A Metrical Analysis of Chinese Regulated Verse*

San Duanmu
University of Michigan

0. Abstract

In generative phonology, English verse has been analyzed in terms of a set of metrical rules that map a line to the rhythmic template of the verse. In this article I use the model to examine Chinese verse. This study has both practical and theoretical goals.

On the practical side, I provide a description of 1460 lines of Chinese verse from the Tang Dynasty, whose word categories and syntactic structure have been labeled. The corpus can be used to construct and test various theories of meter, as I will demonstrate.

Theoretically, Chinese is of interest for several reasons. First, it is typologically unrelated to English and offers a good test for a general theory of stress and meter. Second, a preliminary study suggests that Chinese is different from English (Chen 1979). It is worth knowing whether the conclusion holds for a larger corpus. Third, constraints on English meter typically hold for polysyllabic words, while monosyllabic words are often free to occur in either strong or weak positions. Since most Chinese words are monosyllabic, there is the question of what constraints there are in Chinese meter.

I argue that there are important similarities between English and Chinese. In both languages stress plays a central role. In addition, whereas word stress in English is often more important than phrasal stress, in Chinese, where most words are monosyllabic, the

*For comments and judgments on some examples, I thank Uijin Ang, Rob Burling, Francois Dell, Shengli Feng, Nigel Fabb, Chris Golston, Carlos Gussenhoven, Morris Halle, Bruce Hayes, Ben Hermans, Yu-chau Hsiao, Zhiqiang Li, Yen-hwei Lin, Bingfu Lu, Michael Marlo, Fu-hsing Su, I-Ping Wan, Hongjun Wang, Jenny Wang, Debao Xu, Moira Yip, some anonymous reviewers, and audiences at University of Michigan, University of Iowa, and University of California Irvine, where earlier versions of this paper were presented. Not all of the comments could be adopted here though.

definition of the stress maximum depends on phrasal stress. Finally, while it is often assumed that a correlation exists between metrical tension and frequency in English verse, there is no evidence for such a correlation in the present corpus.

1. Introduction

In generative phonology, English verse has been analyzed in terms of a set of metrical rules that map the stress pattern of a line to the rhythmic (or metrical) template of the verse. Consider the analysis of the iambic pentameter. According to Halle and Keyser (1971: 169), the iambic pentameter has the basic template in (1), where W is a weak position, S a strong position, x an unstressed syllable, and X a stressed syllable (optional positions are ignored).

(1) Template: WSWSWSWSWS
Example: The curfew tolls the knell of parting day (xXxXxXxXxX)

The example shows an ideal line, where W and S are filled with unstressed and stressed syllables respectively. However, many lines depart from the ideal case. For example, in (2), some S positions are filled with an unstressed syllable and some W positions are filled with a stressed one.

(2) Batter my heart, three-person'd God, for you (XxxXXxXxx)
Prison my heart in thy steel bosom's ward (XxxXxxXXxX)

It is generally assumed that whether a line is metrical (or well formed) cannot be judged simply by the reader's intuition. Rather, if certain patterns are frequently used, they are assumed to be metrical (for the given poet). The job of the linguist is to find out the rules that make such lines metrical. Halle and Keyser propose that lines are not mapped to the metrical template directly, but via a set of 'correspondent rules' (also called 'metrical rules' in later works). Lines that can be so mapped are metrical. Lines that cannot are

unmetrical. The correspondence rules are defined in (3) and (4), from Halle and Keyser (1971: 169)

(3) Correspondence Rules (CR)

CR1 a. A position corresponds to a single syllable

b. Or to a sonorant sequence incorporating at most two vowels (immediately adjoining or separated by a sonorant consonant)

CR2 a. Fully stressed syllables occur in S positions only and in all S positions

b. Or Fully stressed syllables occur in S positions only but not in all S positions

c. Or stress maxima occur in S positions only but not in all S positions

(4) Stress Maximum (Halle and Keyser 1971: 169)

When a fully stressed syllable occurs between two unstressed syllables in the same syntactic constituent within a line of verse, this syllable is called a "stress maximum"

CR1a is the ideal case. CR1b allows two syllables to count as one in certain cases, which need not concern us. CR2a is the ideal case. CR2b allows an unstressed syllable to fill S, and CR2c allows a stressed syllable (but not a stress maximum) to fill W. A syllable is fully stressed if it carries the stress of a polysyllabic word or if it is a monosyllabic lexical word (instead of a functional or grammatical word). Given the correspondence rules, both lines in (2) can be mapped to the metrical template. This is shown in (5) and (6).

