Elastic Words in Chinese
San Duanmu and Yan Dong
University of Michigan

Summary
Elastic words are those whose length can vary from monosyllabic to disyllabic, without a change in meaning. They have been known since Karlgren (1918), but both theoretical and empirical questions remain. We offer a precise definition of elastic words, which distinguishes them from regular compounds. Then we calculate the percentage of elastic words in modern Chinese. Finally, we compare theories of why elastic words are created in Chinese and show that Foot Binarity, rather than homophony, offers the best explanation.

1. What are elastic words?

Chinese has many long-short pairs of synonymous word forms. Some examples are shown in (1). The long form is made of the short one plus another morpheme, which can itself be a word. However, the extra part of the long form contributes little new meaning; therefore, it is shown in parentheses.

(1)　　　Examples of elastic words in Chinese

Short		Long	
见	'see'	(看)见	'(look)-see'
学	'study'	学(习)	'study-(practice)'
技	'skill'	技(术)	'skill-(method)'
煤	'coal'	煤(炭)	'coal-(charcoal)'
泪	'tear'	(眼)泪	'(eye)-tear'
蒜	'garlic'	(大)蒜	'(big)-garlic'
虎	'tiger'	(老)虎	'(old)-tiger'
教	'teach'	教(书)	'teach-(book)'
写	'write'	写(字)	'write-(word)'

Since the two forms share the same basic part and have the same meaning, many people consider them to be variants of the same word (Karlgren 1918; 1923; Guo 1938; Chao 1948). We follow Guo (1938) and refer to them as elastic words, or words with elastic length. Theoretically, there is nothing unusual for a word to have two (or more) phonetic forms (Bloomfield 1933). For example, *'m* [m] (as in *I'm*) and *am* [æm] are different forms of the same English word, so are *math* and *mathematics*. What is unusual in Chinese is that the long form is often in the shape of a compound, a point we shall return.

Elastic words in (1) represent the most common cases. Less common cases are also found, shown in (2)-(4). In (2), the long form is trisyllabic. In (3), the two parts of the long form have similar meanings and either can be used alone. In (4), a short form has more than one long form. In this article, however, we shall focus on the most common cases, such as those in (1).

(2)　　　Monosyllabic-trisyllabic pair

Length		Example	
Long	加拿大 'Canada'	加拿大元 'Canadian dollar'
Short	加 'Canada'	加元 'Canadian dollar'

(3)　　　Either part of the long form can be the short form

Length		Example	
Long	需要 'need-want'	不需要 'not need-want'
Short	需 'need'	不需 'not need'
Short	要 'want'	不要 'not want'

(4)　　　More than one long form

Length	Example	
Short	泪	'tear'
Long	泪(水)	'tear-(water)'
Long	泪(花)	'tear-(flower)'
Long	(眼)泪	'(eye)-tear'
Long	泪(珠)	'tear-(pearl)'
Long	泪(液)	'tear-(liquid)'
Long	(热)泪	'(warm)-tear'

The use of elastic words in Chinese is extremely common, so common that most Chinese speakers, scholars or average people alike, hardly feel anything unusual about it. However, when Chinese expressions are translated into English literally, the oddity becomes apparent. Consider the example in (5)-(7).

(5)　　　An ordinary Chinese sentence and its English translation
我　刚(刚)　　(看)见　　他　(回)到　　　　家(里)。
'I　just-(just)　(look)-see　him　(return)-reach　home-(inside).'

(6)　　　Object required in a Chinese answer

Question	你 想 做 什么？ you want do what 'What do you want to do?'
Odd answer	教/写/走/吃. 'Teach/Write/Walk/Eat.'
Normal answer	教书/写字/走路/吃饭. 'Teach book/Write word/Walk road/Eat meal.'

(7)   Object not required in an English answer

Question	'What do you want to do?'
Normal answer	'Teach/Write/Walk/Eat.'
Odd answer	'Teach books/Write words/Walk roads/Eat meal.'

In (5), four of the six Chinese words are repetitive (although the sentence can be used without the repetitive parts, too). In (6) an object is required in a Chinese answer, even though it is obvious from the verb. For example, in the normal case, what else do you write if not words (or books)? In contrast, as seen in (7), corresponding English answers are fine without an object, and preferably so.

To the Western eye, therefore, Chinese seems highly repetitive or redundant. Naturally, when Bernhard Karlgren started learning Chinese, he immediately noted this oddity and reported it (Karlgren 1918). Since then, many others have discussed the topic (e.g. Jespersen 1930; Guo 1938; Chao 1948; Kennedy 1951; 1955; Lü 1963; Li and Thompson 1981; Pan 1997; Ke 2006; Jin 2011), yet many important questions remain unanswered. In this article we address such questions. In particular, we examine the issues in (8).

(8)   Questions to be discussed with regard to elastic words
       a.   How are elastic words defined?
       b.   How many words are elastic in modern Chinese?
       c.   Why does Chinese have so many elastic words?
       d.   Impact of elastic words on poetry

In section 2 we consider (8a). In section 3 we consider (8b). In section 4 we consider (8c). In section 5 we consider (8d). In section 6 we offer concluding remarks.

