian words and disyllabicity can be better accounted for if the sound pattern of Splitting-sound words and the pattern of Diéyùn are connected. Previous research simply ignored this association, or considered their similarity as a random coincidence.

Second, with close examination and comparison of the preserved Splitting-sound words in OC and Middle Chinese documents and the colloquial speech of Modern Chinese dialects, this study discovers a clear connection between the rise of Splitting-sound words and the simplification of *CL- onset in LOC. The loss of liquids in complex *CL- onsets involved simplification and neutralization of CL > C. Awareness of this neutralization provided a

phonological motivation to preserve the lost liquids in the cluster and so, to preserve the contrast between *CL- and *C- onsets. Zhengzhang (2003) and Pan (1999) assume that there is a type of sesqui-syllabic onsets *C.C- in OC phonology. Since in the *C.C- (including *C.L-) structure, the minor syllable onset C- is usually lost, while the major syllable onset C always preserved, the simplification of *C.C- (*C.L-) did not necessarily generate any Splitting-sound patterns. This analysis, then, if correct, confirms the existence of two onset types *C.L- and *CL-.

Third, my proposal that motivation for Splitting-sound words should be true complex onsets *CL- which were being simplified contributes to theories of sound change. The evolutionary approach is adopted to provide a likely scenario giving rise to Splitting-sound words which are not found as the product of onset simplification in other languages.

During the gradual sound change *CL- > *C-, older speakers, with CL- representation attempt to maintain CL- by (i) writing CL-words with two characters, — one to represent the initial C, another to represent the L; and by (ii) pronouncing them slowly, in a fortis way, so that the Head C is expanded to a full syllable, and the rest of the original syllable with the L and the rhyme is preserved. The reason Splitting-sound words associated with this precise kind of sound change are not found in Tibeto-Burman and Mon-Khmer languages, is because the ancient writing systems of both families writes the CL cluster, while the lost liquid cannot be recovered from a single written Chinese character. If this hypothesis is correct, Splitting-sound words can be seen as a kind of "unnatural" development (Blevins, 2004, 2006) where literacy and language-conservation interact with regular phonetically-based sound change.

Fourth, in the exploration of external evidence to support the relation between the simplification of *CL- and the rise of Splitting-sound words in OC, the study investigates the onset correspondence between the Splitting-sound words in Modern Chinese varieties and the

monosyllables in Mien dialects [ium]. This investigation has led to the discovery of complex onset correspondences between some Splitting-sound words in OC and potential cognates in Mon-Khmer languages. The correspondences show that OC has strong phonological and lexical correlations with Hmong-Mien and Mon-Khmer languages. There is no evidence to show any genetic relations between Proto-Chinese/OC and Hmong-Mien and between Proto-Chinese/OC and Austroasiatic languages. As such, these correspondences may provide further support for the language contact model proposed by DeLancey that 'Proto-Chinese could have been widely used as a lingua franca by non-Chinese outside the Chinese state at that time (before Zhou or even Shang dynasty) until the widespread lingua franca version replaced the original everywhere ultimately' (DeLancey, 2011, p. 60).

Last but not least, this research makes use of the results of the computational model of lexical change over time proposed by Wedel (2004, 2007) to the analysis of the factors leading to the dramatic increase of disyllables from LMC to Modern Chinese. The disyllabic-compound outputs that the disyllabic template $\sigma_{a+}\sigma_b$ created reinforced learners' perception, which in turn created a loop with production. This self-feeding loop promotes the usage of the template as a word-formation tool, producing more and more disyllables in the Modern Chinese lexicon.

8.3 Future research

This study focuses on the discussion on consonant-cluster onset simplification in OC phonology and its influence on the rise of disyllables. It does not, however, address the influence of coda simplification in OC on the development of disyllabicity. For example, I found that the difference between the two patterns $O_aR_bLR_c$ (Splitting-sound pattern) and $O_aR_bLR_b$ (Di éyùn) did not lie in their rhymes but in their codas. While some linguists argue that the coda in OC is connected with tonogenesis (Mei, 1980; Sagart, 1999; Zhengzhang, 1987), my analysis claims

that the evolution of the disyllabic template is independent of tonogenesis. The current model is forced to invoke reduplication to explain the rise of Shuāngshēng (another kind of Lianmian words), unable to provide a unified account for the rise of both Shuāngshēng and Diéyùn words. If this is a weakness of the model, it may be because coda simplification in OC has not been explored in depth. Future detailed research on coda changes in OC may contribute to understanding the rise of Shuāngshēng words and some of the additional phonological factors involved in the transition from Splitting-sound patterns to Diéyùn patterns. Such research will shed light on the evolution of Chinese disyllables and hopefully contribute to a better understanding of more general aspects of Chinese prosody.

APPENDIX I: SHUANGSHENG DATA IN OLD CHINESE

Shuāngshēng words are OC disyllables where the onsets of the first and second syllables are identical or near-identical but where rhymes of the two syllables can differ (Chapter 3). These words are attested as early as around 1000 BC. The forms in the table below are the basis of discussion of Shuāngshēng words in this dissertation.

Nouns		**Reconstructed OC (*)**				**Pinyin**	**Gloss**
		Zhengzhang	**Pan**	**Li**	**Baxter**		
1	棠棣	daaŋ l'ɯɯds[47]	g·laaŋ b ɬɯɯds	daŋ diədh	daŋ ləts	tang2/di4	Chinese bush cherry
2	荎蕷	di ta	di k ɬa	drjid trjag	drjij trja	chi2/zhu1	the fruit of Chinese magnolia vine
3	薤荭	kree kooʔ	kree kooʔ	krig kugx	kre koʔ	xie4/hou4	water chestnut
4	枇杷	piʔ braa	piʔ braa	pjidx brag	pjijʔ bra	pi2/pa2	loquat
5	蚍蚾	bi bu	bi bɯ	bjid bjəg	bjij bjə	pi2/fu2	a kind of grass
6	鼎薑	teeŋʔ tooŋʔ	teeŋʔ tooŋʔ	tiŋx tuŋx	teŋʔ toŋʔ	ding3dong3	root of lotus
7	蘼蕪	mril ma	mril ma	mjid mjag	mrjij mja	mi2wu2	a kind of grass
8	蘡薁	qeŋ qug	qeŋ qug	ʔjiŋ ʔjəkw	ʔjeŋ ʔjuk	ying1/yu4	a kind of grapes
9	镐侯	gaawʔ goo	gaawʔ goo	gagwx gug	gawʔ go	gao3/hou2	Cyperus rotundus Linn.
10	铫弋	lew lɯg	[g]lew lɯg	ragw rək	ljew ljək	yao2/yi4	carambola
11	薏英	qrab qraŋ	qrab qraŋ	ʔjap ʔjiaŋ	ʔrjap ʔrjaŋ	yi4/ying1	Job's tears
12	藷蔗	tja tjaags	klja kljags	tjag tjiagh	tja tjʌks	zhu1/zhe4	sugar cane
13	蠨蛸	sɯɯw sreew	slɯɯw sreew	siəgw sragw	siw srew	xiao1/shao1	a kind of spider
14	蠐螬	zliil zluu	ziil zuu	dzid dzəgw	ɦtshəj dzu	qi2/cao2	grub
15	至掌	tjigs tjaŋʔ	tjigs tjaŋʔ	tjidh tjaŋx	tjits tjaŋʔ	zhi4/zhang3	leech
16	蚍蜉	bi bu	bi bu	bjid bjəgw	bjij bju	pi2/fu2	big ants
17	鼅鼄	te to	k ɬe to	trjig trjug	trje trjo	zhi1/zhu1	spider
18	蛣蚰	k^hid k^hlud	k^hid k^hlud	khjit khwjət	khjit khjut	jie2/qu1	scorpion
19	渠螅	ga k^hwed	ga k^hwed	gjag khwjat	gja kh^wjet	qu2/que1	a kind of bug
20	蚾蚗	qi qod	qi qod	ʔjid ʔuat	ʔjij ʔjot	yi1/jue2	a kind of cicada
21	蠛蠓	meed mooŋ	meed mooŋ	miat muŋ	met moŋ	mie4/meng3	a kind of fly

To be continued

[47] *l’ stands for the obstruent allophone of *l-.

22	㠯螽	buʔ ban	buʔ ban	bjəgwx bjan	bjuʔ bjan	fu4/fan2	grasshoppers
23	悉蝨	sid sruds	slid sruds	sjit srjədh	sjit srjuts	xi1/shuai4	crickets
24	蜌威	qlil qul	qlil qul	ʔjid ʔwjəd	ʔjij ʔjuj	yi1/wei1	a kind of bug
25	虎蝓	slal lo	slal lo	star rug	saj ljo	si1/yu2	snail
26	蟦蜓	lin l'eenʔ	lin leenʔ	rin dianx	ljin lenʔ	dian4/ting2	dragonfly
27	蝤蠐	s ɬlu zliil	s ɬlu ziil	tsjəgw dzid	tsju ɦtshəj	qiu2/qi2	lonicorn's larva
28	牟母	mu mɯʔ	mu mɯɯʔ	mjəgw məgx	mju məʔ	mou2/mu3	a kind of Turnix species
29	鴛鴦	quun qaaŋ	quun qaaŋ	ʔwən ʔaŋ	ʔaŋ ʔun	yuan1/ yang2	mandarin duck
30	鶟□	kiig ŋed	kiig ŋed	kit ŋjat	kik ŋjet	jie2/jia2	wild ducks
31	鸐鵝	sɯwɢ sraŋ	sliwg sraŋ	sjəkw srjaŋ	sjiwk srjaŋ	su4/shuang1	a kind of water fowl
32	鶩離	m ɾu reels	m ɾu b ɾeels	ljəgw liarh	c-rju c-rejs	liu2/li2	oriole
33	鶹鳩	klud ku	klud ku	kwjət kjəgw	kjut kju	jue2/jiu1	a kind of bird
34	韃鷥	gaans gruug	gaans gruug	ganh grəkw	gans grəkw	han4/xue2	titmouse
35	秸鶵	kriil kug	kriil kug	krid kjəkw	krij kjuk	jia2/ju2	cuckoo
36	詹諸	tjam tjaa	kljam klja	tjam tjiag	k-ljam tjᴀ	zhan1/zhu1	toad
37	蝙蝠	peen pɯg	peen pɯg	pian pjək	pen pjək	bian1/fu2	bat
38	溥浮	phaaʔ bu	phaaʔ bu	phagx bjəgw	phaʔ bju	pu3/fu2	black finless porpoise
39	猲獢	ŋ̊ad ŋ̊ew	ŋ̊ad ŋ̊ew	ŋ̊jat ŋ̊jagw	ŋ̊jat ŋ̊jaw	xie1/xiao1	a kind of hunting dog
40	麃貑	kɯʔ kraa	kɯʔ kraa	kjəgx krag	kjəʔ kra	jiu4/jia1	a kind of deer
41	騊駼	l'aa l'uu	luu laa	dəgw dag	b-lu la	tao2/tu2	name for a kind of good horse
42	貑玃	kraa kwag	kraa kwag	kwjak	kra kʷjak	jia1/jue2	a large female ape
43	町疃	tʰeenʔ l̥ooŋ	tʰeenʔ kʰ·looŋ	thianx thuŋ	thenʔ hloŋ	ting1/tuan3	tracks of deer
44	鎕锑	l'aaŋ diil	g·laaŋ liil	daŋ did	g-laŋ dij	tang2/ti1	a kind of mineral
45	螮蝀	teeds tooŋ	teeds tooŋ	tiadh tuŋ	tets toŋ	di4/dong1	rainbow
46	霢霂	mreeg moog	mreeg moog	mrik muk	mrik mok	mai4/mu4	light rain
47	椌楬	kʰrooŋ kʰraad	kʰrooŋ kʰraad	khruŋ khrat	khroŋ khrat	qiang1/jie1	a kind of musical instrument
48	牝服	bilʔ bɯg	bilʔ bɯg	bjidx bjək	bjijʔ bjək	pin4/fu2	the bars on the

