

Word and Wordhood, Modern

San Duanmu

February 20, 2016

In this article, I review issues in defining words, such as distinctions between roots and affixes, between morphemes and words, between single words and compounds, and between compounds and phrases. I also discuss possible solutions to some problems and compare Chinese with English.

Keywords: affix, morpheme, word, compound, elastic word length, morpheme inventory, foot

1. Introduction

In English, it seems obvious what a word is: it is a meaningful unit written between spaces. In Chinese, there is a similar unit, also written between (invisible) spaces: it is called *zì* 字, a monosyllabic graph that in most cases has a meaning. Naturally, many people equate “word” in English with *zì* in Chinese. For example, Mǎ (1898), a pioneer of modern linguistics in China and the first native scholar to write a grammar of Chinese, calls a verb as *dòng zì* 动字 ‘action *zì*’, a noun *míng zì* 名字 ‘name *zì*’, a conjunction *lián zì* 连字 ‘connection *zì*’, and so on.

However, there are problems if we equate *word* with *zì*. For example, Karlgren (1918, 1923), Kennedy (1951), and Lín (1952) point out that *zì* is rarely used as a free word in modern Chinese; instead, most Chinese words are disyllabic. In addition, many scholars, such as Chao (1968), Lǚ (1981), Wáng Hóngjūn (1999), Chéng (2003), Pān *et al.* (2004), and Xú (2005), have argued, quite compellingly, that there is nothing in Chinese that corresponds to the notion of word. Indeed, the term is not found in the vocabulary of traditional Chinese linguistics, or in the vocabulary of ordinary Chinese speakers.

A crucial issue in defining words in Chinese is whether Chinese has affixes, which I shall discuss first. Then I discuss two views on words: (a) words can be defined for both English and Chinese, and (b) languages can differ in the basic unit of grammar, where it is *word* in English but *zì* in Chinese. Next, I discuss a well-known property of Chinese, which is elastic word length. I shall show that this property offers an explanation why in Chinese it is hard to distinguish morphemes from words on the one hand and words and compounds or phrases on the other (see also the articles on modern morphology, Old Chinese morphology, and prosodic morphology).

2. Affixes in Chinese

A word in English is often made of a root and one or more affixes. In the English lexicon *CELEX*, 80% of uninflected non-compound words (called ‘lemmas’) are made this way.

A typical affix has a grammatical function but no referential meaning, and it cannot be used alone. For example, the English suffix *-ly* (as in *slowly*) has the function of changing an adjective to an adverb and it cannot be used alone. Some English affixes have some referential meaning, such as *-er* in *writer* (which refers to a person), but these are in the minority and not used as roots.

Linguists who look for words in Chinese have proposed various affixes, too. Pān *et al.*

(2004:464-487) reviews 14 such studies, which have proposed a total of 355 affixes in Chinese. However, most of them, such as *cài* 菜 ‘vegetable’, *chǎng* 厂 ‘factory’, and *chē* 车 ‘vehicle’ have a referential meaning and no grammatical function; in addition, they can serve as roots or words. Naturally, linguists hardly agree on whether they all count as affixes. For example, of the 355 proposed affixes, only 11 are agreed upon by nine or more of the 14 studies. Let us take a look at them, shown in (1).

(1) Chinese affixes that have been proposed by nine or more of 14 studies

Affix	Studies	Example
<i>mén</i> 们 ‘plural’	9	<i>wǒmén</i> 我们 ‘i-plural (we)’
<i>huà</i> 化 ‘change’	11	<i>gōngyèhuà</i> 工业化 ‘industry-change (industrialize)’
<i>xìng</i> 性 ‘nature’	12	<i>gòngxìng</i> 共性 ‘common-nature’
<i>jiā</i> 家 ‘expert’	13	<i>zuòjiā</i> 作家 ‘write-expert (writer)’
<i>zhě</i> 者 ‘person’	13	<i>zuòzhě</i> 作者 ‘write-person (author)’
<i>yuán</i> 员 ‘member’	11	<i>jiàoyuán</i> 教员 ‘teach-member (faculty)’
<i>dù</i> 度 ‘degree’	10	<i>yìngdù</i> 硬度 ‘hard-degree (hardness)’
<i>tóu</i> 头 ‘head’	11	<i>shítóu</i> 石头 ‘stone’
<i>zǐ</i> 子 ‘son’	12	<i>zhuōzǐ</i> 桌子 ‘table’
<i>ér</i> 儿 ‘son’	11	<i>hāuér</i> 花儿 ‘flower’
<i>lǎo</i> 老 ‘old’	9	<i>lǎoshī</i> 老师 ‘teacher’