2.   How are elastic words defined?
In previous studies, it is often left to the intuition of the reader whether two forms, such as 店 'store' and 商店 'business-store', are different words or variants of the same word. Many scholars assume that pairs like 店 'store' and 商店 'business-store' are variants of the same word (e.g. Karlgren 1918; 1923; Jespersen 1930; Guo 1938; Chao 1948; Kennedy 1951; 1955; Lü 1963; Li and Thompson 1981; Pan 1997; Ke 2006; Jin 2011). For example, Chao (1948) points out that, while most Chinese morphemes are monosyllabic, most Chinese words are disyllabic. Chao's statement makes sense only if many disyllabic forms are simple words, and not true compounds. The reason is that a true compound is made of two (or more) words that contribute different meanings. For example, 书包 'book bag (backpack)' and 煤店 'coal store' are true compounds, because 'book bag' is not the same as 'book' or 'bag', and 'coal store' is not the same as 'coal' or 'store'. If all disyllabic forms are true compounds in Chinese, we cannot reasonably say that Chinese words are disyllabic, such 书 'book', 包 'bag', 煤 'coal', and 店 'store'. Chao's statement makes sense only if many disyllabic forms are also simple words, such as 书(本) 'book-(volume)', 煤(炭) 'coal-(charcoal)', and (商)店 '(business)-store'.

Perhaps because the long form of an elastic word looks like a compound, some scholars treat all disyllabic forms as true compounds. In addition, they propose that, even for related pairs, the long form always has a different meaning from the short one (e.g. Li 1990; Wang 2002; Wu 2003; Ke 2007). Some examples have been offered to support their view, such as that in (9).

3

(9)   Semantic difference between short and long forms (Li 1990)

Length			Basic meaning	Additional meaning
Short	死	'die'	'die'	abstract
Long	死亡	'die-perish'	'die'	concrete

Li (1990) observes that, when used as a verb, the short form 'die' has a concrete meaning, whereas the long form 'die-perish' has an abstract meaning. Therefore, 'die-perish' can be used for an abstract entity, such as 'society', but not for a concrete entity, such as a person or an animal. In contrast, 'die' has a concrete meaning and can be used for concrete entities, such as a person or an animal. However, examples like that in (9) are sporadic, and quantitative data are lacking. In addition, it can be shown that most elastic words do not have such meaning distinctions between the long and short forms. Moreover, as is often observed before, and as we shall see below with quantitatively data, while both languages use compounds, elastic words occur extensively in Chinese but not in English. To treat elastic words as regular compounds is to miss an important different between the two and an important difference between Chinese and English.

Let us now consider an explicit definition of elastic words, shown in (10) and (11), based on (Duanmu 2012; Huang and Duanmu 2013).

(10)   Defining elastic words in Chinese:
An elastic word has two length forms A and B where
a. A is monosyllabic and B disyllabic
b. A and B share the same base (morpheme)
c. A and B have the same meaning
d. A and B are interchangeable in some contexts

Strictly speaking, (10a) is not necessary, but it represents a property of the majority of elastic words in Chinese. Given (10a), (10b) means that B contains A. (10c) means that the extra part of B adds no new meaning to A and is therefore semantically redundant. (10d) is expected if (10c) is the case.

Of the four conditions in the definition, (10a) and (10b) are easy to verify, so is (10d). An example of (10d) is shown in (11).

(11)   Three interchangeable expressions for 'skilled worker'
a. 技术 工人 'skill-method   worker-person'
b. 技术 工   'skill-method   worker'
c. 技 工   'skill   worker'

(11a) and (11b) show that 工(人) 'worker-(person)' and 工 'worker' are interchangeable. (11b) and (11c) show that 技(术) 'skill-(method)' and 技 'skill' are interchangeable.

Let us now consider (10c). It can often be a subtle judgment whether two forms have exactly the same meaning. However, we can rely on an objective criterion, which is mutual annotation in dictionary entries. For example, if A is defined as B in a dictionary and B defined as A, then we consider A and B to have the same meaning. An example is shown in (12). The

definitions are found in *Modern Chinese Dictionary* (Xiandai Hanyu Cidian 2005), where … is a description of 'coal'.

(12)   Mutual dictionary annotation of 煤 'coal' and 煤炭 'coal-charcoal'

Entry	Annotation
煤	noun, …, also called 煤炭 'coal-charcoal'
煤炭	noun, 煤 'coal'.

Our definition of elastic words is both explicit and implementable. Its application is illustrated in (13).

(13)   Word pairs that do or do not form elastic words

Short	Long	Same meaning	Same base	Elastic
煤 'coal'	煤炭 'coal-(charcoal)'	Yes	Yes	Yes
虎 'tiger'	老虎 '(old)-tiger'	Yes	Yes	Yes
乐 'glad'	高兴 'high-mood'	Yes	No	No
包 'bag'	书包 'book-bag'	No	Yes	No

According to our definition, 煤 and 煤炭 are variants of the elastic word 'coal' and 虎 and 老虎 are variants of the elastic word 'tiger'. In contrast, 乐 'glad' and 高兴 'high-mood (glad)' are different words, so are 包 'bag' and 书包 'book-bag'.

3. How many words are elastic in modern Chinese?
Having defined what elastic words are, let us consider how many there are in Chinese. Previous scholars generally agree that there are many. Some even suggests that 'nearly all' Chinese words have elastic length (Pan 1997: 140). However, quantitative data have not been offered until quite recently. Three studies are summarized in (14).