(5) Batter my heart, three-person'd God, for you (XxxXXXxXxx)

CR2c, CR2b, CR2c, CR2b (Complexity = 4)

(6) Prison my heart in thy steel bosom's ward (XxxXxxXXxX)

CR2c, CR2b, CR2b, CR2c (Complexity = 4)

In (5), there are four positions that depart from those of the ideal line (departing from CR2a); each departure is related to a clause in CR2. For example, the first syllable has stress but is not a stress maximum (it is not 'between two unstressed syllables'), and so according to CR2c it can fill W. Similarly, the second syllable in (5) is unstressed, but by CR2b it can fill S. Other deviant positions are analyzed in the same way.

The number of departures from the ideal line (from CR2a) provides a measure of 'metrical tension' (or 'metrical complexity'). Thus, while both (5) and (6) are metrical, they are both complex, with a tension of degree 4. It is hypothesized that lines with more tension are used less frequently than lines with less tension.

The Halle-Keyser model has some variations. First, there are different definitions of the stress maximum (or 'peak'). For example, Hayes (1989) defines it as a combination of two theories, Peak Theory and Compensation Theory. In contrast, Fabb (2002) defines it as in (7).

(7) Stress Maximum (Fabb 2002: 18)

A stressed syllable in a polysyllabic word is a *stress maximum* if it is preceded in a line by a syllable of lesser stress.

Second, some researchers use rules or constraints to derive the metrical template. For example, Fabb (2002) uses rules to construct the metrical template, following the Bracketed Grid Theory of Idsardi (1992) and Halle and Idsardi (1995), whereas Golston (1998) uses prosodic constraints to derive the template. Third, there are different conceptions of the correspondence rules and metrical tension. For example, Golston (1998) uses a constraint-based approach (Prince and Smolensky 1993) and accounts for metrical tension in terms of gradient constraint violations.

But common to all is the idea that there is a mapping relation between the stress pattern of a line and an abstract template, where the relation is specifiable by a set of rules or constraints. The template is abstract because as long as the mapping is successful, the line is metrical. There is, for example, no requirement for syllables in S positions to

be read with stress or for syllables in W positions to be read without stress (although such adjustments can occur in what Kiparsky 1975 calls 'the schoolboy manner' of recitation). This model has been called generative metrics. For convenience, I will simply call it the Halle-Keyser model.

Chinese verse raises two questions for the Halle-Keyser model. First, it is often reported that Chinese speakers do not have a clear intuition for stress in their language (Chao 1968: 38). Indeed, some linguists believe that Chinese has no stress (Hyman 1977, Selkirk and Shen 1991). If Chinese has no stress, what would the Chinese meter be based on? Second, most Chinese words are monosyllabic. Even if Chinese has stress, there is rarely a stress maximum (as defined above), and so almost any line should be metrical. The conclusion is rather surprising, if not unlikely. In any case, it remains to be verified.

A possible response to the questions is that each language can have a different verse convention. The stress-based theory works for English, but a different theory works for Chinese. Traditional studies generally follow this approach. For example, Wang (1958) provides a comprehensive discussion of rhyming, tone, and couplets, but stress and feet are discussed only under western poetry and westernized Chinese poetry, where Chinese poets are thought to imitate European feet.

In a seminal study, Chen (1979) proposes that Chinese regulated verse also has feet. However, whereas English feet are based on stress, Chinese feet are based tone. In addition, the syntactic tree plays a major role in Chinese. Specifically, the syntactic tree of an ideal line should match the prosodic tree. The prosodic tree of a seven-syllable line is shown in (8a), and the ideal syntactic tree is shown in (8b).

(8) Tree-matching theory (Chen 1979)

a. Prosodic tree for a seven-syllable verse line in Chinese

```
Line:        [                        ]
Half-lines:  [              ][        ]
Feet:        [M  M][M  M][M  M  M]        M = monosyllable
```

b. Syntactic tree for a seven-syllable verse line in Chinese

[[[M M][M M]][M M M]]

As in English, lines that deviate from the ideal one are still found. But unlike Halle and Keyser (1971), Chen does not offer metrical rules to map deviant lines to the prosodic template, and therefore, there is no specific prediction of whether a line is metrical or not. Nevertheless, Chen predicts that the more deviant a tree is, the less frequent it is used. The deviation of a tree is measured in terms of the number of nodes that differ between it and the ideal tree, to which we will return below.