The list is rather short. In addition, from an English perspective, only the first looks like an affix, and possibly the second. Besides, *mén* 们 is mostly limited to the first, second, and third person pronouns, and not required otherwise (e.g., ‘those teachers’ can be *zhèxiē lǎoshī* 这些老师 or *zhèxiē lǎoshīmen* 这些老师们), and *huà* 化 can be used as a word. The next five items in (1) all have referential meanings, so that the disyllabic example looks like a compound (a word made of two words). The last four items in (1) add no meaning or function to the root they attach to; indeed, they are often redundant, because the root can be used without them, e.g., *fāng zhuō* 方桌 ‘square table’ (without 子 for ‘table’) and *huā píng* 花瓶 ‘flower vase’ (without 儿 for ‘flower’).

Unlike the paucity of affixes in Chinese, English has hundreds of them, based on the data in *CELEX*. For example, there are three suffixes spelled as *-acy*, as in *supremacy* (where the suffix converts A to N), *papacy* (where the suffix converts N to N), and *conspiracy* (where the suffix converts V to N). Therefore, while Chinese may have some affixes, the number is strikingly small compared with that in English.

3. Defining words in Chinese

Many linguists have offered analyses of words in Chinese in terms of roots and affixes (see Pān *et al.* 2004 for a comprehensive review). However, as just noted, a problem for this approach is the lack of true affixes in Chinese.

Some linguists, such as Jespersen (1922:369), are aware of the problem and have concluded that Chinese words are essentially monosyllabic. However, as Karlgren (1918; 1923), Kennedy

(1951), and Lín (1952) have argued, the question for this view is why most monosyllables are not free and why most Chinese words occur in disyllabic forms.

In yet another approach, both Chinese and English have words, but they differ in morphology (Sproat and Shih 1996; Packard 2000). In English, most words are made of a root plus one or more affixes, whereas in Chinese most words are made of two (or more) roots (called ‘root compounds’). This approach also faces a question, namely, why most roots are not free in Chinese and why root compounds are rare in English.

4. Giving up words in Chinese

Some linguists have come to the conclusion that “word” is not a universal entity for every languages. Instead, languages can differ in the basic units of grammar. For example, according to Chao and Yang (1947) and Chéng (2003), both English and Chinese have morphemes, but only English has words. Unlike English, which uses morphemes to build words, Chinese uses morphemes to build phrases. Some linguists go even further. For example, Pān *et al.* (2004) and Xú (2005) argue that English and Chinese do not need to share any common category at all: English has morphemes and words, and Chinese has *zì* 字, which is neither a morpheme nor a word.

5. Elastic word length

Many Chinese words (or morphemes) can be long (disyllabic) or short (monosyllabic), where the former contains the latter. The property has been discussed by Guō (1938) and many other scholars (see Dong 2015; Duanmu and Dong 2015a, 2015b). Let us call the property elastic word length, which is defined in (2) and exemplified in (3).

(2) Defining elastic word length in Chinese

A word has elastic length if it has two forms A and B where

- a. A is monosyllabic and B disyllabic (or longer);
- b. A and B share the same base (morpheme) (i.e., B contains A);
- c. A and B have the same meaning (i.e., B is semantically redundant or empty);
- d. A and B are interchangeable in some contexts

(3) Sample words with elastic length

Long	Short	
<i>xué-xí</i> 学习	<i>xué</i> 学	‘study-(practice)’
<i>jì-shù</i> 技术	<i>jì</i> 技	‘skill-(technique)’
<i>méi-tàn</i> 煤炭	<i>méi</i> 煤	‘coal-(charcoal)’
<i>lǎo-hǔ</i> 老虎	<i>hǔ</i> 虎	‘(old)-tiger’
<i>yā-zǐ</i> 鸭子	<i>yā</i> 鸭	‘duck-(son)’

