

A JUDGMENT STUDY OF LENGTH PATTERNS IN CHINESE:
PROSODY, LAST RESORT, AND OTHER FACTORS

San Duanmu Shengli Feng* Yan Dong Yingyue Zhang
University of Michigan; The Chinese University of Hong Kong

ABSTRACT

In Chinese VO (verb-object) phrases, 2+1 (disyllabic + monosyllabic) is ill formed when other length patterns are available, such as 2+1 *种植蒜 *zhongzhi suan* vs. the well-formed 2+2 种植大蒜 *zhongzhi dasuan* ‘plant garlic’. However, 2+1 VO is acceptable when alternative length patterns are unavailable, such as 节约水 *jiyue shui* ‘save water’ and 喜欢钱 *xihuan qian* ‘love money’. The conditional acceptance of 2+1 VO is known as the ‘last resort’ effect. However, the predicted judgment difference has not been demonstrated experimentally. In addition, it is unclear whether native speakers find the last resort 2+1 to be as good as 2+2. Moreover, it is unclear what other factors may affect native judgment. To address the questions, we conducted a judgment experiment. Our study shows that (i) 2+2 VO is ranked the best, (ii) 2+1 is ranked the worst when a better form is available, and (iii) the last resort 2+1 is ranked between (i) and (ii). The result (iii) indicates a persistent effect of prosody, even for the last resort 2+1. In addition, we found a collocation effect between V and O. Moreover, we found considerable variation among the subjects, in that some subjects consistently gave higher scores than others, and for some expressions, there is a high degree of disagreement among subjects, which suggests that personal perspectives may play a role.

KEYWORDS

Word-length preference Prosody Acceptability experiment Elastic words Last resort

Acknowledgment We would like to thank helpful comments from an anonymous JCL reviewer and audiences of IACL-22 & NACCL-26 at the University of Maryland for their feedback.

1. INTRODUCTION

Chinese has many monosyllabic-disyllabic word pairs whose meanings are similar. Some examples are shown in (1), where parentheses indicate a semantically redundant syllable or morpheme.

(1) Synonymous word pairs in Chinese

Monosyllabic	Disyllabic	Character	Gloss
jian	(kan) jian	(看) 见	(look) see
ji	ji (shu)	技(术)	skill (technique)
mei	mei (tan)	煤(炭)	coal (charcoal)
suan	(da) suan	(大) 蒜	(big) garlic

In ‘look-see’, the meaning of ‘look’ is redundant, since to see something one has to look at it. In ‘skill-technique’, the meaning of ‘technique’ is repetitive. In ‘coal-charcoal’, the meaning of ‘charcoal’ is absent, since ‘coal-charcoal’ only means coal, not coal and charcoal. In ‘big-garlic’, the meaning of ‘big’ is absent, too, since garlic of any size is called ‘big-garlic’.

The above property has been known for a long time (Karlgren 1918; Guo 郭 1938; Kennedy 1951; Lü 吕 1963; Feng 冯 1996; Duanmu 2007, 2013; Huang 黄 and Duanmu 端木 2013; Dong 2015; Duanmu and Dong 2015). While the disyllabic forms look like compounds, they are different from true compounds in important ways. Specifically, we can define synonymous word pairs by two criteria, shown in (2).

(2) Defining synonymous word pairs in Chinese:

- a. The short and long forms share the same basic morpheme
- b. The short and long forms have the same meaning, as evidenced by mutual annotation in the dictionary.

The first criterion is self-evident. The second criterion can be met by mutual annotation in a dictionary. For example, in *Xiandai Hanyu Cidian* (Modern Chinese Dictionary) (ZSKYYSCB 2005), the annotation for the entry 煤 *mei* ‘coal’ is 煤炭 *meitan* ‘coal-(charcoal)’, and conversely, the annotation for the entry 煤炭 *meitan* ‘coal-(charcoal)’ is 煤 *mei* ‘coal’. The criteria can distinguish true compounds, which do not have a monosyllabic equivalent, from

those that do. Some examples are shown in (3), where parentheses indicate the actual meaning of a disyllabic form.

(3) Word pairs that do and do not pass the defining criteria

Word pair	Character	Gloss	(2a)	(2b)
le, gao-xing	乐, 高兴	happy, high-mood (happy)	No	Yes
gao, gao-xing	高, 高兴	high, high-mood (happy)	Yes	No
shu, shu-bao	书, 书包	book, book-bag (backpack)	Yes	No
mei, mei-tan	煤, 煤炭	coal, coal-charcoal (coal)	Yes	Yes
jian, kan-jian	见, 看见	see, look-see (see)	Yes	Yes

In the examples, ‘high-mood (happy)’ and ‘book-bag (backpack)’ are true compounds, whereas ‘coal-charcoal (coal)’ and ‘look-see (see)’ are not. To distinguish the two cases, we consider word pairs that satisfy the criteria in (2), such as those in (1), to be two length forms of the same word. We follow Guo 郭 (1938) and consider such words to have ‘elastic length’.

Many scholars believe that most Chinese words have elastic length (e.g. Karlgren 1918; Chao 1948; Kennedy 1951; Pan 潘 1997), and some quantitative estimates have been offered (Duanmu 2013; Huang 黄 and Duanmu 端木 2013). According to Duanmu and Dong (2015) and Dong (2015), based on a complete annotation of all Chinese morphemes in *Xiandai Hanyu Cidian* (Modern Chinese Dictionary) (ZSKYYSCB 2005), just 46.8% of Chinese morphemes have elastic length. The data are shown in table 1.

Table 1 POS (part of speech) counts, POS percentages (POS %), and percentages of words that are monosyllabic only (Mono %), polysyllabic only (Poly %), and elastic in length (Elastic %), in all senses in *Xiandai Hanyu Cidian* (Modern Chinese Dictionary).

POS	Count	POS %	Mono %	Poly %	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%
Measure	411	2.1%	91.0%	0.5%	8.5%
Onomatopoeia	291	1.5%	18.6%	74.6%	6.9%
Mood	121	0.6%	96.7%	1.7%	1.7%
Pronoun	116	0.6%	90.5%	6.9%	2.6%
Preposition	103	0.5%	97.1%	0.0%	2.9%
Interjection	79	0.4%	97.5%	2.5%	0.0%
Conjunction	69	0.3%	63.8%	0.0%	36.2%
Numeral	64	0.3%	96.9%	0.0%	3.1%
Affix	32	0.2%	100.0%	0.0%	0.0%
All	19,887	100%	46.0%	7.2%	46.8%

There are three kinds of length types: monosyllabic only, polysyllabic only, and elastic. They are exemplified in (4). As can be seen, most Chinese morphemes are also words.