To be continued

							sides of a cart
49	涒滩	l̥uun n̥aan	k^{h}·l̥uun n̥aan	thən hnan	thun hnan	tun1/tan1	a way of numbering the years in the lunar calendar
50	刮剧	kral kɯd	ral kɯd	kjiar kjət	krjaj kjut	ji1/jue2	a curved knife
51	衹裯	tiil tuuw	tiil tɯɯw	tid təgw	tij tiw	di1/dao1	short clothes
52	廲廔	reel roo	reel g ɾoo	liar lug	c-rej c-ro	li4/lou2	bright windows elegantly carved
53	鎗鏦	s^{h}raaŋ s^{h}looŋ	sk^{h}raaŋ sk^{h}ooŋ	tshraŋ tshuŋ	tshraŋ tshoŋ	cheng1/ cong1	bell sound
Verbs							
54	覲覣	bjɯns ban	bjɯns ban	biənh bjan	bjəns bjan	bin4/fan2	to look shortly
55	首鼠	l̥juʔ l̥jaʔ	q^{h}ljuʔ q^{h}ljaʔ	hrjəgwx hrjagx	hljuʔ hjaʔ	shou3/shu3	to hesitate
56	蹃躅	deg dog	deg dog	drjik drjuk	drjek drjok	di2/zhu2	to hesitate
57	唐逮	gl'aaŋ l'ɯɯds	g·laaŋ b ɬɯɯds	daŋ dədh	g-laŋ ləts	tang2/dai4	to capture
58	歟瘉	le lo	le lo	rig rug	lje ljo	ye1/yu4	to make fun of'
59	覜覷	s^{h}egs s^{h}as	s ƙ^{h}egs s ƙ^{h}as	tshjigh tshjagh	tshjeks tshjas	ci1/qu4	to spy
60	了戾	reewʔ rɯɯds	reewʔ rɯɯds	liagwx liədh	c-rewʔ c-rəts	le/li4	to curl
61	邂逅	grees goos	grees goos	grigh gugh	gres gos	xie4/hou4	to meet without appoinment
62	契闊	k^{h}eeds k^{h}ood	k^{h}eeds k^{h}ood	khiadh khuat	khets kh^{w}at	qi4/kuo4	be separated far away
Adjectives							
63	忼慨	k^{h}lud k^{h}ɯɯds	k^{h}lud k^{h}ɯɯds	khwjət khədh	khjut khəts	kang1/kai3	emotionally aroused
64	尷尬	kreem kreeds	kreem kreeds	kram kriadh	krem krets	gan1/ga4	ill at ease
65	譸張	tu taŋ	tu k·laŋ	trjəgw trjaŋ	trju trjaŋ	zhou1/ zhang1	frightened
66	坎坷	k^{h}oomʔ k^{h}aalʔ	k^{h}oomʔ k^{h}aalʔ	khəmx kharx	khomʔ khajʔ	kan3/ke1	full of frustrations
67	襤褸	g ɾaam roo	g ɾaam g ɾoo	lam lug	g-ram c-ro	lan2/lü3	shabby
68	涔濱	s ɡrɯm ʔ sli	s ɡrɯm s ʇi	dzrjəm tsjid	dzrjəm tsjij	cen2/zi1	sodden
69	憔悴	zew zuds	s ɡlew zuds	dzjagw dzjədh	dzjew sdjuts	qiao2/cui4	wan and sallow, withered
70	婷娥	ŋ·greeŋ ŋaal	ŋ·greeŋ ŋaal	ŋriŋ ŋar	ŋreŋ ŋaj	xing2/e2	beautiful

To be continued

71	黤姁	q^{h}eem go	q^{h}eem go	hiam gjug	xem gjo	xian1/xu3	reddish yellow
72	頡顗	kriid kljod	kriid kljod	krit tjuat	krit tjot	jie2/zhuo1	description of upward cheekbone
73	冶由	laaʔ lɯw	laʔ lɯw	ragx rəgw	ljᴀʔ ljiw	jia1/ya2	seductive, bewitching
74	容與	loŋ laʔ	[g]loŋ laʔ	grjuŋ ragx	ljoŋ ljaʔ	rong2/yu2	leisurely and careless; up and down (of pacing)
75	閼嬝	daab dɯɯ	daab dɯɯ	dap dəg	dap lə	ta4/tai2	slow
76	趍趙	deels l'ewʔ	deels rlewʔ	driarh drjagwx	drejs lrjewʔ	qu1/zhao4	slow (of walking)
77	趑趄	s^{h}li s^{h}jaas	sn̥i sk^{h}jas	tshjid tshjiagh	tshjij tshjᴀ	zi1/ju1	hesitate to advance
78	謰謱	ren roo	b ɾen g ɾoo	ljan lug	c-rjen c-ro	lian2/lou2	inarticulate
79	連遱	ren roo	b ɾen g ɾoo	ljan lug	c-rjen c-ro	lian2/lou2	continuous
80	婘娞	quum qaal	quum qaal	ʔəm ʔar	ʔum ʔaj	an1/e1	hesitant, uncertain
81	簣箸	dɯw tas	dɯw k ɬas	drəgw trjagh	drjiw trjas	chou2/zhu4	hesitant
82	逸豫	lig las	lig las	rit ragh	ljit ljas	yi4/yu4	living a comfortable life
83	詰詘	k^{h}id k^{h}lud	k^{h}id k^{h}lud	khjit khwjət	khjit khjut	jie2/qu1	curved
84	覭髣	meeŋ mu	meeŋ mu	miŋ mjəgw	meŋ mju	ming2/ mao2	forested
85	繽紛	p^{h}in p^{h}ɯn	p^{h}in p^{h}ɯn	phjin phjən	phjin phjən	bin1/fen1	in riotous profusion
86	絪緼	qin quun	qin quun	ʔjin ʔwən	ʔjin ʔun	yin1/yun4	misty, foggy
87	仓卒	s^{h}aaŋ ʔ suud	s ɬ^{h}aaŋ suud	tshaŋ tsət	tshaŋ stut	cang1/cui4	hasty
88	跨區	kralʔ qoo	kalʔ qoo	kjiarx ʔug	krjajʔ ʔo	qi1/qu1	uneven, rough
89	阬坎	k^{h}aaŋs k^{h}oomʔ	k^{h}aaŋs k^{h}oomʔ	khaŋh khəmx	khaŋs khomʔ	keng1/kan3	dangerous, deeply trapped
90	參差	s^{h}luum s^{h}raal	s^{h}uum s ɬ^{h}raal	tshəm tshrar	tshum tshraj	cen1/ci1	uneven, irregular
91	畢沸	pid pɯds	pid pɯds	pjit pjədh	pjit pjəts	bi4/fei4	description of springs coming out
92	仿佛	ben bɯd	ben bɯd	bjian bjət	bjen bjət	fang3/fu2	similar, alike
93	斯须	se so	sle so	sjig sjug	sje sjo	si1/xu13	a little while
94	凓冽	rig red	[b]rig [b]red	ljit ljat	c-rjit c-rjet	li4/lie4	cold

To be continued

95	阢隉	ŋuud ŋeed	ŋuud ŋeed	ŋwət ŋiat	ŋut ŋet	wu4/nie4	unstable
96	駸驙	tɯn ten	k ɬɯn k ɬen	trjiən tran	trjən tran	zhen3/ zhan1	(horse) heavy-loaded

APPENDIX II: DIEYUN DATA IN OLD CHINESE

Di éyùn words are OC disyllables where the rhymes of the first and second syllables are identical or near-identical but where onsets of the two syllables can differ (Chapter 3). These words are attested as early as around 1000 BC. The forms in the table below are the basis of discussion of Di éyùn words in this dissertation.

Nouns		**OC**				**Pinyin**	**Gloss**
		Zhengzhang	**Pan**	**Baxter**	**Li**		
1	蒹葭	tiiʔ niilʔ	tiiʔ niilʔ	tidx nidx	tijʔ nijʔ	di3/ni3	kind of grass
2	牂薖	ʔ·saaŋ naaŋ	skaaŋ naaŋ	tsaŋ nraŋ	tsaŋ nraŋ	zang1/ ning2	a kind of grass
3	稖穅	pʰaaŋ gwaaŋ	pʰaaŋ gwaaŋ	phaŋ gwaŋ	phaŋ waŋ	pang2/ huang2	a variety of panicled millet
4	苜蓿	mood sug	mood slug	muat sjəkw	mot sjuk	mu4/xu4	alfalfa
5	菡萏	guumʔ l'oomʔ	guumʔ g ɬoomʔ	gəmx dəmx	gumʔ domʔ	han2/dan4	lotus bloom
6	芣苢	bɯ lɯʔ	bɯ lɯʔ	bjəg rəgx	bjə ljəʔ	fu2/yi3	snokeweed
7	活苽	kood l̥ood	kood l̥od	kuat thuat	kʷat hlot	huo2/tuo1	bush or small arbor
8	茱萸	djo lo	djo lo	djug rug	djo ljo	zhu1/yu2	Cornus officinalis
9	杜若	l'aaʔ nja	g ɬaaʔ nja	dagx njiag	laʔ njᴀ	du4/ruo4	a kind of sweet herbs
10	茉莀	ʔ sliw gu	skliw gu	tsjəgw gjəgw	stjiw gju	jiao1/qiu2	fruit of cornels
11	菡萏	guumʔ l'oomʔ	guumʔ g ɬoomʔ	gəmx dəmx	gumʔ domʔ	han4/dan4	a second name for lotus flower
12	蓳歷	beg reeg	beg reeg	bjik lik	bjek c-rek	bi4/li4	Stenoloma chusana
13	芎藭	kʰwɯŋ guŋ	kʰwɯŋ guŋ	khwjəŋ gjəŋw	khʷjəŋ gjuŋ	xiong1/ qiong2	a plant name
14	薏苡	qɯg lɯʔ	qɯg lɯʔ	ʔjək rəgx	ʔjək ljəʔ	yi4/yi3	seed of Job's Tears
15	茹蘆	nja raa	nja raa	njag lag	nja c-ra	ru2/lu2	*Rubia cordifolia Linn.*
16	樸樕	pʰroog sloog	phroog sloog	phruk suk	phrok sok	pu3/su4	a kind of small tree
17	楅椮	t ɬuub l'uub	k ɬuub g ɬuub	təp dəp	k-lup dup	da2/ta4	a kind of tree (as plum)