The meaning of the long form can be limited to one of three kinds, which can be called XX, X0, and 0X. In XX, the meanings of the two parts are repetitive, such as *xué-xí* 学习 ‘study-practice’ and *xūyào* 需要 ‘need-want’. In X0, the meaning of the second part is empty, such as *yā-zǐ* 鸭子 ‘duck’. In 0X, the meaning of the first part is empty, such as *lǎo-hǔ* 老虎 ‘tiger’.

There are two views on such long-short pairs. The first view sees the two forms of a pair as essentially synonymous and belong to the same word (e.g., Karlgren 1918; 1923; Guō 1938; Chao and Yang 1947; Lín 1952; Pān 1997; Dong 2015; Duanmu and Dong 2015a, 2015b). The second view holds that the two forms differ in meaning and are different words (e.g., Lǐ 1990; Liu 1992; Wáng Cànlong 2002; Wú 2003; Kē 2007).

It is true, as Lǐ (1990) and Wáng (2002) argue, that in some pairs, the long form is more formal (e.g., *gòumǎi* 购买 vs. *mǎi* 买 ‘buy’), or more abstract (e.g., *sǐwáng* 死亡 vs. *sǐ* 死 ‘death’), or of a larger quantity (e.g., *shūjí* 书籍 vs. *shū* 书 ‘books’). However, such differences do not always hold for other pairs. For example, there is no evidence that *lǎohǔ* 老虎 ‘tiger’ is more formal or abstract, or implies a larger quantity, than *hǔ* 虎. Rather, there is broad consensus that, in most pairs, the two forms have the same referential meaning. This is evidenced by the fact that in dictionaries, the two forms of a pair are either listed under the same entry (e.g., Chao and Yang 1947) or used to annotate each other (e.g., *Xiàndài Hànyǔ Cídiǎn* 現代漢語詞典 [Dictionary of Modern Chinese]).

Karlgren (1918, 1923) suggests that the creation of long forms is motivated by the desire to avoid homophones, a view shared by many people. This view faces three problems. First, while some long forms are indeed expanded from short forms (such as *lǎohǔ* 老虎 ‘tiger’ from *hǔ* 虎), the reverse process is also present, i.e., reducing long forms to short ones, such as *rì* 日 ‘Japan’ from *riběnn* 日本 or *jī* 机 ‘airplane’ from *fēijī* 飞机 ‘fly machine (airplane)’. Second, there are restrictions on where long and short forms can occur. For example, in a corpus study, Duanmu (2012) has shown that in [N N], 1+2 (monosyllable + disyllable) is disfavored, whereas in [V N], 2+1 is disfavored. Such patterns suggest that long forms are preferred in phonologically strong positions. For example, in [N N], the first N is strong, as predicted by the Compound Stress Rule, and in [V N], N is strong, as predicted by the Nuclear Stress Rule (Chomsky and Halle 1968). An analysis in terms of stress assignment and foot binarity can be found in Lu and Duanmu (2002) and (Duanmu 2007). Third, according to the homophone avoidance theory, the more homophones a syllable has, the more likely a disyllabic form will be created for it. However, it has been shown that such a correlation does not exist; instead, syllables with no homophones are just as likely to have elastic length as syllables with many homophones (Dong 2015; Duanmu and Dong 2015a, 2015b).

How many monosyllables in Chinese have a disyllabic form? Pān (1997) suggests that nearly all of them do, but he offers no actual count. Huáng and Duānmù (2013) sampled a random list of 2,000 morphemes (one tenth of those in *Xiàndài Hànyǔ Cídiǎn*) and found that 61% of them have elastic length. Dong (2015) annotated the entire *Xiàndài Hànyǔ Cídiǎn* and found that 47% of Chinese morphemes have elastic length. The data are summarized in (4), divided by parts of speech (POS), where 1-only means a morpheme only has a monosyllabic form, Poly-only means a morpheme only has a polysyllabic form, and Elastic means a morpheme has both a monosyllabic form and a disyllabic (or longer) form.