(14)   Quantitative studies on the percentage of elastic words in Chinese

Study	Data size	Elastic
Duanmu (2013)	60 characters; 84 words	79%
Huang and Duanmu (2013)	2,000 morphemes in MCD	61%
Duanmu and Dong (2014)	All morphemes in MCD	47%

Duanmu (2013, written in 2011) sampled 1/50 of 3,000 most commonly used Chinese characters, totaling 60 characters and 84 words (some characters represent more than one word). It was found that elastic words constitute 79% of the samples.

Huang and Duanmu (2013) sampled 10% of all morphemic words in *Modern Chinese Dictionary* (Xiandai Hanyu Cidian 2005, MCD for short), totaling 2,000 word senses. Items that have related meanings but differ in part of speech are counted as different ones. For example, 学习 'study' can be a noun or a verb, and each is counted as a separate item. Of the 2,000 word senses examined, 61% are found to be elastic (after excluding names).

Duanmu and Dong (2014) examined all morphemes in MCD, totaling some 10,000 word entries and 20,000 word senses (parts of speech distinguished). It is found that 47% of them are

elastic words. The details are shown in Table 1. The label '1-only' refers to monosyllabic morphemes without a disyllabic form, such as 水 'water', 人 'person', and 狗 'dog'. The label 'Poly-only' refers to (mostly disyllabic) morphemes without a monosyllabic form, such as 琥珀 'amber', 葡萄 'grape', and 儒艮 'dugong'. The label 'Elastic' refers to morphemes that can be monosyllabic or disyllabic, such as 虎-老虎 'tiger'-'old tiger', 眉=眉毛 'brow'-'eyebrow', and 煤-煤炭 'coal'-'coal-charcoal'.

Table 1:　POS and length properties of Chinese morphemes.

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%

The result of Duanmu and Dong (2014) is likely to be more accurate, since its data coverage is exhaustive and many times larger than previous ones. The average percentage of elastic words is lower than those reported in previous. The result seems counter to the intuition of some scholars that most Chinese words are or can be disyllabic in use (e.g. Chao 1948; Pan 1997). It turns out that most morphemes that do not have a disyllabic form are either verbs, which often have an object to occur with, or a noun that must occur with another morpheme in a disyllabic set phrase or compound. Three examples are shown in (15).

(15)　　Example: monosyllabic word without elastic length

Morpheme	Use
教 'teach'	教书 'teach book (teach)'
	教英语 'teach English'
辜 'guilt'	无辜 'no guilt (innocent)'
	余辜 'extra guilt'
点 'point'	起点 'start point'
	沸点 'boil point'
	冰点 'ice point (freezing point)'

None of the morphemes in (15) has elastic length and none of them is used alone either. The word 教 'teach' has a default object 'book' if it is to be used alone, regardless of whether a book is used or not. The word 辜 'guilt' is used in a few set expressions only, such as 'no guilt (innocent)' and 'extra guilt'. The word 点 'point' is always used with another word, too, which it is a point of, except in geometry, where it has a special meaning and can be used as an independent noun. If we take such cases into consideration, then most Chinese words can indeed occur in disyllabic forms, either in the long form of an elastic word or in combination with another syllable in a set phrase.

There is no study on how many English words have elastic length. Various examples can be proposed though, such as those in (16).

(16)   Morphologically and semantically related word pairs in English
        Short   Long
        Ben   *Benjamin*
        prof   *professor*
        math   *mathematics*
        dog   *doggy*
        brow   *eyebrow*

The first three pairs are based on truncation, where the extra part of the long form is not an independent morpheme. In contrast, the long form of an elastic word in Chinese is typically made of two independent morphemes. The word *doggy* is arguably made of two morphemes, but it differs in stylistic meaning from *dog*; therefore, *dog* and *doggy* are hardly interchangeable. However, the pair *brow-eyebrow* does satisfy the definition of elastic words, similar to 泪 'tear' and 眼泪 'eye-tear' in Chinese. Nevertheless, it is reasonable to say that word pairs like *brow-eyebrow* are fairly rare in English and constitute far lower than 47% of the vocabulary.

4. Theories on elastic words

Having confirmed the abundance of elastic words in Chinese (and their rarity in English), let us consider why this is the case. Various theories have been proposed. Let us consider five of them, shown in (17).

(17)   Theories on why Chinese has so many elastic words
        a. Speech rate
        b. Processing need
        c. Disyllabic origin
        d. Homophony
        e. Prosody

The 'speech rate' theory is proposed by Guo (1938). The idea is that in some positions we need to speak faster and in some we need to speak slower. Monosyllabic forms are good for the former and disyllabic forms are good for the latter, when one would start to repeat words. It is further proposed that speech rate is a property of spoken Chinese and that written Chinese ought to use monosyllabic words most of the time, unless it tries to reflect the spoken style, as is the case and in modern writing. Moreover, it is proposed that elastic words have existed in Chinese all along, as can be observed on in early poetry

The 'processing need' theory is proposed by Pan (1997: 177). The idea is that the information conveyed by monosyllabic words is too dense for the listener, who cannot process what is hears fast enough. To help the listener, the speaker slows down by using repetitive words.