To be continued

18	螟蠕	meeŋ reeŋ	meeŋ reeŋ	miŋ liŋ	meŋ c-reŋ	ming2/ ling2	silkworm
19	强羊	gaŋ laŋ	gaŋ [g]laŋ	gjaŋ raŋ	gjaŋ ljaŋ	qiang2/ yang2	a small black beetle
20	竈蟊	to mu	to mu	trjug mjəgw	trjo mju	zhu1/mao2	a kind of fly
21	堂蜋	daaŋ raaŋ	g·laaŋ raaŋ	daŋ laŋ	daŋ c-raŋ	tang1/lang2	mantis
22	蠡蛸	p^{h}ew sreew	phew sreew	phjew sragw	phjew srew	xiao1/ shao1	mantis eggs
23	蓊蜙	qlooŋ ʔ·sloŋ	qlooŋ soŋ	ʔuŋ tsjuŋ	ʔoŋ tsjoŋ	weng1/ zong1	a kind of gad-fly
24	蛺蜨	keeb seeb	keeb qeeb	kiap siap	kep sep	jia2/die2	butterfly
25	蒲盧	baa b ɾaa	baa raa	bag lag	ba c-ra	pu2/lu2	a kind of wasp
26	蠣蠃	klool rool	klool rool	kuar luar	k^{w}aj c-roj	guo3/luo3	a kind of wasp
27	蛁蟟	teew reew	k-leew [g]reew	tiagw liagw	tew c-rew	diao1/liao2	a kind of cicada
28	夗蟺	qonʔ djanʔ	qonʔ djanʔ	ʔjuanx djanx	ʔjonʔ djanʔ	wan1/ shan4	earthworm
29	丁蛵	r·teeŋ q^{h}eeŋ	rteeŋ qheeŋ	triŋ hiŋ	rteŋ xeŋ	ding1/ xing2	dragonfly
30	蜻蛉	ʔ·sleŋ reeŋ	skeŋ [m]reeŋ	tsjiŋ liŋ	tsjeŋ c-reŋ	qing1/ling2	a kind of dragon-fly
31	鵂鶹	qhu ru	qhu g ɾu	hjəgw ljəgw	xju c-rju	xiu1/liu2	a kind of owl
32	鵗鶀	gɯ kɯ	gɯ kɯ	gjəgh kjəg	gjəs gjə	ji4/qi1	owlet
33	鴘□	pɯn m̥ɯɯn	pɯn m̥ɯɯn	pjən hnən	pjən hnən	fen1/hun1	a kind of bird
34	鴝鵒	koo log	koo [g]log	kug grjuk	ko ljok	qu2/yu4	mynah
35	鴾鵝	kaal ŋaal	kaal/ŋaal	kar/ŋar	kaj/ŋaj	ge1/e2	goose
36	鸑鷟	ŋroog zroog	ŋroog/r·z roog	ŋruk/dzruk	ŋrok/dzrok	yue4/zhuo2	a kind of phoenix
37	鸛專	koons tjon	koons tjon	kuanh tjuan	kons tjon	guan4/ zhuan1	magpie
38	鴅鵖	bɯb prɯb	bɯb pɯb	bjəp prjəp	bjəp pjəp	fu2/bi2	a kind of bird
39	鷦□	ʔ sews mewʔ	s ɬlews mewʔ	tsjagwh mjiagwx	tsjews mjewʔ	jiao1/ miao3	a kind of small bird
40	倉庚	s^{h}aaŋ kraaŋ	s ɬk^{h}aaŋ kraaŋ	tshaŋ kraŋ	tshaŋ kraŋ	cang1/ geng1	oriole
41	鸕諸	raa tjaa	raa klja	lag tjiag	c-ra tjᴀ	lu2/zhu1	cormorant

To be continued

42	蝦蟆	q^{h}raa mraa	q^{h}raa mraa	hrag mrag	xra mra	ha1/ma2	toad
43	禿黿	rug s^{h}lug	rug s^{h}ug	ljəkw tshjəkw	c-rjuk tshjuk	lu4/cu4	toad
44	蜥易	seeg leegs	sleeg legs	sik righ	sek ljeks	xi1/yi4	lizard
45	蝘蜓	qeenʔ l'eenʔ	qeenʔ leenʔ	ʔianx dianx	ʔenʔ lenʔ	yan3/ting2	wall gecko
46	鼆令	beŋ ren	beŋ [m]ren	bjiŋ ljan	brjeŋ c-rjen	ping2/ling4	mountain rat
47	服翼	bɯg lɯg	bɯg lɯg	bjək rək	bjək ljək	fu2/yi4	bat
48	駏驉	gaʔ q^{h}a	gaʔ q^{h}la	gjagx hjag	gjaʔ xja	ju4/xu1	horse
49	獿獀	nuuw sru	nuu sru	nəgw srjəgw	nu srju	you1/sou1	a kind of dog
50	觰挚	r ɬaaʔ r ɲaa	k ʁaaʔ r ɲaa	tragx nrag	traʔ nra	zha1/na2	a big animal name
51	獨狢	doog log	doog [g]log	duk grjuk	dok ljok	du2/yu4	a kind of animal that looks like a tiger
52	彷約	pljewɢ ʔ ɬewɢs	pljewɢ q ɒlewgs	tjewk ʔjewks	tjakw ʔjagwh	bo2 yue1	meteor
53	蜩蛹	mlaŋʔ raŋʔ	m·gaŋʔ raŋʔ	mjaŋx ljaŋx	mjaŋʔ b-rjaŋʔ	wang3/ liang3	monsters in mountains and rivers
54	縮朒	srug nug	srug nug	srjəkw nrjəkw	srjuk nrjuk	suo1/nü4	crescent in the east
55	劈歷	p^{h}eeg reeg	p^{h}eeg reeg	phik lik	phek c-rek	pi1/li4	thunder
56	厓屭	s·tol ŋal	stol ŋal	tsjar ŋjar	tsjoj ŋaj	zui1/wei2	a mountain peak
57	附娄	bos g-roo	bos [g]roo	bjugh lug	bjos c-ro	fu4/lou2	a small mountain
58	俾倪	peʔ ŋee	peʔ ŋee	pjigx ŋig	pjeʔ ŋe	bi3/ni2	the holes in the city walls
59	滎濘	gweeŋ nees	ɢ^{w}eeŋ nees	gwiŋ nigh	weŋ nes	ying2/ ning4	puddle
60	槫栌	paag raa	paag raa	pak lag	pak c-ra	bo2/lu2	ridge beam
61	桱桯	keeŋs l^{h}eeŋ	keeŋs l̥eeŋ	kiŋh thiŋ	keŋs hleŋ	jing4/ cheng2	a kind of hard wood
62	籧篨	ga r ɬa	gla rla	gjag drjag	gja lrja	ju3/chu2	a rough bamboo mat
63	篋箄	sreel pee	sreel pee	srar pig	srej pe	shai1/pai2	a container made of bamboo
64	桻雙	gruuŋ srooŋ	gruum srooŋ	grəŋw sruŋ	ɦkruŋ cCroŋ	xiang2/ shuang1	a sail made of thin bamboo

To be continued

							stripes
65	罣罷	doog roog	doog roog	duk luk	dok c-rok	zhu3/lu4	a small fishing net
66	茲箕	ʔɯs kɯj	sɯ kɯ	tsjəg kjəg	tsjə kjə	zi1/ji1	the curved handle of hoes
67	兜鍪	too mu	too mu	tug mjəgw	to mju	dou1/mou2	helmet used in battle
68	鋃鍜	qraa graa	qraa graa	ʔrag grag	ʔra gra	ya1/xia2	neck armour
69	槍唐	s^{h}aŋ gl'aaŋ	skhaŋ g·laaŋ	tshjaŋ daŋ	tshjaŋ g-laŋ	qiang1/ tang2	saw
70	玓瓅	pleewg reewg	p ɬeewg [g]reewg	tiakw liakw	tewk g-rewk	di4/li4	the light of pearl
71	鱕□	baan slaan	baan slaan	ban san	ban san	pan2/shan1	ragtag
72	令丁	ren r-teeŋ	[m]ren rteeŋ	ljan triŋ	c-rjen treŋ	ling4/ding1	little bells
73	銼鑹	s^{h}ools rool	skhlools rool	tshuarh luar	tshojs c-roj	cuo4/luo2	small pan
74	鋯錍	ʔ se p^{h}ee	se phee	tsjig phig	tsje phe	zi1/pi1	hatchet
75	鋃鐺	raaŋ taaŋ	raaŋ k·laaŋ	laŋ taŋ	c-raŋ taŋ	lang2/ dang1	lock
76	殺毀	kɯɯ q^{h}ɯɯg	kɯɯ qhlɯɯ	kəg həg	kə xə	gai1/yi3	ornaments for exorcising a ghost
77	穹隆	k^{h}wɯŋ gruuŋ	kh^{w}ɯŋ [g]rum	khwjəŋ ljəŋw	kh^{w}jəŋ c-rjuŋ	qiong2/ long2	a mountain name
78	繫緩	keegs re	keegs re	kigh lig	keks c-re	xi4/li2	waste cotton
79	鞀遼	deew reew	g ɬeew [g]reew	diagw liagw	dew c-rew	tao2/liao2	drum
80	褚衧	taʔ ɢwaʔ	k ɬaʔ ɢ^{w}aʔ	trjagx gwjagx	trjaʔ wjaʔ	zhu3/yu2	big women’s robe
81	裋褕	djoʔ low	gljoʔ low	djugx ragw	djoʔ ljaw	shu4/yu2	coarse clothes of houseboys
82	醬酯	moog l'oo	moog loo	muk dug	mok lo	mu2/tu2	a thick sauce made from elm pods
83	粻餭	taŋ gwaaŋ	k·laŋ g^{w}aaŋ	trjaŋ gwaŋ	trjaŋ waŋ	zhang1/ huang2	provisions
84	饘饎	qwɯɯn ʔ qɯɯn	quunʔ qɯɯn	ʔwənx ʔən	ʔunʔ ʔən	wen4/en4	food (wheat gruel, etc.) for the guest in old times
85	焦僥	ʔ sew k·ŋeewʔ	sklew keewʔ	tsjagw kiagwx	tsjew kewʔ	jiao1/yao2	a legendary Kingdom in the

To be continued

							south where people were very short
86	仳催	p^heʔ q^hwi	pheʔ qh^{w}i	phjigx hwjid	phjeʔ xjuj	pi2/sui1	ugly females
87	嬰婗	qii ŋee	qii ŋee	ʔid ŋig	ʔij ŋe	yi1/ni2	infant
88	膀光	baaŋ kwaaŋ	baaŋ k^{w}aaŋ	baŋ kwaŋ	baŋ k^{w}aŋ	pang2/ guang1	bladder
89	匡當	k^hwaŋ taaŋ	k^hwaŋ	khwjaŋ taŋ	kh^{w}jaŋ taŋ	kuang1/ dang1	frame
Verbs							
90	旁皇	baaŋ gwaaŋ	baaŋ g^{w}aaŋ	baŋ gwaŋ	baŋ waŋ	pang2/ huang2	to walk back and forth
91	旘□	l'ee gwee	g ɬee g^{w}ee	dig gwig	le g^{w}e	di1/xie2	to walk lame
92	瀎泹	meed q^hwaad	meed qh^{w}aad	miat hwat	met hwat	mo4/sa4	to obliterate
93	齞契	greeds kleed	greeds kleed	griadh kiat	grets ket	xie4/gua1	to scrub and brush
94	號咷	ɦ ɬaaw l'aaw	glaaw g ɬaaw	gagw dagw	gaw g-law	hao2/tao2	to cry, of baby
95	滶瀁	l'aaŋʔ laŋʔ	laaŋʔ laŋʔ	daŋx raŋx	daŋʔ ljaŋʔ	dang4/ yang4	(river) to run rapidly
96	滴歷	teeg reeg	teeg reeg	tik c-rek	tek lik	di1/li4	to tick
97	唌嘆	l'aan n^haan	laan n̥aan	dan hnan	dan hnan	xian2/tan4	to sigh
98	魛烇	bruu q^hruu	bruu qhruu	brəgw hrəgw	bru xru	pao2/xiu1	to roar
Adjectives							
99	康寬	k^hlaaŋ raaŋ	k^hlaaŋ raaŋ	khaŋ laŋ	khaŋ c-raŋ	kang1/ lang2	spacious, vacant
100	鏤鑸	quulʔ rulʔ	quulʔ [g]rulʔ	ʔwədx jərx	ʔujʔ c-rjujʔ	wei3/lei3	rough
101	鍖鉦	l^humʔ njɯm	k^h ɬumʔ njum	thrjəmx njəm	kh-ljumʔ njum	chen3/ren3	slow, leisurely
102	嘐啁	gruuw tuuw	[g]ruu tɯɯw	ləgw təgw	c-ru tiw	jiao1/zhao1	Like a bird sound
103	窈窕	qeewʔ l'eewʔ	qeewʔ g ɬeewʔ	ʔiagwx diagwx	ʔewʔ g-lewʔ	yao2/tiao2	reticent, modest
104	謵讂	ljub ʔ ɬjob	sɢlub kljob	rjəp tjap	zjup tjop	xi2/zhe2	loquacious
105	讂讄	tjeb njeb	tjeb ŋgljeb	tjap njap	tjep njep	zhe2/nie4	loquacious