(4) Three length types of Chinese morphemes

Length type	Percentage	Example
Monosyllabic	46.0%	水 <i>shui</i> 'water'
Polysyllabic	7.2%	珊瑚 <i>shanhu</i> 'coral'
Elastic	46.8%	煤(炭) <i>mei(tan)</i> 'coal(charcoal)'

The average percentage of words with elastic length, at 46.8%, is lower than previous estimates, such as those offered by Pan 潘 (1997), Duanmu (2013), and Huang 黄 and Duanmu 端木 (2013). Nevertheless, it represents a fairly large number and is not seen in other languages.

The availability of elastic word length in Chinese creates options of different length patterns. For example, two words of elastic length can yield four length patterns, 2+2, 2+1, 1+2, and 1+1, where 1 is monosyllabic and 2 is

disyllabic. However, not all length patterns are equally good. In particular, it is well known that 1+2 is usually bad for [NN] compounds and 2+1 is usually bad for [VO] (verb-object) phrases. This is exemplified in (5) and (6), which are supported by robust native judgment. For visual clarity, a space is added between the two words in the character column.

(5) Length patterns of [NN] compound: 1+2 is generally bad

Length	Pinyin	Character	Gloss
2+2	jishu gongren	技术 工人	'skill worker'
2+1	jishu gong	技术 工	
*1+2	ji gongren	技 工人	
1+1	ji gong	技 工	

(6) Length patterns of [VO] phrase: 2+1 is generally bad

Length	Pinyin	Character	Gloss
2+2	zhongzhi dasuan	种植 大蒜	'plant garlic'
*2+1	zhongzhi suan	种植 蒜	
1+2	zhong dasuan	种 大蒜	
1+1	zhong suan	种 蒜	

The length preferences are also confirmed by a corpus study (Duanmu 2012), which shows that 1+2 [NN] and 2+1 [VO] are both rare: their rates of occurrences are generally well below 2%.

It can be shown that the restriction against 1+2 [NN] and 2+1 [VO] is phonological (Feng 冯 1996; Lu and Duanmu 2002; Duanmu 2007; Feng 冯 2013). In particular, in [NN], phrasal stress goes to the first N, which forms a monosyllabic foot in 1+2, violating Foot Binararity, a well-known phonological restriction (Prince 1980). Similarly, in [VO], phrasal stress goes to O, which forms a monosyllabic foot in 2+1, violating Foot Binararity again.

What happens to words without elastic length? It has been observed that for [VO] phrases without alternative length options, 2+1 do occur and are judged to be fine (Lu and Duanmu 2002; Duanmu 2007; Feng 冯 2013). Some examples are shown in (7), where the verbs are polysyllabic only and the object nouns are monosyllabic only.

- (7) Acceptable 2+1 [VO] phrase, when no length options are available
 jieyue shui 节约 水 ‘save water’
 xihuan qian 喜欢 钱 ‘love money’
 yanjiu gui 研究 鬼 ‘study ghost’

Cases like those in (7) do not seem unusual. They show the ‘last resort’ effect, which is quite common in linguistics. In addition, Optimality Theory (Prince and Smolensky 1993) offers a straightforward analysis of last resort effects in terms of constraint interaction. For illustration, let us assume two constraints, shown in (8).

- (8) Two constraints on word length options in [VO]:
 Avoid 2+1: Do not use 2+1 [VO]
 Express: Express an idea in a phonological form
 Constraint ranking: Express >> Avoid 2+1

The analysis of a [VO] phrase with alternative length options is shown in (9) and that of a [VO] phrase with no alternative length option is shown in (10). The number of length options of a [VO] phrase can be determined with additional constraints, but for ease of exposition we omit additional constraints here.

- (9) Constraint-based analysis of [VO] with word length options

	‘plant garlic’	Express	Avoid 2+1
√	2+2 zhongzhi dasuan 种植 大蒜		
	2+1 zhongzhi suan 种植 蒜		*
√	1+2 zhong dasuan 种 大蒜		
√	1+1 zhong suan 种 蒜		
	(null)	*	

- (10) Constraint-based analysis of [VO] without word length options

	‘study ghost’	Express	Avoid 2+1
√	2+1 yanjiu gui 研究 鬼		*
	(null)	*	

In (9), both words have elastic length and there are four length pattern options, plus a null option (one without any phonological content). Three of the options have no constraint violation and are optimal outputs (actual expressions), indicated by \checkmark . In (10), neither word has elastic length and there is just one length option, plus the null option. The 2+1 form satisfies Express but violates Avoid 2+1. The null option violates Express but satisfies Avoid 2+1. Since Express is a higher ranked constraint, the 2+1 option is a better form, or the optimal form, and the actual output.

While the problem seems to have been solved in Optimality Theory, some questions remain. In particular, while the predictions of (9) have been confirmed indirectly by corpus data (Duanmu 2012), they have not been confirmed directly through a judgment study. In addition, the acceptability of last resort 2+1 [VO] has not been confirmed experimentally either. Therefore, a judgment experiment is needed to answer a number of questions, shown in (11).

- (11) Questions to be addressed in a judgment experiment
- a. Will last resort 2+1 [VO] be judged as good as 2+2 [VO], or will it be judged slightly worse than 2+2 [VO]?
 - b. Is 2+2 [VO] always judged to be perfect?
 - c. Is 2+1 [VO] with other length options always judged to be bad?
 - d. Are there other factors that affect the acceptability judgment?
 - e. Are there variations among native speakers?

To answer the questions, we carried out a judgment experiment on [VO] expressions, to be described in the next section.

Before we proceed, another question needs to be addressed. As a reviewer points out, not only are there last resort cases for [VO], such as 2+1 *jiēyue shuǐ* 节约水 ‘save water’, there seem to be last resort cases for [NN] as well, such as 1+2 *zhǐ lǎohǔ* 纸老虎 ‘paper tiger’, *mù dìbǎn* 木地板 ‘wood floor’, *xiàn yīyuàn* 县医院 ‘county hospital’, and *xiào língdǎo* 校领导 ‘school administrator’, none of which is problematic. Therefore, we might as well address last resort cases in both [VO] and [NN]. However, the case with [NN] is more complicated. First, as Duanmu (2012) observes, occurring examples of 1+2 [NN] are limited to two special cases, where the first N is either ‘material’ (as in ‘paper tiger’ and ‘wood floor’) or ‘possessor’ (as in ‘county hospital’ and

‘school administrator’). Second, when the first N is ‘material’ or ‘possessor’, 1+2 is fine even if 2+2 is available. For example, 2+2 *mutou diban* 木头地板 ‘wood floor’ is not obviously better than 1+2 *mu diban* 木地板 ‘wood floor’, and 2+2 *xuexiao lingdao* 学校领导 ‘school administrator’ is not obviously better than 1+2 *xiao lingdao* 校领导 ‘school administrator’. Third, when the first N is not ‘material’ or ‘possessor’, the last resort 1+2 [NN] does not seem to be good either. For example, *shui gongsi* 水公司 ‘water company’ does not seem acceptable, even though neither ‘water’ nor ‘company’ has elastic length. The difference between [NN] and [VO] is not completely clear and requires a separate study.