The 'disyllabic origin' theory is proposed by Kennedy (1951; 1955). The idea is that most Chinese words were originally in the first place, as some still are, such as 蝴蝶 'butterfly'. However, because Chinese dictionaries use characters s entries, instead of using words, each disyllabic word is split into two entries. For example, 蝴蝶 would have an entry under 蝴, where the annotation would be 'butterfly', and another entry under 蝶, where the annotation would

again be 'butterfly'. As a result, users of such dictionaries get the misimpression that each entry is a word and start using them as such, mistakenly. Gradually, most Chinese words have become monosyllabic.

The 'homophony' theory is proposed by Karlgren (1918). It is based on the fact that classic Chinese used to have 3,000-4,000 distinct syllables, whereas modern Standard Chinese only has 1,300. Since most Chinese morphemes are monosyllabic, there is a high degree of homophony. If disyllabic words are introduced, homophony can be radically reduced and ambiguity avoided. It is proposed, therefore, that Chinese created disyllabic words precisely to serve this purpose. In addition, since monosyllabic words are still alive, we end up with the situation of many elastic words.

The 'prosody' theory is what we would like to argue for. The idea is that some positions are stressed (by word stress or phrasal stress) and others not, and that stressed positions need a disyllabic word, owing to the phonological requirement of Foot Binarity (Prince 1980). To satisfy Foot Binarity, disyllabic forms are created.

Little evidence has been shown for the 'processing need' theory or the 'disyllabic origin' theory. We shall, therefore, not discuss them any further. Instead, we shall focus on the homophony theory and the prosodic theory and compare their predictions. The comparison hs implications for the speech rate theory, too, to be discussed later.

The homophony theory is intuitively quite plausible. Naturally, it is shared by many other scholars (e.g. Jespersen 1930; Guo 1938; Lü 1963; T'sou 1976; Li and Thompson 1981; Nettle 1995; Ke 2006; Jin 2011). Indeed, it has been called the 'orthodox' (Kennedy 1955). However, apart from sporadic examples, little evidence for it has been shown. Let us consider its predictions and compare them with those of the prosody theory, shown in (18).

(18)   Predictions of the homophony theory and the prosody theory

	Homophony	Rhythm
Minimal word	not relevant	relevant
Homophone	relevant	not relevant
Stress	not relevant	relevant
Spoken	yes	yes
Written	no	yes
Modern	yes	yes
Classic	no	yes

First, the homophony theory assumes that disyllabic words serve to avoid ambiguity. When there is no ambiguity, there is no need to use disyllabic words. Therefore, there ought to be no 'minimal word' effect, i.e. a word spoken alone need not be disyllabic but could be monosyllabic, if the meaning is clear. The homophony theory also predicts a positive correlation between homophone density (the number of homophones per syllable) and the percentage of elastic words. Syllables that have more homophones are more likely to have elastic length. On the other hand, stress ought to have no influence on the percentage of elastic words. In addition, since modern Chinese has much fewer syllables than classic Chinese, there ought to be a lot more elastic words in modern Chinese than in classic Chinese. Moreover, disyllabic words should occur in spoken Chinese only. In written Chinese, the number of distinct graphs remains similar to that in classic Chinese and far exceeds the number of distinct syllables. Therefore, there is little need for disyllabic words in written Chinese.

The prosody theory makes fairly different predictions. First, there ought to be a 'minimal word' effect, because a word spoken alone has stress and should be disyllabic, whether it is ambiguous or not. Second, there is no correlation between homophone density and the percentage of elastic words, since the creation of elastic words is driven by Foot Binarity, not by homophony. Third, stress is relevant for the use of disyllabic words. Fourth, there is no reason written Chinese should not reflect the effect of Foot Binarity; therefore, elastic words ought to be found in both spoken and written Chinese. Finally, prosodic requirement is likely to be the same in both classic Chinese and modern Chinese. Therefore, we expect there to be just as many elastic words in classic Chinese as in modern Chinese.

In summary, the two theories agree that elastic words are extensively used in spoken, modern Chinese. However, they differ in whether the minimal word effect, homophony, stress, written Chinese, and classic Chinese are related to elastic words. Therefore, we examine these factors in detail.

4.1. The minimal word effect

It is well known that Chinese has a 'minimal word' requirement, by which monosyllables are not spoken alone. Many examples have been offered by Lü (1963). In addition, this is the case even if there is no ambiguity. An example is shown in (19).

(19)   Minimal word requirement in Chinese when there is no ambiguity

Question:	你去华山还是泰山？ You go Hua Mountain or Tai Mountain Will you go to Mt. Hua or Mt. Tai?
Normal answer:	泰山。 Tai Mountain.
Odd answer:	泰。 Tai.

A monosyllabic mountain name is usually accompanies by the word 'mountain' in order to become disyllabic, presumably because it would otherwise be ambiguous, owing to many homophones. However, when there is no ambiguity, as in (19), the answer still has to be disyllabic. In general, all monosyllabic names must be used with another syllable when spoken alone, such as 姓王 'named Wang', 沙市 'Sha City', 法国 'France Country', 华山 'Hua Mountain', and 鲤鱼 'carp fish'. In contrast, disyllabic names can be spoken alone with an extra word, such as 欧阳 'Ouyang' (a personal name), 上海 'Shanghai', 荷兰 'Holland', and 峨嵋 'Emei' (a mountain name).