To be continued

106	諈諉	tols nols	tols nols	trjarh nrjarh	trjojs nrjojs	zhui4/wei3	fatigued, onerous
107	謎諧	l'uub duub	g ɬuub duub	dəp dəp	dup dup	ta4/ta4	reckless (of speaking)
108	輸孺	h ɬjo njos	l̥ njos	hrjug njugh	hljo njos	shu1/ru3	slow, retarded
109	掘閱	glod lod	glod [g]lod	gjuat ruat	gjot ljot	jue2/yue4	description of baby insects coming out of the earth
110	娑妙	tems q^hwem	k ɬems q^hwem	trjamh hwam	trjems x^wem	chan1/jin4	smiling
111	嬰娠	qeŋ mreeŋ	qeŋ mreeŋ	ʔjiŋ mriŋ	ʔjeŋ mreŋ	ying1/ ming2	shy
112	媒姽	qloolʔ nloolʔ	qoolʔ noolʔ	ʔuarx nuarx	ʔojʔ nojʔ	wo3/nuo3	graceful, weak
113	妥鹵	gaaʔ raaʔ	gaaʔ raaʔ	gagx lagx	gkaʔ c-raʔ	hu4/lu3	greedy
114	嬋媛	djan ɢwans	djan ɢ^{w}ans	djan gwjanh	djan wjans	chan2/ yuan2	connected, sentimentally attached to someone
115	媻娑	baal saal	baal slaal	bar sar	baj saj	pan2/suo1	a mass of branches and leaves
116	頒妍	peens ŋeen	pens ŋgeen	pjianh ŋian	pjens ŋken	yu3/yan2	beautiful (hair style)
117	披靡	p^hral mralʔ	p^hal mralʔ	phjiar mjiarx	phrjaj mrjajʔ	pi1/mi2	description of being swept by the wind
118	峭嶸	zraaŋ ɢwreeŋ	sgraaŋ ɢwreeŋ	dzraŋ gwriŋ	dzraŋ wreŋ	zheng1/ rong2	steep, high
119	嵯峨	zlaal ŋaal	sgaal ŋaal	dzar ŋar	dzaj ŋaj	cuo2/e2	high and steep (mountain)
120	焦峣	ʔ sew ŋeew	sklew ŋeew	tsjagw ŋiagw	tsjew ŋew	jiao1/yao2	high (mountain)
121	岑崟	s ɡrɯm ŋ·grɯm	sgrɯm ŋ·grɯm	dzrjəm ŋjiəm	dzrjəm ŋrjəm	cen2/yin2	high and steep (mountain)
122	崔嵬	zluul ŋ·guul	zuul ŋguul	sdəd ŋwəd	dzuj ŋuj	cui1/wei2	lofty
123	巂嶵	ruuls zuulʔ	[g]ruuls zuulʔ	lərh dzədx	c-rujs dzujʔ	lei3/zui4	description of mountains
124	磊竚	ruulʔ tuul	ruulʔ tuul	lədx təd	c-rujʔ tuj	lei3/dui4	description of piling up
125	淊湣	t ƙ^{h}ub	k^h ɬub	thrjəp	kh-ljup	chi4/ji2	description of

To be continued

		ʔ sib	skib	tsjəp	tsjəp		water boiling
126	淶沛	naals poobs	naals poods	narh puadh	najs pots	nai4/pei4	description of waves
127	冥巠	meeŋ keeŋ	meeŋ keeŋ	miŋ kiŋ	meŋ keeŋ	ming2/ jing1	description of floody water
128	瀆菐	l'oog boog	loog boog	duk buk	lok bok	du2/pu2	trivial
129	駊騀	paalʔ ŋaalʔ	paalʔ ŋaalʔ	parx ŋarx	pajʔ ŋajʔ	po3/e2	description of shaking head (horse)
130	等趑	ɯɯʔ zlɯɯ	ɯɯʔ zɯɯ	təgx dzəg	təʔ dzə	deng3/cai2	suspiciously
131	忯懝	gje sle	gje sle	grjig stjig	gje sje	qi2/yi2	disregarded
132	顑頷	k^{h}ɯɯmʔ b ɾuumʔ	khoomʔ b-ruumʔ	khəmx ləmx	khomʔ ləmx	kan3/lan3	slim
133	鬒鱗	tjɯnʔ rɯnʔ	kjɯnʔ rɯnʔ	tiənx liənx	tjənʔ c-rənʔ	zhen3/lin4	description of few and scattered hairs
134	埃壒	qlɯɯ nɯɯs	qlɯɯ nɯɯs	ʔəg nəgh	ʔə nəs	ai2/nai4	foggy and dark
135	蒼黄	s^{h}aaŋ gwaaŋ	skhaaŋ g^{w}aaŋ	tshaŋ gwaŋ	tshaŋ g^{w}aŋ	cang1/ huang1	be greenish yellow
136	紛賱	p^{h}ɯn qunʔ	phɯn qunʔ	phjən ʔjənx	phjən ʔjunʔ	fen1/yun2	one by one
137	綢繆	t^{h}uuw mlɯw	thɯɯw [g]rɯw	thəgw mjəgw	thiw mjiw	chou2/ mou2	readily
138	蚴蟉	qɯw g ɾɯw	qrɯw [g]rɯw	ʔjiəgw ljəgw	ʔrjiw c-rjiw	you1/liu2	zigzagged
139	蟉虯	g ɾɯw grɯw	[g]rɯw grɯw	ljəgw gjiəgw	c-rjiw grjiw	liu2/qiu2	zigzagged
140	拘覷	ko shoo	ko skhoo	kjug tshjug	kjo tshjo	ju1/qu1	rough, not dense
141	韡孛	ɢwuds bɯɯds	ɢuds bɯɯds	wədh bədh	gwuts bəts	wei4/bei4	flourishing
142	槍攘	s^{h}aŋ njaŋ	skhaŋ njaŋ	tshjaŋ njaŋ	tshjaŋ njaŋ	qiang1/ rang3	disorderly
143	秧穰	qaŋ njaŋ	qaŋ njaŋ	ʔjaŋ njaŋ	ʔjaŋ njaŋ	yang1/ rang3	description of flourishing crops
144	扶疏	pa s·ŋra	pa sqra	pjag srjag	pja srjag	fu2/shu1	description of flourishing trees
145	窅窔	ʔ ɨmreew qeews	qmreew qeews	ʔragw ʔiagwh	ʔrew ʔews	yao3/yao4	deep and dark
146	窵窅	tɯɯws ʔ ɨmreew	tɯɯws qmreew	tiəgwh ʔragw	tiws ʔrew	diao4/yao3	deep, distant
147	佝瞀	q^{h}oos	qhoos	hugh	xos moks	kou4/mao4	ignorant

To be continued

		moogs	moogs	mugh			
148	萎蕤	qrol lal	qol [g]lal	ʔjiar rar	ʔrjoj ljaj	wei3/yi2	description of grass shaking and waving
149	逶迤	qrol lal	qol lal	ʔjiar rar	ʔrjoj ljaj	wei1/yi2	meandering
150	剌发	raad bood	raad bood	lat buat	c-rat bat	la4/ba2	unable to walk straight
151	趠趫	r tʰeewɢs lowɢ	rteewgs lowg	thragwh rakw	threwks ljawk	zhuo1/yue4	description of walking rapidly
152	齷齲	ʔ·sru ŋoo	skru ŋoo	tsrjəgw ŋug	tsrju ŋo	zou1/yu2	uneven (teeth)
153	倚移	qralʔ lal	qalʔ [g]lal	ʔjiarx rar	ʔrjajʔ ljaj	yi1/yi2	graceful and gentle
154	屠戾	zruub dɯb	sgruub dɯb	dzrəp drjəp	dzrup drjəp	qi4/zhe2	sneaky
155	莽沆	maaŋʔ gaaŋ	maaŋʔ gaaŋ	maŋx gaŋ	maŋʔ gaŋ	mang3/ hang4	vast (river)
156	儻莽	tʰaaŋʔ maaŋʔ		thaŋx maŋx	thaŋʔ maŋʔ	tang3/ mang3	broad
157	燕婉	qeen qonʔ	qeen qonʔ	ʔian ʔjuanx	ʔen ʔjonʔ	yan4/wan3	beautiful
158	虺隤	ŋʰuul l'uul	ŋ̊huul g ɬuul	hŋwəd dəd	hŋuj luj	hui3/tui2	sick
159	從容	zloŋ loŋ	shoŋ [g]loŋ	tshjuŋ grjuŋ	tshjoŋ ljoŋ	con1/rong2	at leisure, casually
160	望洋	maŋ ljaŋ	maŋ sɢlaŋ	mjaŋ rjaŋ	mjaŋ zljaŋ	wang4/ yang2	description of looking up
161	強陽	gaŋ laŋ	gaŋ laŋ	gjaŋ raŋ	gjaŋ ljaŋ	qiang2/ yang2	description of movement
162	強梁	gaŋ raŋ	gaŋ raŋ	gjaŋ ljaŋ	gjaŋ c-rjaŋ	qiang2/ liang2	fierce, intrepid
163	優遊	qu lu	qu lu	ʔjəgw rəgw	ʔju lju	you1/you2	at leisure
164	孟浪	mraaŋs raaŋ	mraaŋs raaŋ	mraŋh laŋ	mraŋs c-raŋ	meng4/ lang4	rash, impetuous
165	逍遙	sew lew	slew [g]low	sjagw grjagw	sjew ljaw	xiao1/yao2	free and unfettered
166	靡靡	s qʰwalʔ mralʔ	sqhʷalʔ mralʔ	swjarx mjiarx	swjajʔ mrjajʔ	huo4/mi2	description of soft grass
167	差池	sʰraal l'aal	skhraal laal	tshrar dar	tshraj daj	ci1/chi2	uneven, not uniform
168	翍翏	g ʁuub tʰuub	[g]ruub thuub	ləp thəp	g-rup hlup	la4/ta4	description of bird flying
169	崛峋	glud mɯd	glud mɯd	gwjət mjət	gjut mjət	jue2/wu4	towering

To be continued

170	蟠蜿	baan qoon	baan qoon	ban ʔuan	ban ʔon	pan2/wan2	winding
171	杳窈	ʔ ɯeewʔ ʔ ɬ'ɯɯwʔ	qmeewʔ qlɯɯwʔ	ʔiagwx tiəgwx	ʔewʔ tiwʔ	yao3/tiao3	deep, dark, profound
172	鞅掌	qaŋ tjaŋ	qaŋ tjaŋ	ʔjaŋ tjaŋx	ʔjaŋ tjaŋʔ	yang1/ zhang3	busy and tired

APPENDIX III: SPLITTING-SOUND WORDS IN OLD CHINESE

(Adapted from Sun 1999)

A Splitting-sound word in OC refers to an OC disyllable that is split from a one-syllable word: the onset of the monosyllabic word is reflected in the first syllable onset of the Splitting-sound word, and the rhyme of the monosyllabic word is preserved in the second syllable rhyme. The onset of the second syllable is a liquid *l- or *r- (Chapter 4). These words are attested as early as around 1000 BC, a little earlier than Diéyùn and Shuāngshēng words. The forms in the table below are the basis of the discussion of Splitting-sound words in this dissertation.