2. METHOD

The goal of our experiment is to obtain native judgment on various length patterns, in order to answer the questions raised above. We focus on two-word [VO] phrases and compare three types of length types, shown in (12) and (13). The verb is always disyllabic, without elastic length (i.e. it is disyllabic only). The object noun has three length types, where ‘Elastic1’ refers to the monosyllabic form of a word with elastic length, ‘Elastic2’ refers to the disyllabic form of a word with elastic length, and ‘1-only’ refers to a monosyllabic word without a disyllabic form. Since what varies is the length type of the object, we also use the object length type to refer to the length type of the [VO] in (13). For visual clarity, a space is added between the verb and the object noun in the character column in (13).

(12) Length types of [VO] phrases in our judgment experiment

Verb	Object length	Object type
Disyllabic only	monosyllabic	Elastic1
Disyllabic only	disyllabic	Elastic2
Disyllabic only	monosyllabic	1-only

(13) Examples of [VO] phrases in three length types

Type	Example	Character	Gloss
Elastic1	yanjiu suan	研究 蒜	‘study garlic’
Elastic2	yanjiu dasuan	研究 大蒜	‘study garlic’
1-only	yanjiu shui	研究 水	‘study water’

Verbs and nouns are selected from *Xiandai Hanyu Cidian* (Modern Chinese Dictionary) (ZSKYYSCB 2005) according to the criteria in (14).

(14) Criteria for the selection of test words

Words	Length	Criteria
6 verbs	disyllabic only	common; combinable with any noun
70 nouns	elastic	common; both length forms are free words
60 nouns	1-only	common; free words

Xiandai Hanyu Cidian (Modern Chinese Dictionary) annotates words with stylistic labels, such as ‘written’, ‘archaic’, ‘same as’, or ‘dialectal’. In addition, it indicates whether a morpheme is free or bound. A free morpheme (word) is given a part of speech label, whereas a bound morpheme is not. All words in our study are commonly used, in that they do not have special stylistic labels. In addition, all words, including both forms of elastic length, are free words. The six verbs and sample nouns are shown in (15)-(17). For ease of reading, we show the actual meaning of disyllabic words and omit their literal translation.

(15) List of six selected verbs

Word	Gloss	Character
yilai	‘depend on’	依赖
fenxi	‘analyze’	分析
aixi	‘take care of’	爱惜
baohu	‘protect’	保护
yanjiu	‘study’	研究
liaojie	‘understand’	了解

(16) Samples of 70 selected nouns with elastic length

Elastic1	Elastic2	Gloss	Character
xue	xueye	'blood'	血, 血液
wa	qingwa	'frog'	蛙, 青蛙
miao	simiao	'temple'	庙, 寺庙
ya	yachi	'tooth'	牙, 牙齿
...			

(17) Samples of 60 selected nouns that are 1-only (monosyllabic only)

Word	Gloss	Character
lang	'wolf'	狼
shu	'book'	书
xue	'snow'	雪
ren	'person'	人
...		

The selected nouns and verbs make a total of 1,200 [VO] combinations. The calculation is shown in (18).

(18) [VO] combinations made of selected words

Elastic1	6 verbs x 70 nouns = 420
Elastic2	6 verbs x 70 nouns = 420
1-only	6 verbs x 60 nouns = 360
Total	1,200

Each [VO] unit was placed in a carrier sentence. In order to make the sentences more natural, depending on the meaning of the verb, two slightly different carrier sentences were used. They are shown in (19) and (20), where 'Zhang San' is a generic name (similar to 'John' in English) and the target [VO] is underlined.

(19) Carrier sentence for the verbs 'depend on' and 'understand'

Example: 张三非常了解青蛙。

Zhang San feichang liaojie qingwa

Zhang San a lot understand frog

'Zhang San understands frogs a lot.'

- (20) Carrier sentence for the verbs ‘analyze’, ‘take care of’, ‘protect’, and ‘study’

Example: 张三应该研究青蛙。

Zhang San yinggai yanjiu qingwa

Zhang San should study frog

‘Zhang San should study frogs.’

The 1,200 sentences, written in Chinese without underline, were randomized and divided into ten test sheets, each containing 120 sentences. Each sheet was given to ten native speakers of Chinese. Each native speaker was instructed in Chinese to rank each sentence according to its ‘naturalness’ (shun bu shun kou 顺不顺口, literally ‘smooth to the mouth or not’), on a seven-point scale, where 1 is ‘completely unnatural’ (wanquan bu shunkou 完全不顺口, literally ‘not smooth to the mouth at all’) and 7 is ‘completely natural’ (wanquan shunkou 完全顺口, literally ‘fully smooth to the mouth’). Altogether 9,600 judgment scores were gathered (two lists were judged by 9 people each).

3. OVERALL RESULTS

According to previous predictions, the three kinds of [VO] phrases should be judged differently, depending on the length type of the object noun. This is shown in (21).

- (21) Predictions of previous analyses, based on length type of the object noun

Length type	Prediction
--------------------	-------------------

Elastic2	perfect or ranked best
----------	------------------------

Elastic1	bad or ranked lowest
----------	----------------------

1-only	perfect or acceptable, ranked the same as Elastic2, or between Elastic1 and Elastic2
--------	--

The results are shown in figure 1. The statistics in table 2 show that the rankings of all length types are significantly different from each other.

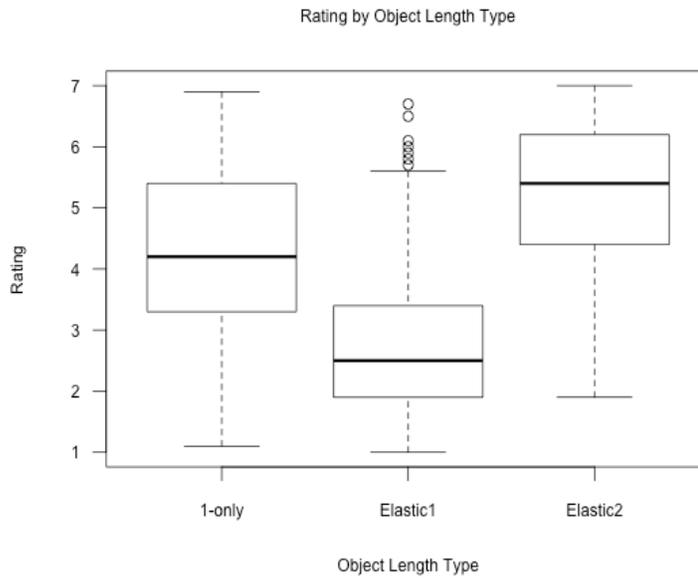


Figure 1 Ratings of [VO] phrases by three length types of the object noun. As previously predicted, Elastic2 is ranked the highest, Elastic1 the lowest, and 1-only in between. However, Elastic2 is not always judged to be perfect and Elastic1 not always judged to be impossible. In addition, there is a wide range of rankings within each length type, to be discussed below.