(4) POS and length properties of Chinese morphemes (Dong 2015)

POS	All	POS %	1-only	Poly-only	Elastic
Noun	9,559	48.1%	32.5%	8.9%	58.6%
Verb	5,904	29.7%	56.1%	1.8%	42.1%
Adjective	2,709	13.6%	53.3%	9.4%	37.2%
Adverb	429	2.2%	72.5%	0.2%	27.3%
Others	1,286	6.5%	75.0%	18.0%	7.0%
All	19,887	100%	46.0%	7.2%	46.8%

Two comments of interest can be made. First, it can be seen that there is a fairly high percentage of elastic forms among content words (nouns, verbs, adjectives, and adverbs). In contrast, the percentage of elastic forms among function words (i.e. ‘others’) is quite low. Second, the average percentage of elastic words, at 46.8%, is lower than previous estimates and contradicts with a popular view that most Chinese expressions, especially nouns, occur in disyllabic forms (e.g., Chao 1948; Pān 1997). However, Duānmù (2015) has shown that most 1-only nouns cannot occur alone. For example, monosyllabic names commonly occur with another syllable, such as *Lǎo Wáng* 老王 ‘Old Wang’ and *Xiǎo Wáng* 小王 ‘Little Wang’, whereas disyllabic names usually occur without the extra syllable, such as *Ōuyáng* 欧阳 ‘Ouyang’. Other 1-only words are often part of a set phrase, such as *gū* 辜 ‘guilt’ in *wú gū* 无辜 ‘without guilt (innocent)’ and *jiā* 佳 ‘fine’ in *jiā rén* 佳人 ‘fine person (beautiful woman)’. Therefore, as far as Chinese expressions are concerned, most of them are disyllabic (or have a disyllabic form) in actual use.

6. zì 字, morphemes, and words

It has been suggested that zì 字 is basically a morpheme (Chao and Yang 1947; Chao 1968; Chéng 2003). Several provisions are needed though. First, sometimes a zì has two or more meanings and represents two (or more) morphemes, such as què 却, which can mean ‘retreat’ or ‘but’. Second, there are some disyllabic names, such as *wúgōng* 蜈蚣 ‘centipede’, where neither half has a meaning by itself. Third, a zì can occasionally represent a disyllabic compound, such as *qiānwǎ* 阡 ‘thousand watt (kilowatt)’, but such cases are new creations and normally have a disyllabic alternative (in this case 千瓦). For the first case, we can consider them to be homographs, i.e. different zì that happen to look (and sound) the same. The second and third cases are fairly rare. Therefore, in most cases, each zì represents a morpheme.

Pān *et al.* (2004) and Xú (2005) argue that zì is more than a morpheme, mainly because zì also includes the graph of a morpheme, which plays an important role in Chinese. For example, because Chinese has more homophones than English, the disambiguating role of zì is greater than that of the English orthography. For example, ‘morning’ 旦, ‘but’ 但, ‘egg’ 蛋, and ‘thin/weak’ 淡 are all pronounced *dàn* but are distinguished by different zì. In addition, a Chinese graph often offers clues to its meaning (e.g., many verbs that involve the use of hands have a graphic component that indicates a hand). Indeed, when a literate Chinese speaker explains a zì to someone, he or she would often instinctively use the index finger to draw the character in the air. However, while orthography may interact with grammar in some ways, it needs not be part of a language (for example, for children or people who have not learned writing). If we set aside the

orthographical aspect of *zì*, then each *zì* corresponds to a morpheme quite well.

It is interesting to compare the morpheme inventory sizes of English and Chinese. Consider the data in (5), calculated by myself, where English is based on *CELEX* and Chinese on *Xiàndài Hànyǔ Cídiǎn*.