4.2. Elastic words and homophony

Next we consider the correlation between homophone density and the percentage of elastic words. The homophony theory predicts the correlation to exist. Some small-scale studies have been offered (T'sou 1976; Ke 2006; Jin 2011), but questions remain (see review in Duanmu and Dong 2014). In contrast, the prosody theory predicts that there is no such correlation. To verify the correlation, we used a large set of lexical data. In particular, we annotated all morphemes in *Modern Chinese Dictionary*, totaling some 20,000 items. For each monosyllabic morpheme, we annotated its homophone density of (the number of homophones it has) within its POS category. For each set of syllables that share the same homophone density, we annotated its percentage of

9

elastic words. The results for nouns (excluding family names), which constitute the largest POS category, are shown in Tables 2 and 3 (Duanmu and Dong 2014).

Table 2:   Homophone density and the percentage of elastic words among noun morphemes in Chinese. Nouns that are polysyllabic only are excluded. A homophone density of 1 means a word has no homophone (but itself).

Homophones	1-only	Elastic	All	Elastic %
1	75	92	167	55%
2	91	125	216	58%
3	101	199	300	66%
4	101	219	320	68%
5	139	226	365	62%
6	129	267	396	67%
7	147	252	399	63%
8	139	237	376	63%
9	141	264	405	65%
10	87	203	290	70%
11	115	248	363	68%
12	132	240	372	65%
13	133	270	403	67%
14	133	203	336	60%
15	105	225	330	68%
16	65	159	224	71%
17	98	174	272	64%
18	72	126	198	64%
19	117	225	342	66%
20	81	139	220	63%
21	54	93	147	63%
22	96	168	264	64%
23	49	89	138	64%
24	59	85	144	59%
25	37	63	100	63%
26	52	104	156	67%
27	77	166	243	68%
28	42	98	140	70%
29	8	21	29	72%
30	17	43	60	72%
31	22	40	62	65%
32	34	62	96	65%
33	11	22	33	67%
34	10	24	34	71%
35	20	50	70	71%

36	26	46	72	64%
37	24	50	74	68%
40	17	23	40	58%
42	15	27	42	64%
44	14	30	44	68%
47	35	59	94	63%
48	19	29	48	60%
49	16	33	49	67%
51	8	43	51	84%
56	23	33	56	59%
62	30	32	62	52%
64	19	45	64	70%
All	3,035	5,671	8,706	65%

Table 3:   Statistics of the data in Table 2, which show no correlation between homophone density and the percentage of elastic words.

Correlation:	0.089
Confidence interval (95%):	-0.066 0.122
Multiple R-squared:	0.008
F-statistic:	0.3628 on 1 and 45 degrees of freedom
P-value:	0.55

It can be seen that, in most homophone density groups, the percentage of elastic words lies between 60%-70%. Statistic results in Table 3 show that the confidence interval of the correlation crosses the zero point, which means there is no correlation between homophone density and the percentage of elastic words.

4.3. Elastic words and stress

To see the effect of stress, let us consider word length preferences in [N N] compounds and [V O] phrases that are made of two words each. If both words have elastic length, there are four length combinations, 2+2, 2+1, 1+2, and 1+1, where '2' is a disyllabic form and '1' a monosyllabic one. However, it is well known that 1+2 is disfavored in [N N] and 2+1 is disfavored in [V O] (Lü 1963). This is exemplified in (20) and (21), where '*' indicates a disfavored length combination.

(20)   In [N N], 1+2 is disfavored
          2+2   技术   工人   skill-skill   worker-man
          2+1   技术   工   skill-skill   worker
        *1+2   技   工人   skill   worker-man
          1+1   技   工   skill   worker
                                'skill   worker'

(21)   In [V O], 2+1 is disfavored
    2+2   种植   大蒜   plant-plant   big-garlic
   *2+1   种植   蒜   plant-plant   garlic
    1+2   种   大蒜   plant   big-garlic
    1+1   种   蒜   plant   garlic
                          'plant   garlic'

The homophony theory has no explanation for such length preferences. In contrast, the prosody theory can offer a natural account in terms of stress rules and Foot Binarity. First, let us consider the phonological requirements, given in (22).

(22)   Phonological requirements (boldface indicates stress)
    a. Compound Stress: [**N** N]
    b. Phrasal Stress: [V **O**]
    c. Foot Binarity: A syllabic foot must have two syllables
    d. Two monosyllables can form a foot in Chinese

The stress rules (22a) and (22b) are similar to those in English, where stress is assigned to the left in [N N] and right in [V O] (Chomsky and Halle 1968). They can be combined to a single rule, too, which has been called Stress-XP by Gussenhoven (1983) and Truckenbrodt (1995) and Nonhead Stress by Duanmu (1990; 2007). Foot Binarity is a well-known requirement in phonology, originally proposed by Prince (1980). Finally, (22d) is a very common process in Chinese, as has been noted by Shih (1986), Chen (2000), and Feng (1995).

Let us now consider the analysis of [N N], shown in (23), where S indicates syllable; boldface indicates stress.