Chen's work suggests that there is probably more in common between English verse and Chinese verse. However, Chen's work is based on just 100 verse lines, and so some claims remain hypothetical. For example, the tree-matching proposal remains to be verified, so is the predicted relation between the deviation of a tree and its frequency of use.

In this study I examine 1460 lines of Chinese regulated verse in order to find out to what extent Chinese verse is different from or similar to English verse, and whether there is a general theory of meter. In section 2 I describe the corpus. In section 3 I evaluate the proposal of Chen (1979). In section 4 I offer an analysis of Chinese verse based on phrasal stress and the stress maximum. In section 5 I discuss further issues. In section 6 I offer concluding remarks.

2. The corpus

The corpus is selected from an anthology of poems composed in the Tang Dynasty (618-907 AD). The anthology is known as 'The 300 Tang Poems', originally compiled in 1763 under the pen name Heng-Tang Tui-Shi 'Hermit of the Heng Pond'. There are many editions of the anthology. I use the one by Qiu (1976), which is the same edition used by Chen (1979). There are a total of 320 poems, which are serially numbered and divided into six categories, shown in (9).

(9)

Category	Serial #	Syllables per line	Lines per poem
5-Gu	1-45	5	variable
7-Gu	46-89	variable	variable
5-Lü	90-169	5	4
7-Lü	170-223	7	4
5-Jue	224-260	5	8
7-Jue	261-320	7	8

Each category may also contain a subcategory called Yuefu 'Music Academy (style)', which means the poems were in imitation of folk songs collected by the Music Academy in the Han dynasty. For example, 5-Gu includes 45 poems, of which 10 are 5-Gu Yuefu. Similarly, 7-Jue includes 60 poems, of which 9 are 7-Jue Yuefu.

The prefixes 5 and 7 refer to the number of syllables per line (except 7-Gu, see below). Gu (literally 'old') poems belong to *Gu-Ti Shi* 'Old-Style Verse'. Lü (literally 'regulated') and Jue (literally 'abrupt' or 'short') belong to *Jin-Ti Shi* 'Recent-Style Verse'. The difference is that Old-Style Verse has no requirements for tonal alternations, whereas Recent-Style Verse does (see section 5 on tonal requirements). In addition, Gu poems do not have consistent syllable or line length. In 5-Gu, each line has five syllables, but some poems have more lines than others. A few examples are shown in (10).

(10)

Poem No.	1	2	5	7	9	11
Total lines	10	8	14	6	24	16

For example, poem 1 has ten lines, poem 7 has six, and poem 9 has twenty-four. Similarly, 7-Gu poems usually have seven syllables per line, but not always. In addition, some 7-Gu poems are longer than others. Some examples are shown in (11).

(11)

Poem No.	Total lines	Syllables per line
46	4	5, 5, 7, 7
47	8	5, 5, 5, 5, 7, 7, 7, 7
79	47	3, 4, 9, 5, 5, 7, 7, 7, 7, 7, 9, 9, 9, 7, 7, …

Poem 47 has four lines, two of which have five syllables each and two have seven each. On the other hand, poem 79 has forty-seven lines, divided into four stanzas. The first stanza has fifteen lines, whose syllable counts are 3, 4, 9, 5, 5, 7, 7, 7, 7, 7, 9, 9, 9, 7, and 7 respectively.

Jue and Lü are both regulated verse and have a more consistent style, with a fixed number of syllables per line and a fixed number of lines per poem. Therefore, only they are included in the corpus. There are a total of 231 poems and 1460 lines, detailed in (12).

(12) Poems in the corpus

Category	# of Poems (serial #)	Total # of Lines
5-Lü	80 (90-169)	640
7-Lü	54 (170-223)	432
5-Jue	37 (224-260)	148
7-Jue	60 (261-320)	240
Total	231	1460

Next consider features of interest. There are many considerations in evaluating a poem. In this study I focus on metrical considerations. Based on discussions in the literature, the following items have been labeled in the corpus. Tonal considerations have been discussed extensively in previous works and are not included here (I return to tonal requirements in section 5).