(5) Sizes of morpheme inventories in English and Chinese

	English	Chinese
Data source	<i>CELEX</i>	<i>Xiàndài Hànyǔ Cídiǎn</i>
Zero derivations excluded	10,000	10,000
Zero derivations included	17,000	17,000

If we exclude zero derivations, e.g., counting *study* (noun) and *study* (verb) as one morpheme, then both English and Chinese have about 10,000 morphemes, of which two thirds are commonly used. If we include zero derivations, e.g., counting *study* (noun) and *study* (verb) as two morphemes, then both languages have about 17,000 morphemes. Thus, the morpheme inventories are strikingly similar in size.

Recall that there are two problems in defining words in Chinese. First, most monosyllables are not free and hence not words. Second, most words are disyllabic, which look like compounds. Given the property of elastic word length, we can account for both problems. First, a monosyllable is not free because phonology requires a minimal word to be a disyllabic foot, in which case we should choose the disyllabic form of a word, which is always free; thus, most words do have a free form. Second, the disyllabic forms of a word may look like a compound, but it is not, since its meaning is XX, X0, and 0X, whereas a true compound should be XY (made of two different meanings); thus, based on semantic structure, we can distinguish disyllabic words from disyllabic compounds.

7. Words vs. phrases

There are three difficulties in distinguishing words from phrases in Chinese. First, as just discussed, the long form of a word looks like a compound. Second, some nominal units look like a phrase, such as *xiǎo chē* 小车 ‘small car’. Third, there are disyllabic units, called *líhécí* 离合词 ‘separable words’, which behave like a compound in some cases but can be split in a phrase elsewhere. For example, *yǒu qián* 有钱 ‘have money (rich)’ behaves like a compound (e.g., it can be modified by ‘very’, as in *hěn yǒu qián* 很有钱 ‘very rich’), but it can be split in *yǒu hěnduō qián* 有很多钱 ‘have a lot of money’.

We can account for the first case by drawing a distinction in semantic structure, where words are XX, 0X, and X0 and phrases are XY. For the second case, a distinction can be made, as discussed in Duanmu (1998) and references therein. Specifically, [M N] is a word or compound and [M 的 N] is a phrase, where M is a modifier of the noun N. For the third case, we can consider 有钱 to be ambiguous between a compound ‘rich’ and a phrase ‘have money’, just as ‘black sheep’ can be a compound (an unusual person) or a phrase (a sheep that is black).

8. Summary

I have reviewed problems in defining words in Chinese, including difficulties in distinguishing roots from affixes, morphemes from words, single words from compounds, and compounds from phrases. I have suggested that many problems may be solvable, especially in view of the elastic property of word length. Therefore, Chinese may not be as different from English as it seems after all.

Primary Sources:

CELEX: Baayen, R. Harald, Richard Piepenbrock, and L. Gulikers, *The CELEX lexical database: release 2* (CD-ROM), Philadelphia: Linguistic Data Consortium, University of Pennsylvania, 1995.

Xiàndài Hànyǔ Chángyòng Cíbiǎo: Xiàndài Hànyǔ Chángyòng Cíbiǎo Kètízǔ 现代汉语常用词表课题组, *Xiàndài Hànyǔ chángyòng cíbiǎo (cǎo'àn)* 现代汉语常用词表 (草案) [A commonly used word list in Modern Chinese], Běijīng 北京: Shāngwù Yīnshūguǎn 商务印书馆, 2008.

Xiàndài Hànyǔ Cídiǎn: Zhōngguó Shèhuì Kēxuéyuàn Yǔyán Yánjiūsuǒ Cídiǎn Biānjíshì 中国社会科学院语言研究所词典编辑室, *Xiàndài Hànyǔ cídiǎn (dì wǔ bǎn)* 现代汉语词典(第五版) [Modern Chinese Dictionary (Fifth Edition)], Běijīng 北京: Shāngwù Yīnshūguǎn 商务印书馆, 2005.

References:

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

Chao, Yuen-Ren, *A grammar of spoken Chinese*, Berkeley, California: University of California Press, 1968.

Chao, Yuen Ren, and Lien Sheng Yang, *Concise Dictionary of Spoken Chinese*, Cambridge, Mass.: Harvard University Press, 1947.