Table 2 Tukey multiple comparisons of means. The statistics show that the judgment ratings of all length types are significantly different from each other.

	difference	lower range	upper range	<i>p</i> value
Elastic1 vs. 1-only	-1.60	-1.81	-1.40	< 0.001
Elastic2 vs. 1-only	0.93	0.72	1.13	< 0.001
Elastic2 vs. Elastic1	2.53	2.33	2.73	< 0.001

Let us summarize the overall results in (22), which confirm the predictions of previous analyses, which we saw in (21).

- (22) Summary of the overall results
 - a. Elastic2 is better than Elastic1, which confirms the previous prediction.
 - b. 1-only is better than Elastic1 but not as good as Elastic2, which shows a persistent prosody effect even in the last resort situation.

Although the overall results offer little surprise, there is a wide range of acceptability rankings in every length types, which calls for an explanation, to be discussed next.

4. VARIATION IN ACCEPTABILITY RANKINGS

In this section, we consider variation in acceptability rankings. Specifically, we consider three issues, shown in (23).

- (23) Variations to be considered
 - a. High scorers vs. low scorers
 - b. Collocation effect
 - c. Inter-subject agreement

4.1. High vs. Low Scorers

There is a lot of variation among the average scores given by different subjects. For example, consider the average scores of the ten subjects for test set 2, shown in (24).

- (24) Subject effect: Average scores of 120 sentences by the subjects of test set 2

Subject	Average score
S1	3.0
S2	3.9
S3	1.6 'low scorer'
S4	3.0
S5	4.2
S6	6.2 'high scorer'

S7	2.5
S8	4.4
S9	3.6
S10	1.6 'low scorer'

It can be seen that S6 is a 'high scorer', giving an average of 6.2 out of 7 for the 120 sentences in test set 2. In contrast, S3 and S10 are 'low scorers', giving an average of just 1.6 out of 7 for the 120 sentences in the same test set. However, as long as each subject offers different rankings for different length patterns, their average ranking score does not affect the overall results.

4.2. Collocation Effect

Although we aimed to select verbs that can freely combine with any object noun, it is possible that some verbs are more productive than others, in the sense that they are more likely to combine with any noun. Let us call it the collocation effect. To find out, we examine acceptability rankings by verb, shown in figure 2 and table 3.

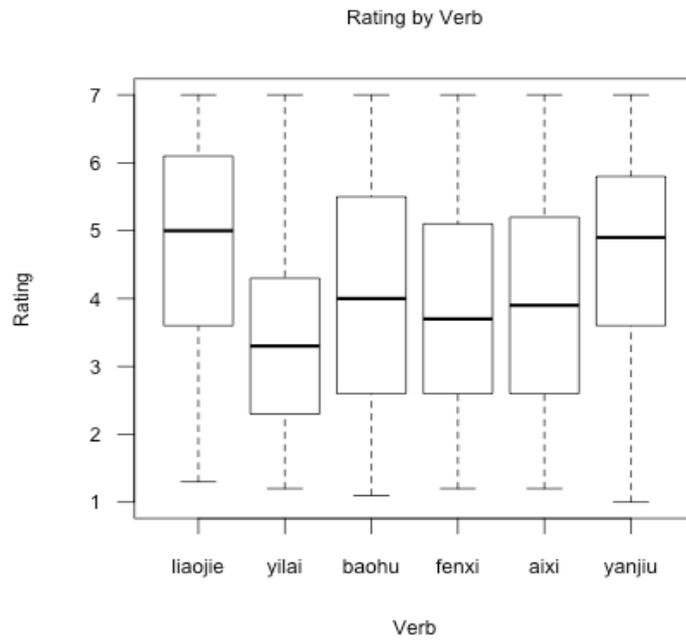


Figure 2 Ratings of [VO] phrases by verbs. The data show that *liaojie* ‘understand’ and *yanjiu* ‘study’ are most compatible with various object nouns, whereas *yilai* ‘depend on’ is the least compatible.

Table 3 Tukey multiple comparisons of means in figure 2 at 95% family-wise confidence level. The statistics show that *liaojie* ‘understand’ and *yanjiu* ‘study’ have significantly higher ratings than any other verb, indicating that they are most compatible with various object nouns. In contrast, *yilai* ‘depend on’ has significantly lower ratings than any other verb, indicating that it is the least compatible.

	difference	lower range	upper range	<i>p</i> value
yilai vs. liaojie	-1.33	-1.78	-0.89	< 0.001
baohu vs. liaojie	-0.64	-1.09	-0.20	< 0.001
fenxi vs. liaojie	-0.95	-1.40	-0.51	< 0.001
aixi vs. liaojie	-0.74	-1.19	-0.30	< 0.001
yanjiu vs. liaojie	-0.10	-0.54	0.35	0.99
baohu vs. yilai	0.69	0.24	1.13	< 0.001
fenxi vs. yilai	0.38	-0.06	0.83	0.14
aixi vs. yilai	0.59	0.14	1.03	< 0.01
yanjiu vs. yilai	1.23	0.79	1.68	< 0.001
fenxi vs. baohu	-0.31	-0.75	0.14	0.36
aixi vs. baohu	-0.10	-0.54	0.35	0.99
yanjiu vs. baohu	0.55	0.10	0.99	< 0.01
aixi vs. fenxi	0.21	-0.24	0.65	0.77
yanjiu vs. fenxi	0.85	0.41	1.30	< 0.001
yanjiu vs. aixi	0.64	0.20	1.09	< 0.001

The results show that there are significant differences among the verbs. For example, *yilai* ‘depend on’, which has the lowest average ranking, is significantly different from four other verbs ($p < 0.001$). Similarly, *liaojie* ‘understand’, which has the highest average ranking, is also significantly different from four other verbs ($p < 0.001$). To highlight the collocation effect, let us look at some outliers, which include those that receive extremely high average rankings (in z-score), shown in (25), and those that receive extremely low rankings, shown in (26).