	Character	**Reconstructed OC (*)**				**Pinyin**	**Gloss**
	Nouns	**Zhengzhang**	**Pan**	**Baxter**	**Li**		
1	薏苡	qɯg lɯʔ	qɯg lɯʔ	ʔjək rəgx	ʔjək ljəʔ	yi1/yi3	Job's tears
2	蒺藜	zid riil	zid b ɾiil	dzjit lid	dzjit c-rij	ji2/li2	thorny caltrops
3	芙蓉	l^{h}uu loŋ	luu g·loŋ	thəgw grjuŋ	ph-lu ljoŋ	fu2/rong2	lotus
4	蒹芙	diil l'iig	diil liig	did dit	dij lit	ti2/die2	weeds
5	胡蝶	gaa l^{h}eeb	gaa k^{h} ɬeeb	gag thiap	ga help	hu2/die2	butterfly
6	蝮蜪	p^{h}ug l^{h}uu	p^{h}ug l̥uu	phjəkw thəgw	phjuk ph-lu	fu4/tao2	larval locust
7	蜦蝣	ga rag	ga rag	gjag ljak	gja c-rjak	qu2/lüe4	mayfly, *Ephemerida*
8	脊令	ʔ seg ren	seg m ɾen	tsjik ljan	tsjek c-rjen	ji2/ling4	waftail
9	鵜鶘	qwaa rlaag	qwaa rlaag	ʔwag drak	ʔ^{w}a lrak	wu1/ze2	pelican
10	鸜鵒	gwa log	gwa g ɬog	gwjag grjuk	g^{w}ja ljok	qu2/yu4	myna
11	伯勞	praag raaw	praag raaw	prak lagw	prak c-raw	bo2/lao2	shrike
12	髑髏	doog roo	doog [g]roo	duk lug	dok c-ro	du2/lou2	skull
13	頑顱	l'oog braa	loog raa	duk lag	log c-ra	duo2/lu2	skull
14	扶搖	pa lews	pa [g]lews	pjag grjagwh	pja ljews	fu2/yao2	whirlwind
15	屋漏	qoog roos	qoog roos	ʔuk lugh	ʔok c-ros	wu1/lou4	north-west corner inside

To be continued

							the room
16	擊蹶	keeg l'ee	keeg lee	kik dig	kek le	ji1/ti2	a kind of thin paper
17	瑚璉	gaa renʔ	gaa b ɾenʔ	gag ljanx	ga c- rjenʔ	hu2/lian3	kind of sacrificial vessel
18	隇陟	qul le	qul le	ʔwjəd rjig	ʔjuj lje	wei1/yi2	a place difficult of access
19	鈐鐫	gram l'oolʔ	gram loolʔ	gjiam duarx	grjam dojʔ	qian2/duo4	a kind of digging tool
20	鏌鋣	maag ljaa	maag lja	mak rjiag	mak zjᴀ	mo4/ye2	a sword name
	Verbs						
21	臾曳	gruds leds	gruds leds	gwjiədh radh	grjuts ljets	yu2/ye4	to tow
22	瞴婁	ma groo	ma g ɾoo	mjag lug	mja c- ro	wu3/lou2	to squint
23	嫪盈	geenʔ leŋ	geenʔ leŋ	gianx riŋ	genʔ ljeŋ	gui1/ying2	to scold
24	戲泄	qʰral lebs	qʰra g ɬebs	hjag rabh	xrja ljeps	xi4/xie4	to rest
25	熠燿	ɢrub lewɢs	ɢrub lewgs	gwjəp ragwh	wjup ljewks	yi4/yao4	sparkling
26	劬勞	go raw	go raaw	gjug lagw	gjo c- raw	qu2/lao2	sick, exhausted
27	狐剌	gwaa raad	gwaa raad	gwag lat	gʷa c- rat	hu2/la4	crooked
28	華離	qʰwraa reels	qʰwraa b ɾeels	hwrag liarh	hwra c- rejs	hua2/li2	crooked
29	豈弟	ŋʰɯɯlʔ diilʔ	kʰɯɯlʔ liilʔ	khədx didx	khəjʔ dijʔ	kai3/ti4	happy and harmonious
30	瓠落	gwlaa graag	gwaa g ɾaag	gwag lak	wa g- rak	hu4/luo4	large in volume
31	搖悅	lew lod	g ɬow g ɬod	grjagw ruat	ljaw ljot	yao2/yue4	very pleased with oneself
32	軥錄	koo rog	koo rog	kug ljuk	ko b- rjok	qu2/lu4	description of working with crooked arms; hard-working

APPENDIX IV: SPLITTING-SOUND WORDS IN MIDDLE CHINESE

(Adapted from Sun 1999)

A Splitting-sound word preserved in Middle Chinese refers to a disyllable that is split from a one-syllable word: the onset of the monosyllabic word is reflected in the first syllable onset of the Splitting-sound word, and the rhyme of the monosyllabic word is preserved in the second syllable rhyme. The onset of the second syllable is liquid a *l- (Chapter 4). These words are attested from 500AD to 1400AD.

Nouns		**Middle Chinese**	**Mono-syllabic form**	**Zheng-zhang**	**Pan**	**Li**	**Baxter**	**Gloss**
1	屈攣	*k^{h}yt lɥan	圈	*k^{h}ron	*k^{h}ron	*k^{h}wjan	*khrjon	circle
2	突欒	*tɦut lwan	團	*doon	*doon	*duan	*don	knot
3	突落	*tɦut lak	鐸	*l'aag	*laag	*dak	*lak	kind of bell
4	勃盧	*pɦut lua	蒲	*baa	*baa	*bag	*ba	rushes, reeds
5	勃籠	*pɦut ləwŋ	蓬	*booŋ	*booŋ	*boŋ	*buŋ	bitter fleabane
6	勃蘭	*pɦut lan	槃	*baan	*baan	*ban	*ban	tray
7	矻落	*k^{h}ut lak	角	*kroog	*kroog	*kruk	*krok	corner
8	滴寧	*tɛjk nɛjŋ	頂	*teeŋʔ	*teeŋʔ	*tiŋx	*teŋʔ	top
9	窟駝	*k^{h}ut tɦa	窠	*k^{h}lool	*k^{h}ool	*k^{h}uar	*khoj	nest, burrow
Verbs								
10	骨露	*kut lua	錮	*kaags	*kaags	*kagh	*kaks	run metal into cracks
Adjectives								
11	步廊	*pɦwa laŋ	旁	*baaŋ	*baaŋ	*baŋ	*baŋ	vast, boundless
12	即零	*tsjak ljaŋ	精	*ʔsleŋ	*skeŋ	*tsjiŋ	*tsjeŋ	smart, astute

APPENDIX V: SPLITTING-SOUND WORDS IN MODERN CHINESE DIALECTS

A Splitting-sound word preserved in Colloquial Modern Chinese dialects refers to a disyllable that is split from a one-syllable word: the onset of the monosyllabic word is reflected in the first syllable onset of the Splitting-sound word, and the rhyme of the monosyllabic word is preserved in the second syllable rhyme. The onset of the second syllable is a liquid l- (Chapter 4).

Taiyun (Jin Dialect) (Adapted from Zhao, B. 1979)			**Gaizhu Dialect (North Min Dialect)** (Adapted from Deng 2007)				**OC Zhengzhang**	**Gloss**
Splitting-sound form		**Monosyl-labic form**	**Splitting-sound form**		**Mono-syllabic form**			
薄拉	pəʔ$_{8}$la$_{1}$	爬 pa$_{1}$	爬□	pa$_{2}$la$_{2}$	爬	pa$_{2}$	*braa	‘to crawl’
圪老	kəʔ$_{8}$lɔ$_{3}$	攪 kjɔ$_{3}$	攪□	ko$_{3}$lo$_{3}$	攪	ko$_{3}$	*kruuʔ	‘to stir’
扑涝	pəʔ$_{8}$lau$_{3}$	跑 p^{h}au$_{3}$	抛□	p^{h}o$_{1}$lo$_{1}$	抛	p^{h}o$_{1}$	*bruu	‘to run’
圪佢	kəʔ$_{8}$lɔ$_{1}$	角 kiɔ$_{1}$	角□	ku$_{7}$lu$_{7}$	角	ku$_{7}$	*kroog	‘corner’
骨拢	kwəʔ$_{8}$luŋ$_{3}$	滾 kuŋ$_{3}$	滾□	kwaŋ$_{3}$lwaŋ$_{3}$	滾	kwaŋ$_{3}$	*kluunʔ	‘to roll’
泼棱	p^{h}əʔ$_{8}$ləŋ$_{1}$	蓬 p^{h}əŋ$_{1}$	蓬□	p^{h}aŋ$_{1}$laŋ$_{1}$			*blooŋ	‘classifier’
窟窿	k^{h}əʔ$_{8}$luŋ$_{1}$	孔 k^{h}uŋ$_{1}$	空窿	k^{h}aŋ$_{1}$laŋ$_{1}$			*k^{h}looŋʔ	‘hole’
圪离	kəʔ$_{8}$li1		胳肢	ku$_{5}$li$_{5}$			*klaag	‘armpit’
刻撩	k^{h}əʔ$_{8}$ljau$_{1}$	翹 k^{h}jau$_{1}$	翹□	k^{h}jau$_{5}$ljau$_{5}$	翹	k^{h}jau$_{5}$	*kloŋʔ	‘to arch’
骨聯	kwəʔ$_{8}$lɥɛ$_{7}$	卷 kɥɛ$_{7}$	□□	ki$_{7}$li$_{7}$			*glonʔ	‘to curl’

Fuzhou Dialect (North Min) (Adapted from Liang, Y. 1982)

Splitting-sound form	**Monosyllabic form**	**OC (Zhengzhang's reconstruction)**	**Gloss**
k^hje_{11} ljɛu$_{213}$	翹 $k^hjɛu_{213}$	*kloŋʔ	'to be placed upside down'
ku$_{31}$ luŋ$_{31}$	滾 kuŋ$_{31}$	*kluunʔ	'to roll about, toss about'
kɛ$_{31}$ lɛiʔ$_4$	夾 kɛiʔ$_4$	*kreeb	'to press from both sides'
kwo$_{31}$ lwoŋ$_{31}$	卷 kwoŋ$_{41}$	*kron	'to roll up'
hu$_{11}$ louʔ$_{23}$	拂 houʔ$_{23}$	*p^hɯd	'to skip over'
so$_{11}$ lɔʔ$_{23}$	□ sɔʔ$_{23}$		'bind round tightly'
t^hu_{31} lu$_{31}$	吐 t^hu_{31}	*l^haaʔ	'to protrude'
a$_{31}$ lau$_{44}$	凹 au$_{44}$	*qriiw	'concave, dented'
li$_{11}$ lɛiʔ$_{23}$	□ lɛiʔ$_{23}$		'speed by'
p^ho_{31} louŋ$_{44}$	□ p^houŋ$_{44}$		'the sound of dropping into water'
ku$_{31}$ luŋ$_{52}$	群 kuŋ$_{44}$	*glun	'group'
t^hwa_{31} lwa$_{44}$	拖 t^hwa_{44}	*l^haal	'a string of'