(23)   Analysis of [**N** N]
    2+2   (**SS**)+(SS)
    2+1   (**SS**)+S   Second N has no stress and need not form a foot
  *1+2   (**S**)+(SS)   First N has stress and must form a foot
    1+1   (**S**+S)   Two monosyllables can form a foot

In 2+2, each noun forms a binary foot. In 2+1, the first noun forms a binary foot; the second noun has no stress and need not form a foot. In 1+2, the second noun forms a binary foot. The first noun has stress and must form a foot, but it has only one syllable, hence Foot Binarity is violated. In 1+1, the two nouns can form a foot. Next we consider [V O], shown in (24).

(24)   Analysis of [V **O**]
    2+2   (SS)+(**SS**)
  *2+1   (SS)+(**S**)   O has stress and must form a foot
    1+2   S+(**SS**)   V has no stress and need not form a foot
    1+1   (S+S)   Two monosyllables can form a foot

In 2+2, each word forms a binary foot. In 2+1, V forms a binary foot; O has stress and must form a foot, but it has only one syllable, hence Foot Binarity is violated. In 1+2, O forms a binary foot; V has no stress and need not form a foot. In 1+1, the two words can form a foot.

The analysis shows that stress is highly relevant for the use of disyllabic words. In contrast, the homophony theory offers no explanation for word length preferences.

4.4. Elastic words in written Chinese

According to Karlgren (1918), homophony is a problem only in spoken Chinese, which can only distinguish some 1,000 distinct syllables. In contrast, the number of distinct graphs in written Chinese has remained the same, totaling about 10,000. Therefore, there is no need for using disyllabic words in written Chinese. Now let us consider word length preferences in written Chinese, using the Lancaster Corpus of Mandarin Chinese (McEnery and Xiao 2004), which total 1.5 million graphs (1.3 million Chinese characters). The results in token frequencies of [N N] and [V O] are shown in (25) and (26), based on Duanmu (2012). The results in type frequencies are similar results and are omitted.

(25)   Word length patterns for [N N] in written Chinese (token frequencies)
        Length   Frequency
        2+2   21.8%
        2+1   21.3%
        1+2   1.1%
        1+1   55.8%

(26)   Word length patterns for [V O] in written Chinese (token frequencies)
        Length   Frequency
        2+2   16.2%
        2+1   1.8%
        1+2   19.1%
        1+1   62.8%

The results show that word length preferences in written Chinese is similar to those in spoken Chinese, where 1+2 [N N] is disfavored, so is 2+1 [V O]. The results are predicted by the prosody theory but not by the homophony theory.

4.5. Elastic words in classic Chinese

Because classic Chinese has three times as many distinct syllables as modern Standard Chinese, the homophony theory predicts that classic Chinese has a much lower percentage of elastic words. In contrast, the prosody theory predicts that prosodic requirements should be the same; therefore, classic Chinese should have just as many elastic words as modern Chinese does.

Dong (forthcoming) examines elastic words in Middle Chinese and Ancient Chinese. The result confirms the prediction of the prosody theory.

4.6. Summary

Let us summarize our discussion on whether various properties are related to elastic words. In (27) we compare how many things each of three theories can correctly account for.

(27)   A comparison of how many things each theory can correctly account for

	Homophony	Speech rate	Prosody
Minimal word	No	No	Yes
Homophone	No	No	Yes
Stress	No	Yes?	Yes
Spoken Chinese	Yes	Yes	Yes
Written Chinese	No	No	Yes
Modern Chinese	Yes	Yes	Yes
Classic Chinese	No	Yes	Yes
POS	No	No	Yes

The homophony theory can account for the high percentage of elastic words in modern spoken Chinese, but none of other effects. The speech rate theory can account for two more effects: the role of stress, if we assume that one needs to be slower when speaking a stressed word, and the high percentage of elastic words in classic Chinese has. In contrast, the prosody theory can account for all the eight properties. Therefore, it is by far the best theory on both the creation and the use of elastic words.

It is worth pointing out that elastic words do not always come from expanding a monosyllabic word to a disyllabic one. The reverse process can happen, too, where a disyllabic word or compound is truncated to a monosyllabic one, so that the latter acquires the full meaning of the original disyllabic one. Some examples are shown in (28).

(28)   Elastic words created by truncation

Original	truncated
轮船 'wheel boat (ship)'	轮 'ship'
荷兰 'Holland'	荷 'Holland'

A powered ship is originally called a 'wheel boat'. The compound is then truncated to 'wheel', which acquires the meaning of 'ship' as well. This process creates the elastic word 轮船 -轮 'ship-(boat)'-'ship'. Similarly, the disyllabic word 荷兰 'Holland' is truncated to 荷, which acquires the full meaning of 'Holland', creating the elastic word 荷兰-荷 'Holland'. The prosody theory can account for both processes of creating elastic words, expansion and truncation, because prosody requires both long and short forms. Specifically, disyllabic forms are needed for stressed positions. Monosyllabic forms are needed, too. For example, the ill-formed 1+2 [N N] can be repaired if the second N is truncated to a monosyllable, so that the [N N] becomes well-formed 1+1. In contrast, the homophony theory assumes that Chinese is already overwhelmed by homophony; therefore, it is hard to explain why Chinese keeps creating more of it.