(13) Properties examined in the corpus

a. Syllable count of each line
b. Word categories
c. Syntactic tree
d. Simplified syntactic tree

Syllables are included because they are the basic counting units in verse. Word categories are relevant for stress in English in that lexical words are stressed and grammatical words are often not. They are labeled for Chinese in order to see whether the same is true. The syntactic tree is included because it is mentioned in both English and Chinese. For example, the definition of the stress maximum in Halle and Keyser (1971) refers to the notion of a 'syntactic constituent'. Similarly, the metrical rules of Kiparsky (1977) refers to 'c-command' relations. In Hayes (1989), meter is analyzed in terms of prosodic categories, which also refer to the syntactic tree. In Chen's (1979) analysis of Chinese verse, the syntactic tree is matched with a prosodic tree. The simplified tree is included for convenience, where word categories are omitted. A sample poem is shown in (14).

(14)

Line	Syntax	Tree
170-1	[NN [A [[V NN] V]]]	[2 [1 [[1 2] 1]]]
170-2	[NN [A [V [NN N]]]]	[2 [1 [1 [2 1]]]]
170-3	[NN [[A V][B VV]]]	[2 [2 [1 2]]]
170-4	[[NN NN][A VV]]	[[2 2][1 2]]
170-5	[[NN VV][NN N]]	[[2 2][2 1]]
170-6	[[NN VV][NN N]]	[[2 2][2 1]]
170-7	[[NN NN][AA V]]	[[2 2][2 1]]
170-8	[[NN NN][V [N V]]]	[[2 2][1 2]]

The first column shows the serial number of a poem (according to Qiu 1976) and the line number. For example, 170-5 means the fifth line of poem 170. The second column shows the syntactic tree (bracketing) and word categories, where NN is a disyllabic compound noun (or occasionally, a disyllabic a noun), N a monosyllabic noun, B a monosyllabic

negation word, A a monosyllabic adjective or adverb, and so on. The third column shows the simplified tree, where 1 is a monosyllabic unit and 2 a disyllabic unit.

Some comments are in order. First, since Chinese lacks inflection, the category of a word is not always easy to determine. For example, *hong* can be an adjective 'red' or a noun 'redness', 'red color', or 'red thing'. Similarly, *dui* can be a verb 'to face' or a preposition 'towards'. The category B (negation) also contains several kind of words, such as *bu* 'not', *wei* 'not yet', *wu* 'without' or 'not have', and *xiu* 'don't (do)'. The category W (wh-word) is also mixed, such as *shei* 'who', *he* 'what', and *ji* 'how many'. Sometimes a nominal wh-word is simply coded as N. Overall, the number of B or W is rather small.

Second, syntactic distinctions that are irrelevant are not made. For example, since [A N] (such as 'yellow crane') and [N N] (such as 'smoke wave') are both compounds in Chinese (similar to *blackboard* and *rice wine* in English, see Dai 1992 and Duanmu 1998), they are both coded as NN. This decision avoids the problem of deciding whether the modifier is A or N. For example, *jin* can mean 'gold' or 'metal'. Now is *jin jia* 'golden armor' [A N], 'gold armor' [N N], or 'metal armor' [N N]? When a disyllabic noun or compound had a modifier, I tried to make a distinction between [A NN] and [N NN]. When in doubt, I favored [N NN] over [A NN]. However, since both A and N are content words, the decision is not crucial (see below). Similarly, adverbials are mostly coded as A or AA whether they are adverbs (such as *yi* 'already'), adjectives (such as *kong* 'empty/in vain'), or nouns (such as *zhao* 'morning/in the morning'). Also, adjective predicates are coded as verbs. An example is shown in (15), where 170-6 refers to poem 170, line 6.

(15) [[[fang cao] qiqi][yingwu zhou]] 170-6
[[[fragrant grass] thick][parrot island]]
[[NN VV][NN N]]
'The fragrant grass is thick, on the parrot island'

Here 'fragrant grass' is coded as NN, and '(is) thick' is coded as VV.

Third, [N NN], such as ‘official tour-person’ (94-1), and [NN N], such as ‘yellow-crane tower’ (170-2), are also compounds. In order to indicate their different tree structures, they are not represented as NNN. Similarly, [NN NN] compounds are not represented as NNNN. Another point of interest is that a verb pair can be either a phrase [V V] or a compound VV. The difference is exemplified in (16).

(16)	[V V] phrase, 171-3	VV compound, 109-5
	yu san	fen-san
	want disperse	split-disperse
	‘want to disperse’	‘separate’

Overall, the number of [V V] or VV is rather small.