Shunping Dialect (Mandarin) (Adapted from Sun 1999:127)

Splitting-sound form	**Monosyllabic form**	**OC (Zhengzhang's reconstruction)**	**Gloss**
$k^h ə_{214}$ ləw_0	口 $k^h ə w_{214}$ 'mouth'	*k^hooʔ	'back area inside of shoes'
ta$_{55}$ lian$_0$	掂 tjan$_{55}$ 'to carry or lift'	*tiim	'a long bag sewn up at both ends with an opening in the middle'
$t^h u_{55}$ lwo$_0$	脫 t^hwo$_{55}$ 'to take off'	*l^hood	'to hang loose'
xu$_{11}$ lwo$_0$	和 xwo$_{11}$ 'to mix flour with water'	*gool	'to sweep flour with broom or hand '
$t^h a_{55}$ la$_0$	拖 t^hwo$_{55}$ 'drag'	*l^haal	'to wear shoes with the back turned in'
xu$_{11}$ lan$_0$	環 xwan$_{11}$ 'hoop'	*gwraan	'an area around the speaker'

Taiyuan Dialect (Jin) (Adapted from Zhao, B. 1979 and Sun 1999)

Splitting-sound form	**Monosyllabic form**	**OC (Zhengzhang's reconstruction)**	**Gloss**
kwəʔ$_{54}$ lə$_{53}$	裹 kwɤ$_{53}$	*klool?	'to wrap'
kwəʔ$_{54}$ la$_0$	刮 kwaʔ$_2$	*krood	'to scrape'
məʔ$_{54}$ ləŋ$_{11}$	蒙 məŋ$_{11}$	*mooŋ	'to cover'
ts^həʔ$_{54}$ la$_2$	擦 ts^haʔ$_2$ 'to wipe'	*s^hlaad	'to rub against something to remove dirt '
$t^h a_{55}$ la$_0$	拖 t^hwo$_{55}$ 'drag'	*l^haal	'to wear shoes with the back turned in'
t^hwəʔ$_{54}$ l ã$_{11}$	團 t^h w ã$_{11}$ 'something like a ball'	*doon	'measure word for ball'

Pingyao Dialect (Jin) (Adapted from Hou, J. 1989)

Splitting-sound form	**Monosyllabic form**	**OC (Zhengzhang's reconstruction)**	**Gloss**
pʌʔ$_{54}$ læ$_{53}$	擺 pæ$_{53}$	*preelʔ	'to swing'
tjʌʔ$_{54}$ li$_{13}$	提 ti$_{13}$	*dee	'to lift'
kʌʔ$_{54}$ ljɔ$_{13}$	翹 tɕjɔ$_{13}$	*gew	'to lift up (one leg)'
kʌʔ$_{54}$ ləŋ$_{53}$	埂 kəŋ$_{53}$ 'a low bank of earth between fields'	*kraaŋ	'slope between terraced fields '
pʌʔ$_{54}$ li$_{53}$	秕 pi$_{53}$	*piʔ	'blighted grain'

Yimeng Dialect (Jin) (Adapted from Li, Z. 1991)

Splitting-sound form	**Monosyllabic form**	**OC (Zhengzhang system)**	**Gloss**
p^{h}əʔ$_{21}$ lan$_{44}$	盤 p^{h}an$_{44}$	*blaan	'dish'
t^{h}uʔ$_{21}$ luŋ$_{44}$	臀 t^{h}uŋ$_{44}$	*duun	'hip'
p^{h}əʔ$_{21}$ liau$_{53}$	漂 p^{h}jau$_{53}$	*p^{h}ew	'to rinse'
pəʔ$_{54}$ lən$_{53}$	笨 pən$_{53}$	*pɯɯnʔ	'stupid '

APPENDIX VI: ONSET COMPARISON OF 55 BASIC STABLE VOCABULARY ITEMS FROM OLD CHINESE, WITH MIDDLE CHINESE AND MODERN CHINESE REFLEXES

Reconstructed Old Chinese (*)		**Reconstructed Middle Chinese (*)**	**Modern Standard Chinese**	**Characters**	**Gloss**
	Pan system	**Pan system**	**In Pinyin**		
1	**srid**	dʑit	shī	虱（蝨）	louse
2	**njis**	ɲiih	èr	二	two
3	**q^{h}wlji?**	ɕɥix	shuǐ	水	water
4	**mljɯ?**	ɲix	ěr	耳	ear
5	maŋ	maŋ	wáng	亡	die
6	la/ŋaa	jo/ŋo	yǘ/wǔ	余（予）/吾	I
7	kaan	kan	gān	肝	liver
8	**ŋ-grɯɯn?**	mjuk	mù	目	eye
9	**m-l̥ɯw?**	sju	shǒu	手	hand
10	mɯn	mjon	wén	闻	hear
11	**djo?/moog**	dʑyoh/muk	shù/mù	树/木	tree
12	ŋa	ŋio	yǘ	鱼	fish
13	meŋ	mieŋ	míng	名	name
14	**gljag**	dʑek	shí	石	stone
15	**ŋ-raa/k^{h}jɯ?**	ŋa/tɕhix	yá/chǐ	牙/齿	tooth
16	**njo?**	ɲɥox	rǔ	乳	breasts
17	baa	ɲex	ěr	爾	you
18	luu?/keeŋs	dau/keŋ	dào/jìng	道/径	path

To be continued

19	kuud	kwot	gǔ	骨	bone
20	**grood**	ʑet	sh é	舌	tongue
21	pa	pju	fū	肤	skin
22	**[g]lags**	jah	y è	夜	night
23	**q^{h}ljeb**	jiep	y è	叶	leaf
24	**ɢwaʔ**	ɦyox	yǔ	雨	rain
25	**sreeds**	ʂet	shā	杀	kill
26	q^{h}wiig	huet	x è	血	blood
27	**kroog**	luk	jǐao	角	horn
28	**njiŋ**	ɲin	r én	人	person
29	**slig**	sit	xī	膝	knee
30	**qlig**	ʔjit	yī	一	one
31	**blids**	bi	b í	鼻	nose
32	**pruuʔ/moonʔ**	brau/mwan	bǎo/mǎn	饱/满	full
33	rɯɯ	lai	l ái	来	come
34	**s-qeeŋ**	seŋ	xīng	星	star
35	**sreen**	sen	shān	山	mountain
36	**q^{h}waalʔ**	hwa	hǔo	火	fire
37	ŋaa	ŋo	wǔ	吾	we
38	**qrɯmʔ**	ʔimx	y ìn	饮	drink
39	keens	kenh	j ìan	见	see
40	bods	pɥoi	f èi	吠	bark
41	**sliŋ**	sin	xīng	新	new
42	**k^{h}weenʔ**	k^{h}wen	quǎn	犬	dog

To be continued

43	**mljig**	ɲit	r ì	日	sun
44	pɯl	pyoi	fēi	飞	fly
45	**kji**	tɕji	zhǐ	脂	grease
46	ŋod	ŋɥat	y ùe	月	moon
47	laʔ	jox	yū	与	give
48	**slɯm**	sim	xīn	心	heart
49	**ɢwaʔ**	ɦɥox	yǔ	羽	feather
50	**braag**	bak	b ái	白	white
Number of words		55	55	55	
Number of words with CC onset		1	32	0	
Percentage of words with CC		1.81%	58.18%	0	

Note: Words in bold contain consonant-cluster onsets.

BIBLIOGRAPHY

Acordia, Giorgio Francesco. (2007). Chinese: A Language of Compound Words? In F. Montermini, G. Boyé & N. Hathout (Ed.), Selected Proceedings of the 5th Décembrettes: *Morphology in Toulouse*. (pp. 79-90). Somerville, MA: Cascadilla Proceedings Project.

Anderson, S. R. (1985). Typological distinctions in word-formation. In Shopen, T. (Eds). *Language typology and syntactic description*, Volume III Grammatical categories in the lexicon. New York: Cambridge. pp. 3-56.

Baxter, William H. (1992). *A Handbook of Old Chinese Phonology*. Berlin: Mouton de Gruyter.

Baxter, William & Laurent Sagart. (1997). "Word formation in Old Chinese". In J. Packard (Ed.), *New approaches to Chinese word formation*. Berlin: Mouton de Gruyter, 35–76.

Bickel & Nichols. (2007). Inflectional Morphology. In T. Shopen (Ed.), *Language typology and syntactic description*. Cambridge: Cambridge University Press.

Blevins, Juliette. (2004). *Evolutionary phonology: the emergence of sound patterns.* Cambridge: Cambridge University Press.

Blevins, J. & Wedel, A. (2009). Inhibited sound change: An evolutionary approach to lexical competition. *Diachronica* 26:2, pp. 143-183.

Blust, R. (2007). Disyllabic attractors and anti-antigemination in Austronesian sound change. *Phonology* 24, pp. 1-36.

Blench, Roger. (2004). Stratification in the peopling of China: how far does the linguistic evidence match genetics and archaeology? Paper for the Symposium "Human migrations in continental East Asia and Taiwan: genetic, linguistic and archaeological evidence". Geneva: Université de Genève.

Bodman, Nicholas C. (1980). Proto-Chinese and Sino-Tibetan: data towards establishing the nature of the relationship. In van COETSEM, Frans & WAUGH, Linda R. (Ed.), *Contributions to historical linguistics: issues and materials*. Leiden: E.J. Brill. pp. 34-199.

Boltz, William G. (2002). Zhengzhang Shangfang: The phonological system of Old Chinese. *Cahiers de linguistique – Asie orientale*, vol. 31: 105-116.

Campbell, Lyle (2004). *Historical Linguistics: an Introduction* (2nd edition). Edinburgh: Edinburgh University Press, and Cambridge, MA: MIT Press.

Casacchia, G. (2006). "Chinese Linguistic Tradition". *Encyclopedia of Language & Linguistics*. Second Edition. London: Elsevier.

Chao, Y. -R. 趙元任. (1930). A system of tone letters. *Le Maitre Phonetique* 45: 24-7.

Chen, Jiewen 陳潔雯. (1984). Shangguyin Fushengmu: Yuefangyan yigeban yinjie de zi suo tigong de zuozheng 上古音復聲母:粵方言一個半音節的字所提供的佐證 [Onset clusters in Old Chinese: evidence from the sesqui-syllabic words in Yue language]. *Fangyan*. Vol. 4: 300-313.
Chen, Kejiong 陳克炯. (1978). Zuozhuan Fuyin ci Chutan. [Attempted Studies of Disyllables in Zuozhuan]. *Journal of Huazhong Normal College*. Vol. 4: 233-241.

Chen, Matthew Y. (2000). *Tone Sandhi: Patterns across Chinese Dialects.* Cambridge, UK: Cambridge University Press.

Cheng, Xiangqing 程湘清. (1981). Shilun shanggu hanyu shuangyinci he shuangyin cizu de qufen biaozhun 試論上古漢語雙音詞和雙音詞組的區分標準. [Tentative analysis of the distinction between disyllables and disyllabic phrases in Old Chinese]. *Dongyue luncong*. Vol. 4: 44-53.

Cheng, Xiangqing 程湘清. (1992). *Xianqin Hanyu Yanjiu*. [Studies on Chinese of Pre-Qin]. Jinan: Shandong Educational Publishing House.

Chen, Yan 陳燕. (1992). Guangyun Shuangsheng Dieyun Lianmianzi de yuyin yuanjiu 廣韻雙聲疊韻連綿字的語音研究. [Phonological study of Lianmian words in Guangyun]. *YYXLC* 17: 60-121.