5. Elastic words and poetry
The wide use of elastic words no doubt makes Chinese seem repetitive. Consider two examples in (29), where the first was given earlier. Repetitive words are shown in parentheses.

(29)   Repetitive words (indicated by parentheses) in two Chinese sentences
        我   刚(刚)   (看)见   他   (回)到   家(里)。
        'I   just-(just) (look)-see him (return)-reach home-(inside).'

14

保(持)　　　　车(厢)　　　　(清)洁
Ensure-(keep)　train-(box)　(clear)-clean

Chinese speakers do not find such sentences repetitive though, because they often ignore the internal structure of disyllabic units, an effect called 'Foot Shelter' by Duanmu (2007), and treat each the unit as a single word. Thus, to a Chinese speaker, the first sentence in (29) has six words, not ten, and the second sentence has three words, not six.

The presence of elastic words has made it very easy to compose rhythmic poems. Specifically, when a position needs a one syllable, we can use the short form of an elastic word, and when a position needs two syllables, we can use the long form of an elastic word. An example is shown in (30) and (31), where S and W represent strong and weak beats.

(30)　A poem with extra positions to fill (three missing syllables per line)
　　　S W　　　S W　　　S W
　　　保　　　　车　　　　洁
　　　皮　　　　扔　　　　外
　　　Ensure　train　　clean
　　　Peel　　　throw　　out

(31)　Elastic words can fill extra positions
　　　S W　　　　　S W　　　　　S W
　　　保(持)　　　　车(厢)　　　　(清) 洁
　　　(果) 皮　　　　扔(出)　　　　(窗) 外
　　　Ensure-(keep)　train-(box)　(clear)-clean
　　　(Fruit)-peel　　throw-(exit)　(window)-out

The poem has six syllable positions per line (three strong beats and three weak beats), but there are just three words each. If we use monosyllabic words, we are missing three syllables per line. If we use elastic words, all positions are filled.

Because it is so easy to make rhythmic poems in Chinese, poetry is used everywhere in China, as noted by Link (2013). One can find it in conversations, in speeches, in advertisements, in social media postings (such as WeChat), in slogans, in public notices, in government documents, etc. Everyone can make poems and everyone can appreciate them. In contrast, in English speaking societies, apart from nursery rimes and artistic performances (e.g. classic plays or rap music), it is rare to see or hear rhythmic language in public or private places, much less in government documents. In addition, even highly educated people may not be able to make or appreciate poetry. For example, Professor Rob Burling (personal communications), who has done research on nursery rimes, said that he had no intuition for iambic pentameter and could not appreciate its rhythm at all.

6. Concluding remarks
Elastic words are those whose length can vary from monosyllabic to disyllabic, without a change in meaning, as shown in (1). Elastic words have been observed in Chinese for a long time, at least since Karlgren (1918). However, both theoretical and empirical studies have been lacking. Theoretically, it is not clear whether elastic words can be distinguished from regular compounds,

nor has it been demonstrate what motivated the creation of elastic words. Empirically, it has not been shown how many elastic words there are in modern Chinese, or how many there were in classic Chinese.

We have offered a clear definition of elastic words that is simple, objective, and feasible. Based on the definition, we have calculated the percentage of elastic words in modern Standard Chinese. Our result shows that about 50% of Chinese morphemes, mostly monosyllabic words, have elastic length. The percentage is lower than most previous estimates but much higher than that in English (or in any other language, to our knowledge). In addition, we have argued that the percentage of elastic words in classic Chinese is just as high, contrary to the popular view that elastic words are the property of modern Chinese only. Finally, we have compared two theories of why elastic words are created in Chinese, the homophony theory and the prosody theory. According to the former, elastic words are created to avoid ambiguity caused by homophones. In contrast, according to the prosody theory, elastic words are created to satisfy Foot Binarity, so that stressed syllables can form a binary foot with another syllable. We have shown that most predictions of the homophony theory are incorrect, whereas all predictions of the prosody theory are borne out.

Biodata of San Duanmu:
San Duanmu is Professor of Linguistics, University of Michigan. He received his PhD in Linguistics from MIT in 1990 and has held teaching posts at Fudan University, Shanghai (1981-1986) and the University of Michigan, Ann Arbor (1991-present). He is the author of *The phonology of Standard Chinese*' (2nd edition, Oxford 2007) and *Syllable structure: the limits of variation*' (Oxford 2008). Personal website: www.umich.edu/~duanmu.

Biodata of Yan Dong:
Yan Dong is a Ph.D. candidate of Linguistics, University of Michigan. She received her MA in Linguistics and Applied Linguistics from Renmin University of China in 2008 and has enrolled in the Ph.D. program in University of Michigan since 2009. She expects to complete her dissertation titled '*The motivation of creating elastic word length in Chinese: homophone avoidance or phonology'* and receive her Ph.D. degree in Winter 2015.

References:

Bloomfield, Leonard. 1933. *Language*. New York: Holt.

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

Chen, Matthew Y. 2000. From tone to accent. *Cashiers de Linguistique—Asie Orientale* 29.2: 179-203.

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

Dong, Yan. Forthcoming. *The motivation of creating elastic word length in Chinese: homophone avoidance or phonology*. Doctoral dissertation, University of Michigan.