(25) [VO] samples that receive extremely high average z-score rankings

Length	Chinese	Gloss	Av. z-score
1-only	爱惜书 aixi shu	take care of book	4.70
Elastic2	保护田地 baohu tiandi	protect land	3.56
1-only	依赖酒 yilai jiu	depend on wine	3.04

(26) [VO] samples that receive extremely low average z-score rankings

Length type	Chinese	Gloss	Av. z-score
Elastic1	依赖瓶 yilai ping	depend on bottle	-3.99
Elastic1	保护麦 baohu mai	protect wheat	-3.70
Elastic2	依赖棕榈 yilai zhonglü	depend on palm	-2.53
Elastic2	分析肚子 fenxi duzi	analyze belly	-3.99
Elastic2	分析蜂蜜 fenxi fengmi	analyze honey	-3.71
Elastic2	分析工厂 fenxi gongchang	analyze factory	-3.99

The high scores of the cases in (25) seem natural: they have good length types and reasonable meanings. The low scores of Elastic1 cases in (26) seem attributable to their bad length type. The low scores of Elastic2 cases in (26) require an explanation: there is no known problem of prosody in the length type of Elastic2; therefore, the low scores of cited cases must be due to their collocation. This seems to be true for ‘depend on palm’, the meaning of which is not very obvious. On the other hand, it is unclear why the three cases with ‘analyze’ are problematic.

4.3. Inter-Subject Agreement

Ideally, all subjects would rank the best expressions with the highest score and the worse expressions with the lowest score, and no expression should

receive the highest score from some subjects and the lowest score from other subjects. However, all cases are found. Some examples are shown in table 4, where subject scores are separated by a comma and X indicates a missing score.

Table 4 Subject rankings of five sample expressions, which show three cases: (i) a set of top scores, as in ‘protect antelope’, (ii) a set of bottom scores, as in ‘study hoe’, and (iii) a mixed set that includes both top scores and bottom scores. X indicates a missing score.

Length	VO	Scores (10 subjects)
1-only	了解 米 liaojie mi ‘understand rice’	1, 7, 1, 7, 7, 7, 1, 6, 1, 1
Elastic2	爱惜 寺庙 aixi simiao ‘take care of temple’	7, 6, 1, 7, 7, 1, 1, 1, 7, 7
1-only	依赖 鹿 yilai lu ‘depend on deer’	7, 7, 1, 1, 7, 4, 1, 1, 7, X
Elastic2	保护 羚羊 baohu lingyang ‘protect antelope’	7, 7, 7, 7, 7, 7, 7, 7, 7, 7
Elastic1	研究 锄 yanjiu chu ‘study hoe’	1, 1, 1, 1, 1, 1, 1, 1, 1, X

For expressions that show a high degree of inter-subject agreement, the average score is usually determined by length type. Expressions that receive consistently low scores are usually Elastic1. Expressions that receive consistently high scores are usually Elastic2. Expressions that receive consistently medium scores are usually 1-only.

On the other hand, there are expressions that show a high degree of disagreement among subjects. It seems to us that, in such cases, the subject judgment is based on personal experiences or perspectives, rather than the effect of prosody or grammatical structure. For example, the length type of ‘depend on deer’ is 1-only, whose prosody ought to be acceptable, evidenced by top scores from four of the subjects. On the other hand, semantically, it might not be obvious for what purpose one may reply on deer, especially for people who have not seen deer outside of a zoo (as is the case for most of the subjects). Therefore, they may have given a low score to the ‘naturalness’ of such an expression.

In order to exclude such subject-dependent ‘semantic’ effect, we recalculated the results in a new way. The procedure is shown in (27).

- (27) Procedure to reduce inter-speaker disagreement
- a. For each expression, obtain the average score of the 10 subjects.
 - b. Exclude scores that is more than 2 degrees above or below the average.
 - c. Exclude expressions that have less than 5 remaining scores.

For example, if the average score of an expression is 4, we will remove 1 and 7, while keeping 2, 3, 4, 5, and 6. If the average is 4.5, we will remove 1, 2, and 7, while keeping 3, 4, 5, and 6. Out of a total of 1,200 expressions, the procedure has removed 264, which means the majority are kept. The new results are shown in figure 3 and table 5.

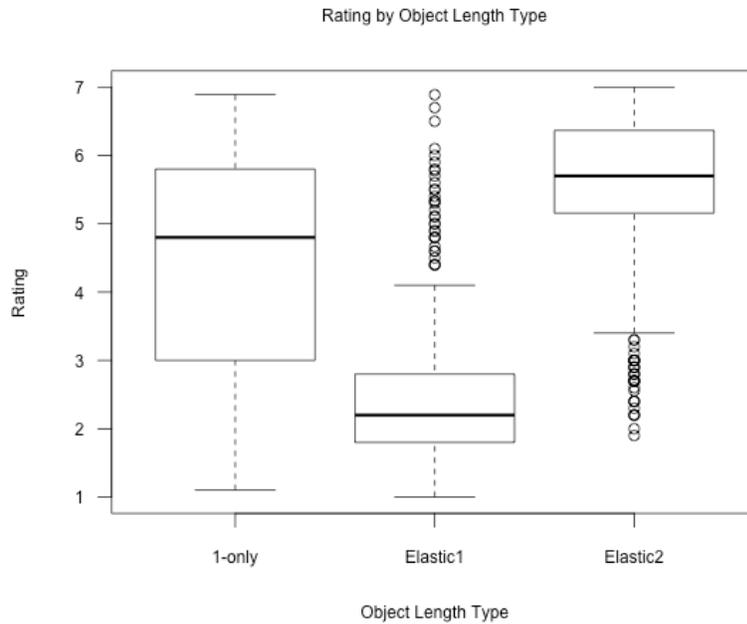


Figure 3 Ratings of [VO] phrases (936 in all) by three length types of the object noun, after removing 264 expressions that have a high degree of disagreement among the subjects. As previously predicted, Elastic2 is ranked the highest, Elastic1 the lowest, and 1-only in between. The variation in rating scores is now smaller for Elastic1 and Elastic2, although the variation in rating scores remains large for 1-only.

Table 5 Tukey multiple comparisons of means for the data in figure 3. The statistics show that the rating scores of all length types are significantly different from each other.

	difference	lower range	upper range	<i>p</i> value
Elastic1 vs. 1-only	-1.91	-2.15	-1.67	0.000
Elastic2 vs. 1-only	1.08	0.84	1.33	0.000
Elastic2 vs. Elastic1	3.00	2.77	3.22	0.000

The results again show that Elastic2 is ranked the best, Elastic1 the worse, and 1-only in between. We also see that, after excluding items with high inter-subject disagreement, the variation in rating scores is now smaller for Elastic1 and Elastic2, although the variation remains large for 1-only.

Besides the pooled results for all verbs, we also examined length effect on each of the six verbs, which are shown in figures 4-9 and tables 6-11.