Chéng Yǔmín 程雨民, *Hànyǔ zì jī yǔfǎ : yǔsù céng zào jù de lǐlùn hé shíjiàn* 汉语字基语法: 语素层造句的理论和实践 [A zì-based grammar of Chinese: theory and practice in building sentences with morphemes], Shànghǎi 上海: Fùdàn Dàxué Chūbǎnshè 复旦大学出版社, 2003.

Chomsky, Noam, and Morris Halle, *The sound pattern of English*, New York: Harper and Row, 1968.

Dong, Yan, *The prosody and morphology of elastic words in Chinese: annotations and analyses*, Doctoral dissertation, University of Michigan, Ann Arbor, 2015.

Duanmu, San, Wordhood in Chinese, Jerome L. Packard, ed., *New approaches to Chinese word formation: morphology, phonology and the lexicon in modern and ancient Chinese*, Berlin: Mouton de Gruyter, 1998, 135-196.

Duanmu, San, *The phonology of Standard Chinese*, 2nd Edition, Oxford: Oxford University Press, 2007.

Duanmu, San, Word-length preferences in Chinese: a corpus study, *Journal of East Asian Linguistics* 21.1, 2012, 89-114.

Duānmù Sān 端木三, 2015. Hànyǔ dānyǔsù míngcí de cízhǎng tánxìng yánjiū: cídiǎn zhōng hé shíjì shíyòng zhōng “zhǐ dān bù shuāng” míngcí 汉语单语素名词的词长弹性研究: 词典中和实际使用中的“只单不双”名词 [A study of elastic length in Chinese monomorphemic

- nouns: monosyllabic-only nouns in the dictionary and in actual use], Féng Shènglì 冯胜利 ed., *Hànyǔ yùnlǜ yǔfǎ xīntàn* 汉语韵律语法新探 [New explorations on Chinese prosodic grammar], [yè. Shànghǎi 上海: Zhōng Xī Shūjú 中西书局, 2015, 4-21.
- Duanmu, San, and Yan Dong, Homophone density and word length in Chinese, Yuchau E. Hsiao and Lian-Hee Wee, eds., *Capturing phonological shades within and across languages: inspirations from the Theoretical Phonology Conferences*, Newcastle upon Tyne: Cambridge Scholars Publishing, 2015a, 213-242.
- Duanmu, San, and Yan Dong, Elastic Words in Chinese, Sin-Wai Chan, ed., *Routledge Encyclopedia of the Chinese Language*, New York: Routledge, 2015b (in press).
- Guō Shàoyú 郭紹虞, Zhōngguó yǔcí zhī tánxìng zuòyòng 中國語詞之彈性作用 [The function of elastic word length in Chinese], *Yānjīng Xuébào* 燕京學報 24, 1938. Reprinted in Guō Shàoyú 郭紹虞, *Yǔwén tōnglùn xùbiān* 語文通論續編 [Papers on Chinese language and literature, new edition], Xiānggǎng 香港: Tàipíng Shūjú 太平書局, 1963, 1-40.
- Huáng Lìjūn 黃麗君 and Duānmù Sān 端木三, Xiàndài Hànyǔ cícháng tánxìng de liànghuà yánjiū 現代漢語詞長彈性的量化研究 [A quantitative study of elastic word length in Modern Chinese], *Yǔyánkēxué* 語言科學 12.1, 2013, 8-16.
- Jespersen, Otto, *Language: its nature, development and origin*, New York: Macmillan, 1922.
- Karlgren, Bernhard, *Ordet och pennan i Mittens Rike* [The word and the pen in the Middle Kingdom], Stockholm: Svenska Andelsförlaget, 1918.
- Karlgren, Bernhard, *Sound & symbol in Chinese*, London: Oxford University Press, 1923.
- Kē Háng 柯航, Xiàndài Hànyǔ dān shuāng yīnjié dāpèi yánjiū 現代漢語單雙音節搭配研究 [A study of monosyllabic and disyllabic usage in modern Chinese], Zhōngguó Shèhuì Kēxuéyuàn Yǔyán Yánjiūsuǒ bóshì lùnwén 中國社會科學院語言研究所博士論文 [Doctoral dissertation, Institute of Linguistics, Chinese Academy of Social Sciences], 2007.
- Kennedy, George A., The monosyllabic myth, *Journal of the American Oriental Society* 71.3, 1951, 161-166.
- Lǐ Líndìng 李臨定, Dòngcí fēnlèi yánjiū shuōlùè 動詞分類研究說略 [A brief discussion on verb classification], *Zhōngguó Yǔwén* 中國語文 4, 1990, 248-257.
- Lín Hàndá 林漢達, Hànyǔ shíbùshì dānyīnjié yǔ? 漢語是不是單音節語? [Is Chinese a monosyllabic language?], *Zhōngguó Yǔwén* 中國語文 11, 1952, 6-11.
- Liu, Feng-his, Verb and syllable in Chinese, paper presented at the 25th International Conference on Sino-Tibetan Languages and Linguistics, Berkeley, 1992.
- Lu, Bingfu, and San Duanmu, Rhythm and syntax in Chinese: A case study, *Journal of the Chinese Language Teachers Association* 37.2, 2002, 123-136.
- Lǚ Shūxiāng 呂叔湘, *Yǔwén cháng tán* 語文常談 [Casual remarks on language], Běijīng 北京: Sān Lián Shūdiàn 三聯書店, 1981.
- Mǎ Jiànzhōng 馬建忠, *Mǎshì wéntōng* 馬氏文通 [Mr. Mǎ's grammar], Shànghǎi 上海: Shāngwù Yìnshūguǎn 商務印書館, 1898.
- Packard, Jerome L., *The morphology of Chinese: A linguistic and cognitive approach*, Cambridge, UK: Cambridge University Press, 2000.
- Pān Wénguó 潘文國, *Hàn Yīng yǔ duìbǐ gāngyào* 漢英語對比綱要 [A comparative outline of Chinese and English], Běijīng 北京: Běijīng Yǔyán Wénhuà Dàxué Chūbǎnshè 北京語言文化大學出版社, 1997.