Fourth, sometimes the syntax is either fragmented or not fully clear. For example, the line in (15) can be interpreted as ‘the fragrant grass on the parrot island is thick’ (a complete sentence), or ‘thick fragrant grass; parrot island’ (syntactic fragments). As I will argue below, higher-level syntactic relations are often inconsequential; in (15), the syntax [[NN VV][NN N]] is sufficient to determine the foot structure. Another example of fragmented syntax is shown in (17).

(17) [[feng ji][tian gao][[yuan xiao] ai]] 186-1

[[wind fast][sky high][[ape cry] sad]]
[[N V][N V][NN V]]

‘The wind is fast; the sky is high; the ape cries are sad’

Here it is hard to determine additional relation among the three parts, so they are coded with a flat structure.

Fifth, [Numeral Classifier Noun] structure is coded as [NN N]. The reason is that the numeral-classifier part forms a foot, similar to NN.

Finally, sometimes an expression has two (or more) interpretations. An example is shown in (18).

(18) [[jing xi][zi [bei qiu]]] 163-8

[[whole night][self [mourn autumn]]]
[NN [N [V N]]] or [AA [A [V N]]]

'All night, I myself mourn the autumn'

Here the word 'self' can either serve as the subject noun or as an adverbial (to mean 'by oneself'). Also, 'all night' can be viewed as a separate fragment NN or as an adverbial AA. Since such differences often do not affect the resulting foot structure, an arbitrary choice is made (usually according to Qiu's 1976 semantic translation).

Despite some uncertainties, most cases seem to be uncontroversial, and the overall results should be reasonably accurate. A summary of tree frequencies is given in (19) and (20). A summary of syntax frequencies is given in the Appendices.[1]

(19) 5-syllable lines (5-Jue and 5- Lü, 117 poems, 788 lines)

	Tree	Frequency	%
1	[2 [1 2]]	424	53.81%
2	[2 [2 1]]	193	24.49%
3	[1 [1 [2 1]]]	85	10.79%
4	[1 [1 [1 2]]]	32	4.06%
5	[1 [[1 2] 1]]	31	3.93%
6	[[2 2] 1]	12	1.52%
7	[[1 [1 2]] 1]	5	0.63%
8	[[2 1] 2]	4	0.51%
9	[[[2 1] 1] 1]	1	0.13%
10	[[1 2] 2]	1	0.13%
	TOTAL	788	100.00%

[1] The full corpus is available at [www.umich.edu/~duanmu/TangPoems04.xls].

(20) 7-syllable lines (7-Jue and 7- Lü, 114 poems, 672 lines)

	Tree	Frequency	%
1	[[2 2][1 2]]	215	31.99%
2	[[2 2][2 1]]	119	17.71%
3	[2 [1 [1 [2 1]]]]	63	9.38%
4	[2 [2 [1 2]]]	47	6.99%
5	[[1 [1 2]][1 2]]	43	6.40%
6	[1 [1 [2 [1 2]]]]	26	3.87%
7	[1 [1 [2 [2 1]]]]	20	2.98%
8	[[1 [1 2]][2 1]]	19	2.83%
9	[[[2 1] 1][1 2]]	18	2.68%
10	[1 [[1 2][1 2]]]	18	2.68%
11	[2 [1 [1 [1 2]]]]	15	2.23%
12	[[[2 1] 1][2 1]]	13	1.93%
13	[2 [2 [2 1]]]	13	1.93%
14	[2 [1 [[1 2] 1]]]	9	1.34%
15	[2 2 [2 1]]	4	0.60%
16	[1 [1 [[2 2] 1]]]	3	0.45%
17	[[[1 2] 1][1 2]]	2	0.30%
18	[[[2 2] 1] 2]	2	0.30%
19	[[[2 2] 2] 1]	2	0.30%
20	[1 [[1 2][2 1]]]	2	0.30%
21	[1 [1 [[[2 1] 1] 1]]]	2	0.30%
22	[1 [1 [1 [[1 2] 1]]]]	2	0.30%
23	[1 [1 [1 [1 [1 2]]]]]	2	0.30%
24	[1 [1 [1 [1 [2 1]]]]]	2	0.30%
25	[2 [[[2 1] 1] 1]]	2	0.30%
26	[2 [[2 2] 1]]	2	0.30%
27	[[[1 2] 1][2 1]]	1	0.15%

28	[[1 2][1 [1 2]]]	1	0.15%
29	[[2 [2 2]] 1]	1	0.15%
30	[[2 1][1 [1 2]]]	1	0.15%
31	[[2 1][1 [2 1]]]	1	0.15%
32	[1 [[1 [2 2]] 1]]	1	0.15%
33	[2 [1 [2 2]]]	1	0.15%
	TOTAL	672	100.00%

The data show a wide range of syntactic trees. For any analysis that assumes a relation between syntax and metrics, directly or indirectly, the challenge is to account for the variation.