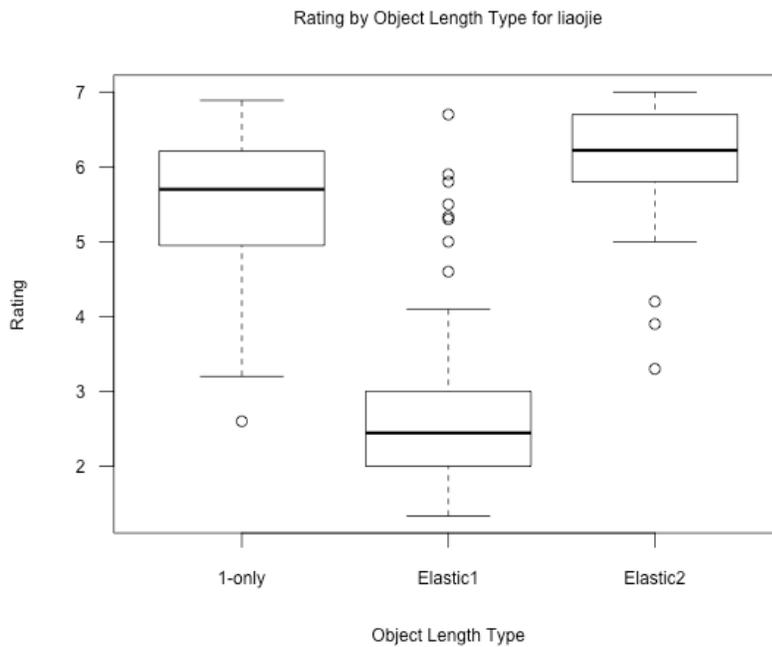


Figure 4 Ratings of phrases made by the verb *liaojie* ‘understand’ and three length types of the object noun, after removing expressions that have a high degree of disagreement among the subjects. As predicted, Elastic2 is ranked the highest, Elastic1 the lowest, and 1-only in between.

Table 6 Tukey multiple comparisons of means for the data in figure 4. The statistics show that the rating scores of all length types are significantly different from each other.

	difference	lower range	upper range	<i>p</i> value
Elastic1 vs. 1-only	-2.61	-3.13	-2.09	0.000
Elastic2 vs. 1-only	0.65	0.15	1.14	0.006
Elastic2 vs. Elastic1	3.26	2.78	3.73	0.000

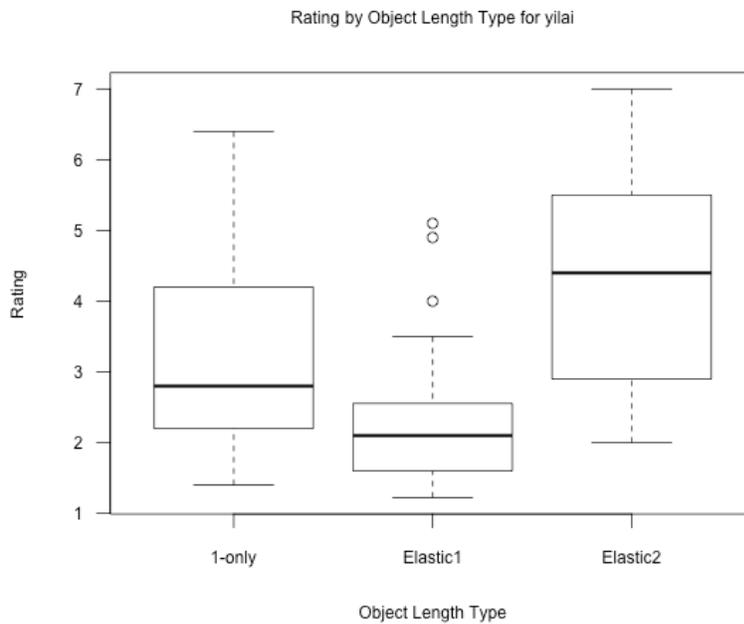


Figure 5 Ratings of phrases made by the verb *yilai* ‘depend on’ and three length types of the object noun, after removing expressions that have a high degree of disagreement among the subjects. While all ratings are lower than the averages for all verbs, Elastic2 is still ranked the highest, Elastic1 the lowest, and 1-only in between.

Table 7 Tukey multiple comparisons of means for the data in figure 5. The statistics show that the rating scores of all length types are significantly different from each other.

	difference	lower range	upper range	<i>p</i> value
Elastic1 vs. 1-only	-1.20	-1.79	-0.61	0.000
Elastic2 vs. 1-only	0.89	0.28	1.51	0.002
Elastic2 vs. Elastic1	2.09	1.54	2.63	0.000

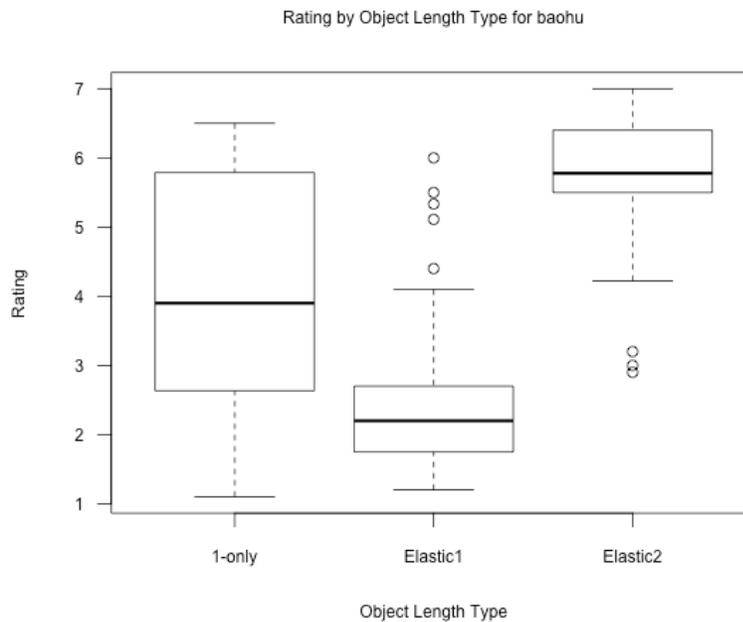


Figure 6 Ratings of phrases made by the verb *baohu* ‘protect’ and three length types of the object noun, after removing expressions that have a high degree of disagreement among the subjects. As predicted, Elastic2 is ranked the highest, Elastic1 the lowest, and 1-only in between.