- Pān Wénguó 潘文国, Yè Bùqīng 叶步青, and Hán Yáng 韩洋, *Hànyǔ de gòucífǎ yánjiū* 汉语的构词法研究 [A study of Chinese morphology], Shànghǎi 上海: Huádōng Shīfàn Dàxué Chūbǎnshè 华东师范大学出版社, 2004.
- Sproat, Richard, and Chilin Shih, A corpus-based analysis of Mandarin nominal root compound. *Journal of East Asian Linguistics* 5.1, 1996, 49-71.
- Wáng Cànlong 王灿龙, Jùfǎ zǔhé zhōng dān shuāng yīnjié xuǎnzé de rènzhī jiěshì 句法组合中单双音节选择的认知解释 [A cognitive explanation of the choice between monosyllables and disyllables in syntactic units], Zhōngguó Yǔwén Zázhi 中国语文杂志社, ed., *Yǔfǎ Yánjiū hé Tànsuǒ*, dì 11 Jí 语法研究和探索, 第 11 辑, Běijīng 北京: Shāngwù Yìnshūguǎn 商务印书馆, 2002, 151-168.
- Wáng Hóngjūn 王洪君, *Hànyǔ fēixiànxìng yīnxìxué* 汉语非线性音系学 [Chinese non-linear phonology], Běijīng 北京: Běijīng Dàxué Chūbǎnshè 北京大学出版社, 1999.
- Wú Wéishàn 吴为善, Shuāngyīnhuà, yǔfǎhuà hé yùnlǜcí de zài fēnxī 双音化、语法化和韵律词的再分析 [A reanalysis of disyllabification, grammaticalization, and prosodic words], *Hànyǔ Xuéxí* 汉语学习 2, 2003, 8-14.
- Xú Tōngqiāng 徐通锵, *Hànyǔ jiégòu de jīběn yuánlǐ — zì běnwèi hé yǔyán yánjiū* 汉语结构的基本原理 — 字本位和语言研究 [Fundamentals of the structure of Chinese: the zì-base and language study], Qīngdǎo 青岛: Zhōngguó Hǎiyáng Dàxué Chūbǎnshè 中国海洋大学出版社, 2005.