Coblin, W. South. (1986). *A Sinologist's Handlist of Sino-Tibetan Lexical Comparisons, Momumenta Serica monograph series* XVIII, Nettetal: Steyler.

DeFrancis, J. (1984). *The Chinese Language: Fact and Fantasy*. Hawaii: University of Hawaii Press.

DeLancey, S. (2011). On the origins of Sinitic. In Zhuo Jing-Schmidt (Ed.), *Proceedings of the 23rd North American Conference on Chinese Linguistics* (NACCL-23), Vol. 1, pp. 51-64, Eugene: University of Oregon.

Deng, Xiangzhang 鄧享璋. (2007). Minbei Minzhong fangyan Fenyinci de xingzhi yu laiyuan. *Yuwen Yanjiu* [Chinese Studies]. Vol.102: 61-64.

Ding, Bangxin 丁邦新. (1987). The syllable structure of Old Chinese. *Journal of Historical Linguistics Institute*. 50-4: 717-739.

Dixon and Aikhenvald. (2002). Word: A typological framework. *Word: A cross-linguistic typology*, ed. by Dixon and Aikhenvald, 1-41. Cambridge: Cambridge University Press.

van Driem, George. (1998). Neolithic correlates of ancient Tibeto-Burman migrations. In Roger Blench and Matthew Spriggs, (Ed), *Archaeology and Language II*. London: Routledge. pp. 67-102.

van Driem, George. (2005). Sino-Austronesian vs. Sino-Caucasian, Sino-Bodic vs. Sino-Tibetan, and Tibeto-Burman as Default theory. In Yogendra Prasada Yadava, Govinda Bhattarai, Ram Raj Lohani, Balaram Prasain & Krishna Parajuli, (Ed.), *Contemporary Issues in Nepalese Linguistics*. Kathmandu: Linguistic Society of Nepal. pp. 285-338.

Duanmu, San 端木三. (1999a). Metrical Structure and Tone: Evidence from Mandarin and Shanghai. *Journal of East Asian Linguistics* 8: pp. 1-38.

Duanmu, San 端木三. (1999b). Stress and the Development of Disyllabic Words in Chinese. *Diachronica* XVI: 1.1-35.

Duanmu, San & Zhang, Huili. (2010). Phonology and the Chinese Lexicon. Presented on CUNY Word Conference.

Feng, Shengli 馮勝利. (1998). Prosodic structure and compound words in Classical Chinese, In J.L. Packard, (Ed.), *New Approaches to Chinese Word Formation*. Berlin – New York: Mouton de Gruyter. pp. 197-260.

Feng, Shengli 馮勝利. (2001). Lun hanyuci de duowei xing 論漢語詞的多維性 [On the multidimensionality of "word" in Chinese]. *Dangdai Yuyanxue [Contemporary Linguistics]* 3, 161-174.

Feng, zhiwei 馮志伟. (2004). The Semantic Loanwords and Phonemic Loanwords in Chinese Language Proceedings for 11th International Symposium of NIJLA. Tokyo: 200–229.

Gernet, Jacques. (1996). *A History of Chinese Civilization* (Second ed.). Cambridge: Cambridge University Press

Guo, Shaoyu 郭紹虞. (1938). Zhongguo yuci zhi tanxing zuoyong [The elastic function of Chinese word length]. *Yen Ching Hsueh Pao* 24. (Reprinted in Shao-yu Guo 1963, 1-40).

Guo, Xiaowu.郭小武. (1993). 試論疊韻聯綿字的統諧規律 *Zhongguoyuwen* Vol. 3: 209.

Haudricourt, A.-G. (1954). Comment reconstruire le chinois archaique. *Word*, 10, 351-364.

Hoosain, R. (1992). Psychological reality of the word in Chinese. In H. C. Chen & O. J. L. Tzeng, (Ed.), *Language processing in Chinese*, 111-130. Amsterdam: North-Holland.

Hou, Jingyi 侯精一. (1989). *Jinyu yanjiu*. Tokyo: Tokyo University of Foreign Studies.

Hu, Yilu 胡以魯. (1923). *Beginning Mandarin* (Guoyuxue Caochuang 国语学草创). Shanghai: Commercial Press.

Jacob, J.M. (1979). Observations on the uses of reduplication as a poetic device in Khmer. In T.L. Thongkum, V. Panupong, P. Kullavanijava and M.R. Kalaya Tingsabadh (Ed.), *Studies in Tai*

and Mon-Khmer Phonetics and Phonology In Honour of Eugénie J.A. Henderson, 111-130. Chulalongkorn University Press.

Janssens, G. (1972). *Stress in Arabic and Word Structure in the Modern Arabic Dialects*. Peeters: Belgium.

Jeon, Kwang-jie 全廣鎮. (1996). *Hanzangyu tongyuanci zongtan*. [Etymological studies of Sino-Tibetan cognate words], Taipei: Taiwan xuesheng shuju.

Jin, Lixin 金理新. (2002). *Shanggu Hanyu Yinxi*. [Old Chinese Phonology]. Anhui: Huangshan Publishing House.

Kalgren, B. (1923). *Analytic Dictionary of Chinese and Sino-Japanese*. Paris. Reprinted in New York: Dover Publications, Inc., 1974.

Karlgren, B. (1949). *The Chinese Language, an essay on its nature and history*. New York: Ronald Press Co.

Karlgren, B. (1954). *Compendium of Phonetics in Ancient and Archaic Chinese*, Museum of Far Eastern Antiquities, Stockholm.

Kosaka, R. (2007). Consonantal lenition and vocalic transfer in Kadai languages: with disyllabic proto-Be-Tai as supporting evidence. In I. Shoichi, A. Simpson, K. Adams and P. Sidwell (Ed.), *SEALS XIII Papers from the 13th Annual Meeting of the Southeast Asian Linguistics Society 2003*, Canberra, Australia, (pp. 93-103). Pacific Linguistics, Research School of Pacific and Asian Studies, the Australian National University.

Lewis, M. Paul (Ed.). (2009). Ethnologue: Languages of the World, Sixteenth edition. Dallas, Tex.: SIL International. Retrieved Dec. 27, 2012, Online version: <http://www.ethnologue.com>.

Li, Fang-Kuei 李方桂. (1980). *Shangguyin Yanjiu* 上古音研究. [A study of Old Chinese phonology]. Beijing: Shangwu Chubanshe [Commercial Press].

Li, Shichun 李仕春. (2007). Cong Fuyinci Shuju Kan Zhonggu Hanyu Cihui de Fanzhan. [Analysis of Middle Chinese Morphology based on disyllable data]. *Journal of Ningxia University*. Vol.29, No.3: 1-8.

Li, Charles N. & Thompson, Sadra A. (1981). Mandarin Chinese. A Functional Reference Grammar. Berkeley: Univerisity of California Press.

Li, Yu 李玉. (1999). Guangxi pingnan minnanhua de yuyin xitong ji tongyi zihui. [The phonology of Pingnan dialect of South Min in Guangxi and the word list]. *Guangxi shiyuan xuebao* 1999, vol.1: 50-57.

Liang, Yuzhang 梁玉璋. (1982). Fuzhou Fangyan de Qiejiaoci. [Splitting-sound words in Fuzhou Dialect]. *Fangyan* 1:37- 46.

Lin, Hua 林華(2001). *A grammar of Mandarin Chinese*. München: LINCOM EUROPA.

Lin, Yen-hwei. (1989). *Autosegmental treatment of segmental processes in Chinese phonology*. Unpulished Doctoral Dissertation, University of Texas, Austin.

Li, Zhiguo 李治國. (1991). Yimeng Fangyan de Fenyinci. [Splitting-sound words in Yimeng Dialect]. *Fangyan* 3: 206- 210.

Lü, Shuxiang呂叔湘. (1963). "Xiandai Hanyu dan shuang yinjie wenti chu tan" [A preliminary study of the problem of monosyllabism and disyllabism in modern Chinese]. *Zhongguo Yuwen* 1963.1:11-23.

McCarthy, J. & Prince, A. (1993). *Prosodic Morphology I – Constraint Interaction and Satisfaction*. University of Massachusetts and Rutgers University.

Matisoff, James. (1973). Tonogenesis in Southeast Asia. In L. M. Hyman, ed. *Consonant types & tones*. Los Angeles: The Linguistic Program, University of Southern California. 71-95.

Matisoff, James. (2003). *Handbook of Proto-Tibeto-Burman*: *system and philosophy of Sino-Tibetan Reconstruction.* University of California Press.

Mei, T.-l. (1970). Tones and prosody in Middle Chinese and the origin of the Rising tone. *Harvard Journal of Asian Studies*, 30: 86-100.

Mei, T.-l. 梅祖麟. (1980). Sisheng bieyi zhong de shijian cengci 四聲別義中的時間層次 *Zhongguo Yuwen*, Vol. 6. 427-433.

Mei, T.-l. 梅祖麟. (1989). Shanggu hanyu s- qianzhui de gouci gongneng shitan 上古漢語 s- 前綴的構詞功能試探, *Zhongyanyuan dier jie hanxue huiyi gunwenji*. 23-51. Taipei: Zhongyang yanjiuyuan

Miao, Ruiqin 苗瑞琴. (2005). *Loanword adaptation in Mandarin Chinese: Perceptual, phonological and sociolinguistic factor*. Ph. D. dissertation, Stony Brook University, Stony Brook, New York.

Miyake, Marc. (2001). Laurent Sagart : The roots of Old Chinese. *Cahiers de linguistique - Asie orientale*, vol. 30: 257-268.

Moravcsik, Edith, A. (1978). Reduplicative constructions. *Universals of Human Language, vol. 3: Word structure*, Joseph H. Greenberg, ed., 297-334. Stanford: Stanford University Press.

Nhàn, Ngô Thanh. (1984). *The syllabeme and patterns of word formation in Vietnamese*. Ph.D dissertation, New York University.

Nie, Hongyin 聶鴻音. (1987). *Zhongshang Gu Hanyu Yinyun Gangyao*, translated from *Compendium of Phonetics in Ancient and Archaic Chinese*, by Karlgren, B. Qilu Publishing House: Ji'nan.

Norman, Jerry. (1988). *Chinese*. Cambridge University Press.

Norman, J. & Mei, T. (1976). The Austroasiatics in ancient south China: Some lexical evidence. *Monumenta Serica* 32:274-301.

Packard, Jerome L. (1998). *New Approaches to Chinese Word Formation – Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*. Mouton de Gruyter: New York.

Packard, Jerome L. (2000). *The Morphology of Chinese*: A Linguistic and Cognitive Approach. Cambridge University Press.

Pan, Weishui 潘渭水. (1999). Fujian jianou niaoyu weitan 福建建甌"鳥語"探微. *Zhongguo Yuwen*, 3, 205-207.

Pan, Wuyun 潘悟雲. (1999). Hanzangyu zhong de Ciyao Yinjie. [Sesqui-syllables in Sino-Tibetan]. *Zhongguo yuyanxue de xintuozhan*. Hong Kong: The City University of Hong Kong Press.

Pan, Wuyun 潘悟雲. (2000). *Hanyu lishi yinyunxue* 漢語歷史音韻學. [Chinese historical phonology]. Shanghai: Shanghai Education Press.

Pan, Wuyun 潘悟雲. (2002). Dui huaao yuxi jiashuo de ruogan zhichi cailiao, [Several supporting evidences for the Sino-Austric Family hypothesis]. *Pan wuyun zixuanji*. [Selected works of Pan, Wuyun]. Hefei: Anhui Publishing House.