Duanmu, San. 1990. *A formal study of syllable, tone, stress and domain in Chinese languages*. Doctoral dissertation, MIT, Cambridge, Mass. Distributed by MIT Working Papers in Linguistics.

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

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

Duanmu, San. 2013. How many Chinese words have elastic length? In *Eastward flows the great river: Festschrift in honor of Prof. William S-Y. Wang on his 80th birthday*, ed. PENG Gang and SHI Feng, 1-14. Hong Kong: City University of Hong Kong Press.

Duanmu, San, and Yan Dong. 2014. Homophone density and word length in Chinese. To appear in *Capturing phonological shades within and across languages: inspirations from the Theoretical Phonology Conferences*, ed. Yuchau Hsiao and Lian-Hee Wee, XXX-XXX. Cambridge Scholars Publishing.

Feng, Shengli. 1995. *Prosodic structure and prosodically constrained syntax in Chinese*. Doctoral dissertation, University of Pennsylvania, Philadelphia, PA.

Guo, Shaoyu. 1938. Zhongguo yuci zhi tanxing zuoyong [The function of elastic word length in Chinese]. *Yen Ching Hsueh Pao* 24: 1-34.

Gussenhoven, Carlos. 1983. Focus, mode and the nucleus. *Journal of Linguistics* 19.2: 377-417.

Huang, Lijun, and San Duanmu. 2013. Xiandai Hanyu ci chang tanxing de lianghua yanjiu [A quantitative study of elastic word length in Modern Chinese]. *Yuyan Kexue* [Linguistic Sciences] 12.1: 8-16.

Jespersen, Otto. 1930. Monosyllabism in English: Biennial Lecture on English philology. *Proceedings of the British academy* 14: 341-368. London: British academy.

Jin, Wen. 2011. A statistical argument for the homophony avoidance approach to the disyllabification of Chinese. *Proceedings of the 23rd North American Conference on Chinese Linguistics* (NACCL-23), Volume 1, edited by Zhuo Jing-Schmidt, 35-50. University of Oregon, Eugene. http://naccl.osu.edu/sites/naccl.osu.edu/files/NACCL-23_1_03.pdf.

Karlgren, Bernhard. 1918. *Ordet och pennan i Mittens Rike*. Stockholm: Svenska Andelsförlaget.

Karlgren, Bernhard. 1923. *Sound & symbol in Chinese*. London: Oxford University Press.

Ke, Hang. 2007. *A study of monosyllabic and disyllabic usage in modern Chinese*. Doctoral dissertation, Institute of Linguistics, Chinese Academy of Social Sciences, Beijing.

Ke, Jinyun. 2006. A cross-linguistic quantitative study of homophony. *Journal of Quantitative Linguistics* 13.1: 129-159.

Kennedy, George A. 1951. The monosyllabic myth. *Journal of the American Oriental Society* 71.3: 161-166.

Kennedy, George A. 1955. The Butterfly Case. *Wennti papers* No. 8 (March, 1955), pp. 1-47. New Haven: Far Eastern Publications, Yale University.

Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: a functional reference grammar*. Berkeley and Los Angeles: University of California Press.

Li, Ningding. 1990. Dongci fenlei yanjiu shuolue [A brief discussion on verb classification]. *Zhongguo Yuwen* 1990.4: 248-257.

Link, E Perry. 2013. *An anatomy of Chinese: rhythm, metaphor, politics*. Cambridge, Mass.: Harvard University Press.

Lü, Shuxiang. 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.

McEnery, Tony, and Richard Xiao. 2004. The Lancaster Corpus of Mandarin Chinese: A corpus for monolingual and contrastive language study. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC) 2004*, ed. M. T. Lino, M. F. Xavier, F. Ferreire, R. Costa, & R. Silva, 1175-1178. Lisbon, May 24-30, 2004.

Xiandai Hanyu Cidian [Modern Chinese Dictionary]. 2005. Compiled by the Institute of Linguistics, Chinese Academy of Social Sciences. Beijing: Shangwu Yinshuguan.

Nettle, Daniel. 1995. Segmental inventory size, word length, and communication efficiency. *Linguistics* 33: 359 – 367.

Pan, Wenguo. 1997. Han Ying duibi gangyao [An outline comparison of Chinese and English]. Beijing: Beijing University of Languages and Cultures Press.

Prince, Alan. 1980. A metrical theory for Estonian quantity. *Linguistic Inquiry* 11:511-562.

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

Truckenbrodt, Hubert. 1995. *Phonological phrases--their relation to syntax, focus, and prominance*. Doctoral Dissertation, MIT, Cambridge, Mass.

T'sou, Benjamin K. 1976. Homophony and internal change in Chinese. *Computational Analysis of Asian & African Languages* 3: 67 – 86.

Wang, Canlong. 2002. Jufa zuhe zhong danshuang yinjie xuanze de renzhi jieshi [A cognitive explanation of the choice between monosyllabic and disyllabic words in syntax]. *Yufa Yanjiu he Tansuo*, Volume 11, ed. Zhongguo Yuwen Zhazhi She, 151-168. Beijing: Shangwu Yinshuguan.

Wu, Weishan. 2003. Shuangyinhua, yufahua, he yunlüci de saifenxi [A reanalysis of disyllabification, grammaticalization, and prosodic words]. *Hanyu Xuexi* 2003.2: 8-14.