3. Evaluation of Chen (1979)

Given the corpus, metrical proposals can be evaluated quite easily. In this section we discuss the proposal of Chen (1979), which focuses on seven-syllable lines. Chen's main idea is that the syntactic tree should match the prosodic tree. Assuming binary branching, a seven-syllable line has two possible prosodic trees, abbreviated as [[2 2][2 1]] and [[2 2][1 2]] and represented in (21).

(21) Two prosodic trees for seven-syllable lines (M = monosyllable)

	[[2 2][2 1]]	[[2 2][1 2]]
Line:	[ ]	[ ]
Half-lines:	[ ][ ]	[ ][ ]
Feet:	[M M][M M][M M][M]	[M M][M M][M][M M]

The trees differ in the last three syllables ([2 1] vs. [1 2]). Chen proposes that the lines of a poem are divided into couplet pairs, where each pair should have the same prosodic tree and adjacent pairs should have different prosodic trees. Therefore, four-line poems have

two patterns of alternation, shown in (22). In eight-line poems, the four-line pattern will be repeated another time.

(22)

Alternation A	Alternation B
[[2 2][2 1]]	[[2 2][1 2]]
[[2 2][2 1]]	[[2 2][1 2]]
[[2 2][1 2]]	[[2 2][2 1]]
[[2 2][1 2]]	[[2 2][2 1]]

A prediction of the analysis is that in seven-syllable lines, [[2 2][2 1]] and [[2 2][1 2]] trees will be most common. In addition, the two kinds of trees should be about equal in number. Now consider the four most frequent trees, shown in (23).

(23) Top four most frequent trees in 7-syllable lines

	Tree	Frequency	%
1	[[2 2][1 2]]	215	31.99%
2	[[2 2][2 1]]	119	17.71%
3	[2 [1 [1 [2 1]]]]	63	9.38%
4	[2 [2 [1 2]]]	47	6.99%

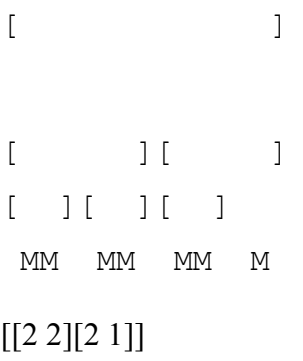
While [[2 2][2 1]] and [[2 2][1 2]] are indeed the two most common trees, they differ sharply in number. This means that there is no systematic couplet alternation in the last three syllables.

Perhaps because of this problem, Chen (1979: 396) suggests that the last three syllables should probably be treated as one foot. The same view is expressed in Chen (1980: 17, 37). However, no evidence for the trisyllabic foot is provided.

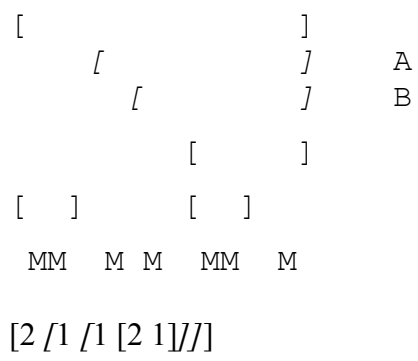
Next consider the relation between the deviation of a tree and the frequency of its use. Chen (1979) proposes that the more deviant a tree is, the less frequently it is used. The deviation of a tree is measured in terms of how many nodes (or pairs of parentheses) differ between it and an ideal tree (for the sake of argument, we ignoring [1 2] vs. [2 1] in

the last three syllables between the two ideal trees). An example is shown in (24).