Pan, Yunzhong 潘允中. (1989). *Hanyu Cihuishi Ganyao*. [Overview of the history of Chinese vocabulary]. Shanghai: Shanghai Guji Publishing House.

Pittayaporn, Pittayawat. (2009). *The phonology of Proto-Tai*. Unpublished Ph. D. dissertation, Cornell University, Ithaca, New York.

Pulleyblank, E. G. (1962). The consonantal system of Old Chinese. *Asia Major*, 2(9), 58-144 & 206-265.

Pulleyblank, E.G. (1984). *Middle Chinese: A study in historical phonology*, Vancouver: University of Bristish Columbia Press.

Pulleyblank, E.G. (1992). How do we reconstruct Old Chinese? *Journal of American Oriental Society*, 112, 365-382.

Purnel, H. C. (1988). Youmian minyaojian geyao de yunlǜjiegou “優勉”民謠間歌謠的韻律結構 [The prosodic structure of the You-Mien ballads]. 1986 International Symposium on Yao Nationality 瑤族研究論文集, Mingzu Publishing House.

Pi, Hongming 皮鴻鳴. (1992). Hanyu Cihui Shuangyinha Yanbian de Xingzhi he Yiyi, *Studies on Old Chinese*, 14, 58-61.

Qiu, Bing 邱冰. (2010). Zhonggu Hanyu Cihui Shuangyinhua Yanjiu. [On disyllabicity in Middle Chinese vocabulary]. *Journal of Yanshan University*. 11(1), 30-33.

Sagart, L. (1995). Some remarks on the ancestry of Chinese. In Wang (Ed.) The Ancestry of Chinese Language, *Journal of Chinese Linguistics monograph series* No. 8 (1995), pp. 195-223.

Sagart, Laurent. (1999). *The roots of Old Chinese*. John Benjamins: Amsterdam and Philadelphia.

Schessler, A. (1974). R and L in archaic Chinese, *Journal of Chinese Linguistics* (Berkeley), 2(2), 186-199.

Schessler, A. (2000). Book review: The roots of Old Chinese, *Language and Linguistics* 1(2), 257-267.

Selkirk, E. (1984). On the major class features and syllable theory. In Aronoff & Oehrle (Eds.) *Language sound structure: Studies in phonology*, 107-136. Cambridge: MIT Press.

Serruys, Paul L-M. & Yang, Hsiung. (1959). *The Chinese dialects of Han time according to Fangyen.* Berkeley: University of California Press.

Shangguyin Chaxun, Retrieved Feb. 20, 2011, from *Comparative Linguistics EISU, Eastling Web site*: <http://www.eastling.org/oc/oldage.aspx>

Shi, Yuzhi 石毓智. (2002). *The establishment of Modern Chinese grammar: The formation of the resultative construction and its effects*. Amsterdam – Philadelphia: John Benjamins.

Shih, Chi-lin 石基琳. (1986). The prosodic domain of tone sandhi in Chinese. Doctoral dissertation, University of California, San Diego.

Shih, Chilin 石基琳. (1997). Mandarin Third Tone Sandhi and Prosodic Structure. In J. Wang and N. Smith (Eds). *Studies in Chinese phonology*, 81-123. Mouton de Gruyter.

Shorto, Harry, et al. (2006). *A Mon-Kmer comparative dictionary*. In Paul Sidwell (Ed.), *Pacific Linguistics* (Vol. 579). Canberra: Australian National University. Accessed via SEAlang database <http://www.sealang.net/monkhmer/> on 03-23-2011.

Sun, Hongkai 孫宏開. (1982). *Dulongyu jianzhi* 独龙语简志 [Description of De’ang language]. Minzu Publishing House.

Sun, H. 孫宏開. (2001). Yuanshi hanzangyu zhong de jieyin wenti 原始汉藏语中的介音问题 [The medial problem in Proto-Sino-Tibetan]. Paper presented in the 34th International Conference of Sino-Tibetan. Kunming, China.

Sun, J. 孫景濤. (1999). *Reduplication in Old Chinese*. Ph.D dissertation. Vancouver: The University of British Columbia.

Tang, Z. 唐作藩. (2011). *Hanyu yuyinshi jiaocheng* 漢語語音史教程 [A course in the history of Chinese phonology]. Beijing: Publishing House of Beijing University.

Thompson, Becky Butler. (2010). The sesqui-syllable as a disyllabic word, *CUNY Conference on the Word in Phonology*. New York.

Trommelen, M. & Zonneveld, W. (1989). *Klemtoon en metrische fonologie*. Muiden: Coutinho.

Ting, Pan-Hsin 丁邦新. (1979). Shanggu Hanyu de Yinjie Jiegou [The syllable structure in Archaic Chinese], *Bulletin of the Institute of History and Philology*, 50. Taipei: Institute of History and Philology, Academia Sinica: 717-739.

Vaux, B. & Wolfe, A. (2009). The appendix. In Raimy, E. & Cairns, C. (Ed.), *Comtemporary views on architecture and representations in Phonology*, pp. 101-143. Cambridge, MA: MIT Press.

Wang, Fusheng 王福生. (1998). *La formazione delle parole nella lingua cinese contemporanea.* Trieste: Universit à degli studi di Trieste – Scuola Superiore di Lingue.

Wang, Guowei 王國維. (1923). Lianmian Zipu 連綿字譜. Reprinted in *Wang Guowei xiansheng quanji* 王国维先生全集, 9, 3449-3518. Taipei: Wenhua Publishing House.

Wang, Jingliu 王敬騮(1984). Wayu shuyu 佤語熟語匯譯[Translation of idioms in Wa language]. *Minzu diaocha yanjiu*. Vol. 2.

Wang, Li 王力. (1944). *Zhongguo yufa lilun*, [Chinese Grammar Theory] Beijing: Shangwu Yinshu Guan.

Wang, Li 王力. (1953). *Zhongguo xiandai yufa*, [Chinese Modern Grammar] Beijing: Zhonghua Shuju.

Wang, Li 王力. (1972). *Hanyu yinyun xue*, [Chinese Phonology] Beijing: Zhonghua Shuju.

Wang, Li 王力. (1980). *Hanyu shigao* [A historical grammar of Chinese]. Beijing: Zhonghua Shuju.

Wang, Shaoxin 王紹新. (1992). Jiagu Keci Shidai de Cihui, [Vocabulary on Oracle Bones]. 1-21. In Cheng, Xiangqing (Eds.), *Xianqin hanyu yanjiu*, [Studies on Chinese of Pre-Qin]. Jinan: Shandong Educational Publishing House.

Wedel, A. (2004). Category competition drives contrast maintenance within an exemplar-based production/perception loop. In Goldsmith, J. & Wicentowski, R., (Eds.), *Proceedings of the Seventh Meeting of the ACL Special Interest Group in Computational Phonology* 7: 1-10.

Wedel, A. (2007). Feedback and regularity in the lexicon. *Phonology* (2007), 24: 147-185.

van de Weijer, J. & Zhang, Jisheng 張吉生. (2008). An X-bar approach to the syllable structure of Mandarin. *Lingua* 118, 1416–1428.

Wichmann, Søren & the ASJP Consortium. (2009). "State of the art of the ASJP project". The Swadesh Centenary Conference, MPI-EVA, Jan. 17-18, 2009.

Xiang, Xi 向熹. (1987). *Shijing yuyan yanjiu* [The study of the languages in Shijing]. Chengdu: Sichuan Renmin Publishing House.

Xu, Baohua 許寶華 & Pan, Wuyun 潘悟雲. (1994). Shi erdeng. 釋二等 [Explanations of erdengyun]. *Yinyunxue yanjiu*. Vol.3. Zhonghua Publishing House.

Xu, Zhaohua 徐紹華. (2003). *Shanggu Hanyu Cihui Shi*. [The history of Old Chinese lexicon]. Beijing: Shangwu Jinshu Guan.

Yip, M. (1992). Prosodic morphology in four Chinese dialects. *Journal of East Asia Linguistics* 1(1), 1-35.

Yu, Min 俞敏. (1985). Dengyun Suyuan. [The Origin of Rhymes in Deng]. *Studies of Phonology* 2, Beijing: Zhonghua Shuju.

Yu, Nai-yong 余乃永. (1985). *Shang Gu Yinxi Yanjiu*, [Study of the Old Chinese sound system]. Hong Kong: The Chinese University Press.

Yu, Suisheng 喻遂生 & Guo Li 郭力. (1987). Shuowen jiezi de fuyinci 說文解字的復音詞. *Xi'an shifan daxue xuebao* 1:123-136.

Zaharani, Ahmad. (2005). *Phonology-Morphology Interface in Malay: an Optimality Theoretic Account*. Pacific Linguistics: Australia National University.

Zeng, Xiaoyu 曾曉渝. (2003). Jianmu de shanggu yinzhi 見母的上古音值. *Zhongguo Yuwen*. Vol.2.

Zhao, Binxuan 趙秉璇. (1979). Taiyuan Fangyan Li de Hanyu Pianci. [Chinese Pian-ci in Taiyun Dialect], *Yuwen Yanjiu*. [Studies of Chinese], Vol. 1.

Zhao, Binxuan 趙秉璇 & Zhu, Jia-ning 竺家寧. (1998). *Guhanyu fushengmu lunwenji* 古漢語複聲母論文集 [Proceedings of Initial Clusters in Old Chinese]. Beijing: Beijing Language and Culture University Press.

Zheng, Wei 鄭偉. (2010). Ziran yinbian he yinbian chongjian: guhanyu yimuzi de yanbian fangshi 自然音變和音變重建：古漢語以母字的演變方式 [Natural process and reconstructing sound change: A case of evolutional types of yi initial in Old Chinese]. *IACL-18 & NACCL-22*, Harvard University.

Zheng, Yide 鄭懿德. (1983). Fuzhou Fangyan Danyin Dongci Chongdieshi. [Reduplications of Mono-syllabic Verbs in Fuzhou Dialect]. *Zhongguo Yuwen* 1: 31- 39.

Zhengzhang, Shangfang 鄭張尚芳. (1983). Shunguyin Gouni Xiaoyi. [On the Reconstruction of Old Chinese]. *Yuyanxue Luncong*. [Journal of Linguistics]. Vol. 14. Beijing: Shangwu Publishing House.

Zhengzhang, Shangfang 鄭張尚芳. (1987). Shanggu Yunmu Xitong he Sideng Jieyin Shengtiao de Fayuan Wenti. [The Problems of Rhyme System, the Fourth Division, Medials, and Tonogenesis in Old Chinese]. *Journal of Wenzhou Normal University*. 4.

Zhengzhang, Shangfang 鄭張尚芳. (2003). *Shanggu yinxi* 上古音系 [Old Chinese phonology]. Shanghai: Shanghai Jiaoyu Chubanshe.

Zhou, Jixu 周及徐. (2003). Shanggu Hanyu Zhong de Kw-/>P- Yinbian jiqi Shijian Cengci. [The Phonological Change of Kw-/>P- and its Chronology in Old Chinese].*Yuyan Yanjiu*. 2003, Vol. 4. 183-194.

Zhu, Guangqi 朱廣祁. (1985). *Shijing Shuangyinci Lungao* 雙音詞論稿 [Discussion on disyllables in *Shijing*]. Henan Province: Henan Renmin Publishing House.

Zhu, Xing 朱星. (1988). *Zhongguo Wenxue Yuyan Fazhan Shilue*. [The History of Chinese Literature and Language]. Beijing: Xinhua Publishing House.

Zhu, Jia-ning 竺家寧. (1990). *Chinese Phonology*《聲韻學》.Taipei: Wunan tushu 五南圖書.