Table 8 Tukey multiple comparisons of means for the data in figure 6. The statistics show that the rating scores of all length types are significantly different from each other.

	difference	lower range	upper range	<i>p</i> value
Elastic1 vs. 1-only	-1.52	-2.12	-0.92	0.000
Elastic2 vs. 1-only	1.74	1.23	2.35	0.000
Elastic2 vs. Elastic1	3.26	2.70	3.81	0.000

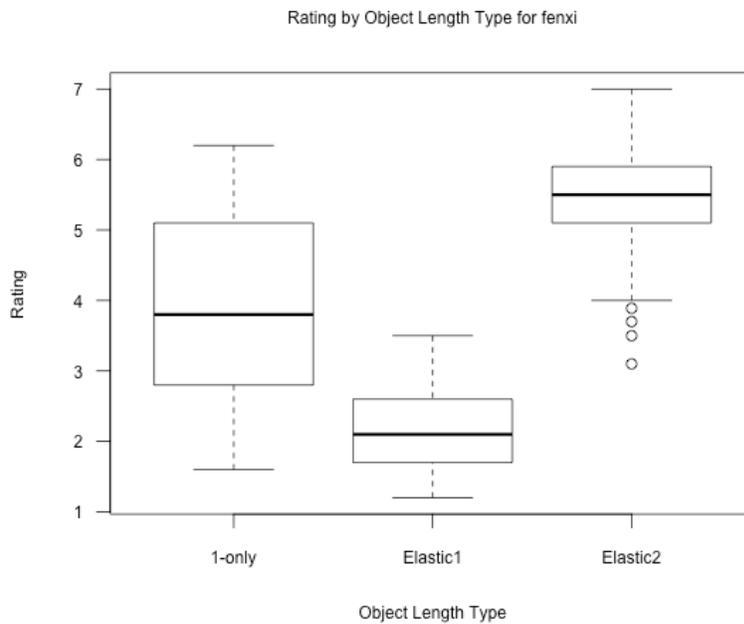


Figure 7 Ratings of phrases made by the verb *fenxi* ‘analyze’ and three length types of the object noun, after removing expressions that have a high degree of disagreement among the subjects. As predicted, Elastic2 is ranked the highest, Elastic1 the lowest, and 1-only in between.

Table 9 Tukey multiple comparisons of means for the data in figure 7. The statistics show that the rating scores of all length types are significantly different from each other.

	difference	lower range	upper range	<i>p</i> value
Elastic1 vs. 1-only	-1.69	-2.14	-1.24	0.000
Elastic2 vs. 1-only	1.57	1.10	2.04	0.000
Elastic2 vs. Elastic1	3.26	2.82	3.70	0.000

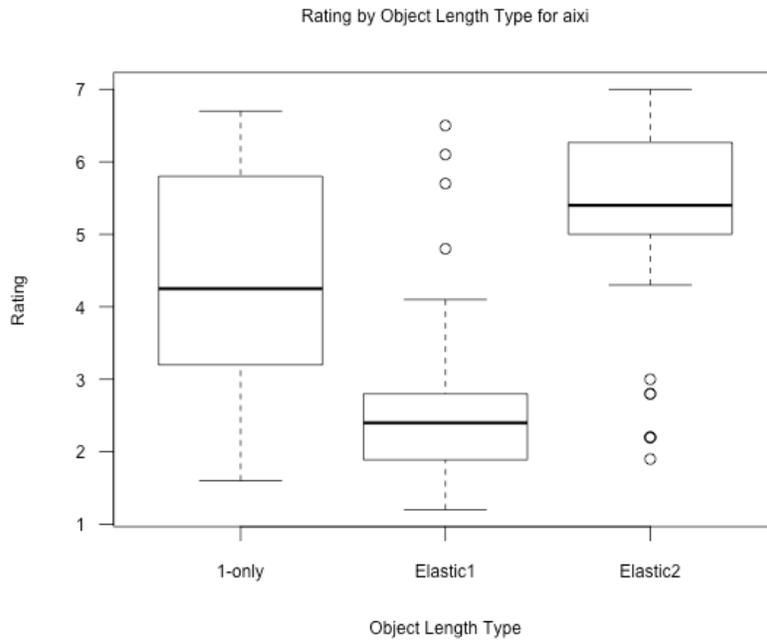


Figure 8 Ratings of phrases made by the verb *aixi* ‘take care of’ and three length types of the object noun, after removing expressions that have a high degree of disagreement among the subjects. As predicted, Elastic2 is ranked the highest, Elastic1 the lowest, and 1-only in between.

Table 10 Tukey multiple comparisons of means for the data in figure 8. The statistics show that the rating scores of all length types are significantly different from each other.

	difference	lower range	upper range	<i>p</i> value
Elastic1 vs. 1-only	-1.89	-2.47	-1.31	0.000
Elastic2 vs. 1-only	0.85	0.25	1.46	0.003
Elastic2 vs. Elastic1	2.73	2.16	3.31	0.000

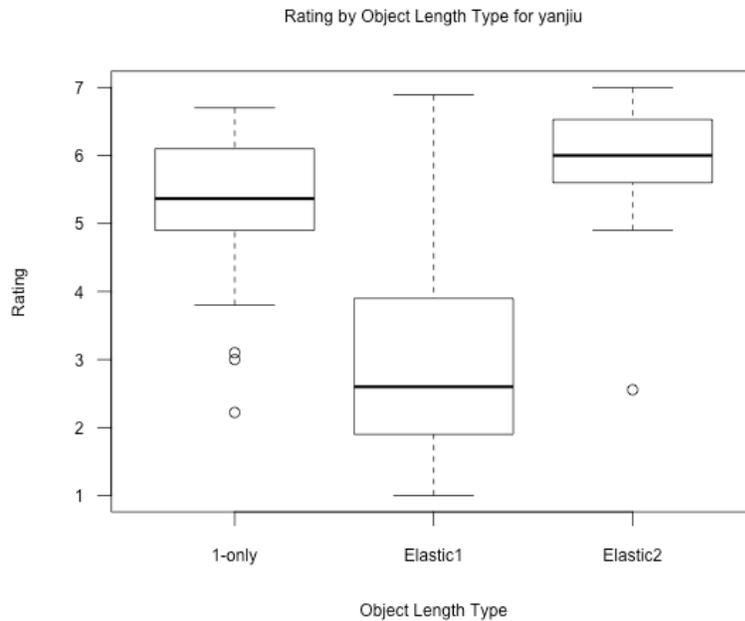


Figure 9 Ratings of phrases made by the verb *yanjiu* ‘study’ and three length types of the object noun, after removing expressions that have a high degree of disagreement among the subjects. As predicted, Elastic2 is ranked the highest, Elastic1 the lowest, and 1-only in between.

Table 11 Tukey multiple comparisons of means for the data in figure 9. The statistics show that the rating scores of all length types are significantly different from each other.

	difference	lower range	upper range	<i>p</i> value
Elastic1 vs. 1-only	-2.30	-2.80	-1.80	0.000
Elastic2 vs. 1-only	0.71	0.22	1.20	0.002
Elastic2 vs. Elastic1	3.01	2.56	3.47	0.000

In summary, after removing items for which there is a high degree of inter-speakers disagreement, the rating scores among the three length types become sharper, both for the pooled data of all verbs and for the data of individual verbs.