(24) Ideal tree

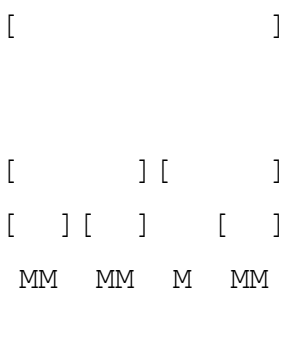


Deviant tree (two bad nodes)

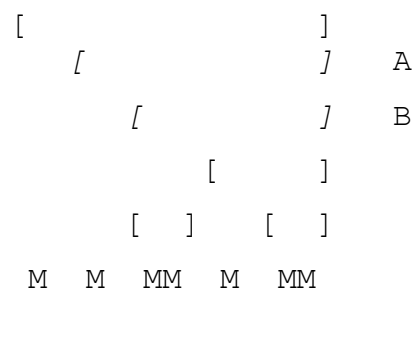


The deviant tree has two bad nodes (shown with italicized parentheses), labeled A and B. The domain of A should be over the first four syllables. The domain of B should be over the third and fourth syllables. Similarly, consider (25).

(25) Ideal tree



Deviant tree (two bad nodes)



The deviant tree again has two bad nodes. The domain of A should be over the first four syllables, and the domain of B should be over the first two syllables.

Having discussed the measure of deviation, let us consider the frequency data. Several patterns are shown in (26), where deviant parentheses are italicized.

(26)	Pattern	Tree	Frequency	%	Bad nodes
	1	[[2 2][1 2]]	215	31.99%	0
	2	[[2 2][2 1]]	119	17.71%	0
	3	[2 /1 /1 [2 1]//]	63	9.38%	2
	4	[2 /2 [1 2]/]	47	6.99%	1
	5	[[1 [1 2]][1 2]]	43	6.40%	1
	6	[1 /1 /2 [1 2]//]	26	3.87%	2
	18	[/[2 2] 1/ 2]	2	0.30%	1
	19	[/[2 2] 2/ 1]	2	0.30%	1
	26	[2 //2 2/ 1/]	2	0.30%	2

The results are hard to account for in Chen's analysis. For example, patterns 1 and 2 have no deviant node and are indeed the two most common trees, yet they differ widely in frequency. Similarly, pattern 3 (the third most common) has two deviant nodes, yet it is more frequent than patterns 4, 5, 18, and 19, which only have one deviant node. Moreover, patterns 3 and 26 both have two deviant nodes, yet one occurs sixty-three times while the other just two.

Although the correlation between the deviation of a tree and its frequency is not fully supported, it is striking that the top two patterns far exceed others. In five-syllable lines, the top two patterns cover 79% of all lines. In seven-syllable lines, the top two patterns cover 50% of all lines. This still requires an explanation. A possible answer is that it is a consequence of grammar. Consider the top two patterns in five-syllable lines, repeated in (27).

(27)	Pattern	Tree	Frequency	% of total (788 lines)
	1	[2 [1 2]]	424	53.81%
	2	[2 [2 1]]	193	24.49%

The most common sentence is perhaps [S [V O]], where S is the subject, V the verb, and O the object. Now, if most nouns (including compound nouns) are disyllabic and most verbs are monosyllabic, [2 [1 2]] should be the most common pattern. To test the

hypothesis, we need to look at the length of nouns and verbs. In addition, we need to look at [2 [1 2]] to see whether the most common lines are indeed [S [V O]]. First, consider word length in five-syllable lines. Following Dai (1992) and Duanmu (1998), [N NN], and [NN N] are counted as compounds. The results are shown in (28).

(28)

Category	# of Words	% of monosyllabic words
N	362	28.53%
NN	756	
N-NN	16	
NN-N	135	
V	952	95.39%
VV	46	

Polysyllabic compound nouns constitute over 70% of all nouns, most of which being disyllabic. In contrast, monosyllabic verbs constitute over 90% of all verbs. Similar proportions are found in seven-syllable lines. In other words, most nouns are disyllabic and most verbs are monosyllabic.

Next consider sentence types in [2 [1 2]] and [2 [2 1]]. The five most frequent sentence structures for each are shown in (29) and (30).

(29) Top five sentence types in [2 [1 2]] (total = 424 lines)

Sentence	Frequency
[NN [V NN]]	145
[[V N][V NN]]	34
[AA [V NN]]	24
[NN [A [V N]]]	20
[[A V][V NN]]	11

(30) Top five sentence types in [2 [2 1]] (total = 193 lines)

Sentence	Frequency
[NN [NN V]]	41
[NN [NN N]]	38
[NN [AA V]]	17
[AA [NN N]]	13
[[N V][NN V]]	10

The sentence type [NN [V NN]] in [2 [1 2]], which is [S [V O]], exceeds all other types by far. This seems to be a major factor that has made [2 [1 2]] the most common five-syllable pattern.