5. SUMMARY

We have offered a judgment study to verify two common views. First, in Chinese [VO] phrases, the length pattern 2+2 is better than 2+1, if both length options are available (Lü 吕 1963; Feng 冯 1996; Lu and Duanmu 2002; Duanmu 2007; and others). For example, because ‘garlic’ has two length forms, the disyllabic *dasuan* and the monosyllabic *suan*, there are two length options for ‘study garlic’: 2+2 *yanjiu dasuan* 研究大蒜, which we call ‘Elastic2’, and 2+1 *yanjiu suan* 研究蒜, which we call ‘Elastic1’, where Elastic2 is good and Elastic1 bad. It has been proposed, too, that 2+2 has a better phonological structure than 2+1 (Lu and Duanmu 2002; Duanmu 2007). The second view is that, when there is no length option, 2+1 is good, such as *yanjiu gui* 研究鬼 ‘study ghost’, which we call ‘1-only’, where ‘ghost’ has no disyllabic form. This is known as the ‘last resort’ effect in the literature (Prince and Smolensky 1993), which can override phonological requirements. The results of our study are summarized in (28).

- (28) Summary of results from the judgment experiment on [VO] phrases
- a. Elastic2 (2+2) is rated the best.
 - b. Elastic1 (2+1) is rated the worst.
 - c. 1-only (last resort 2+1) is rated between 2+2 and 1-only.
 - d. Some subjects consistently give higher scores than others.
 - e. For some expressions, there is a high degree of inter-speaker disagreement.

Our results offer confirmation of the common view that Elastic2 and 1-only are better than Elastic1. Our results also reveal some nuances not reported before. First, (28c) indicates a persistent effect of phonological requirement, even for the last resort case 2+1. Second, subject perspectives also plays a role, in that for a small number of expressions, some subjects give them highest scores, whereas other subjects give them the lowest scores. Finally, some subjects consistently give high scores, while other subjects consistently give low scores.

In summary, our study shows the value of experimental work, which not only can confirm what seems to be obvious intuitively, but reveal important nuances not available to casual introspection.

REFERENCES

- CHAO, Yuen-Ren. 1948. *Mandarin primer, an intensive course in spoken Chinese*. Cambridge, MA: Harvard University Press.
- DONG, Yan. 2015. The prosody and morphology of elastic words in Chinese: annotations and analyses. PhD dissertation, University of Michigan, Ann Arbor.
- DUANMU, San. 2007. *The phonology of Standard Chinese*. 2nd Edition. Oxford: Oxford University Press.
- _____. 2012. Word-length preferences in Chinese: a corpus study. *Journal of East Asian Linguistics* 21(1):89-114.
- _____. 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. 2016. Elastic words in Chinese. In *The Routledge Encyclopedia of the Chinese Language*, ed. CHAN Sin-Wai, 452-468. London and New York: Routledge.
- GUO, Shaoyu 郭绍虞. 1938. Zhongguo yuci zhi tanxing zuoyong 中国语词之弹性作用 (The function of elastic word length in Chinese). *Yen Ching Hsueh Pao* 燕京学报 24:1-34.
- FENG, Shengli 冯胜利. 1996. Lun Hanyu de yunlüci 论汉语的韵律词 (On the prosodic word in Chinese). *Zhongguo Shehui Kexue* 中国社会科学 1996(1):161-176.
- _____. 2013. *Hanyu yunlü jufaxue (zengding ben)* 汉语韵律句法学 (增订本) (Chinese prosodic syntax (expended edition)). Beijing: Shangwu Yinshuguan 商务印书馆.
- HUANG, Lijun 黄丽君, and San Duanmu 端木三. 2013. Xiandai Hanyu ci chang tanxing de lianghua yanjiu 现代汉语词长弹性的量化研究 (A quantitative study of elastic word length in Modern Chinese). *Yuyan Kexue* 语言科学 12(1):8-16.
- KARLGRÉN, Bernhard. 1918. *Ordet och pennan i Mittens Rike*. Stockholm: Svenska Andelsförlaget.
- KENNEDY, George A. 1951. The monosyllabic myth. *Journal of the American Oriental Society* 71(3):161-166.

- LU, Bingfu, and San Duanmu, 2002. Rhythm and syntax in Chinese: A case study. *Journal of the Chinese Language Teachers Association* 37(2):123-136.
- LÜ, Shuxiang 吕叔湘. 1963. Xiandai Hanyu dan shuang yinjie wenti chu tan 现代汉语单双音节问题初探 (A preliminary study of the problem of monosyllabism and disyllabism in modern Chinese). *Zhongguo Yuwen* 中国语文 1963(1):10-22.
- PAN, Wenguo 潘文国. 1997. *Han Ying duibi gangyao* 汉英对比纲要 (An outline comparison of Chinese and English). Beijing: Beijing Yuyan Wenhua Daxue Chubanshe 北京语言文化大学出版社.
- PRINCE, Alan. 1980. A metrical theory for Estonian quantity. *Linguistic Inquiry* 11(3):511-562.
- PRINCE, Alan, and Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. Ms., Rutgers University and University of Colorado.
- ZSKYYSCB (Zhongguo Shehui Kexueyuan Yuyan Yanjiu Suo Cidian Bianjishi 中国社会科学院语言研究所词典编辑室), ed. 2005. *Xiandai Hanyu Cidian* 现代汉语词典 (Modern Chinese Dictionary). 5th edition. Beijing: Shangwu Yinshuguan 商务印书馆.

汉语词长搭配在接受度：韵律，唯一选择及其他因素

端木三 冯胜利* 董岩 张瀛月

密歇根大学；*香港中文大学

摘要

汉语动宾短语 VO 的词长搭配中，2+1（双音节+单音节）的搭配通常不好，例如 2+1 * “种植蒜”不如 2+2 “种植大蒜”好。但是如果没有任何其他搭配可选，2+1 也可以接受，例如“节约水”、“喜欢钱”等。这种现象叫做“唯一选择”效果。不过，有的问题仍然有待探讨。比如，以上所说的两种接受度，能否通过实验加以证明？“唯一选择”的 2+1 的接受度，是否跟 2+2 的一样好？除了词长的搭配以外，是否还有其他因素影响语句的接受度？针对这些问题，我们进行了 VO 的接受度实验。结果显示（i）2+2 VO 的接受度最高；（ii）当有其他它搭配可选时，2+1 VO 的接受度最低；（iii）当无其他它搭配可选时，“唯一选择”的 2+1 VO 的接受度在（i）与（ii）之间。结果（iii）表明，韵律对语句的接受度始终存在影响，即使“唯一选择”的 2+1 VO 也受其影响。我们还发现，动词与名词之间的搭配关系也有影响。还有，参与者之间也存在差异：（i）有的人打分大大高于平均水平、有的大大低于平均水平，（ii）有少数结构得分差异很大，有的人给的是满分、而有的人给的确是最低分，说明个人观点也可能影响对语句的判断。

关键词

词长选择 韵律 接受度实验 弹性词 唯一选择