In summary, the bigger database reveals a number of empirical effects that are missed under Chen's account. In particular, there is insufficient evidence for the proposal that the syntactic tree should match the prosodic tree, nor is there enough evidence for the proposal that the most common lines have better metrical structure than less common ones. Instead, the dominance of certain line types seems to be just that, namely, they are the most common sentences, in particular the [S [V O]] sentence. Such findings cannot be understood unless we re-conceptualize the entire approach.

Given the wide range of syntactic trees and the lack of a clear correlation between the tree structure and its frequency of use, one might wonder if syntax is relevant for Chinese verse at all; this is a point raised by Schlepp (1980a, b). However, I will argue in section 4 that syntax does play a role in Chinese verse, as it does in English.

4. The present analysis

In this section I propose that, with minor modifications, the Halle-Keyser model can be applied to Chinese verse. I outline the proposal in section 4.1, followed by an illustration of the analysis in section 4.2. In section 4.3 I discuss whether there is a relation between metrical tension and frequency of occurrence.

4.1. Outline of the analysis

Following the Halle-Keyser model, I will discuss the following issues.

(31) Metrical requirements
  a. Templates for five-syllable lines and seven-syllable lines
  b. Stress maximum
  c. Relation between the stress pattern of a line and the template

(32) Stress rules
  a. Word and compound stress
  b. Phrasal stress

The templates I propose for five-syllable lines and seven-syllable lines in Chinese verse are given in (33), where the last position is an empty beat, or rest, which has been proposed in many analyses (Liu 1927, Burling 1964, Liberman 1975, Schlepp 1980a, and Hayes 1995).

(33) Templates for Chinese verse
  a. Five-syllable lines: SWSWSWSØ
  b. Seven-syllable lines: SWSWSØ

There are several arguments for the templates. First, as I will show, stress maxima occur in odd-numbered positions, which are S positions in the present analysis. Second, most lines start with a disyllabic compound. I will argue that, as in English, the first syllable of a compound has greater stress. Since such compounds 'strongly favor' SW in English (Kiparsky 1977: 191), the null hypothesis is that they favor SW in Chinese as well. Third, the final syllable is S, in agreement with the fact that it is the rhyming syllable, which should be stressed. Fourth, with the empty beat, both templates contain binary feet only:

the five-syllable line has three binary feet (SW)(SW)(SØ), and the seven-syllable line has four (SW)(SW)(SW)(SØ). In addition, the empty beat helps the analysis of the final trisyllabic unit, to be seen below.

The templates are similar to those proposed by Liu (1927) and Schlepp (1980a, b), whose arguments are based on recitation and modern Chinese verse. However, as Kiparsky (1975: 585) points out, there should perhaps be a 'sharp distinction between the metrical organization of a poem and the way it is cited', because stress 'is highly variable in speech'. In addition, the manner of recitation largely reflects the reader's judgment, which may or may not be the same as that of the original poet. Reference to modern Chinese verse should be avoided, too, unless the same pattern can be observed in the original verse. In the present analysis, no reference to recitation or modern verse is used.

Chen (1979) and Yip (1980) propose that the Chinese foot is iambic, so that the template is WSWS…. They offer two arguments, shown in (34).

(34) Arguments for WSWS… template (Chen 1979, Yip 1980)

a. In recitation, even-numbered syllables are read with a longer duration.

b. In Chinese regulated verse, even-numbered syllables have fewer choices for tone that odd-numbered positions.

The first argument is based on recitation, which should be avoided. The second argument relies on the assumption that the restriction on tone is placed on S positions, rather than W positions, but the assumption remains a hypothesis (we return to this issue in section 5). Thus, there is no clear evidence for the iambic analysis.[2]

Next consider stress. Since there is no traditional literature on stress in classic Chinese, argument is indirect.[3] I will make two assumptions. First, I assume that, as in

---

2 Perhaps because of the difficulty in deciding the stress pattern of the Chinese foot, Chen (2000) suggests that the Chinese 'foot' is something new—a disyllabic unit with two equal parts, where neither is stronger than the other. To distinguish it from a regular foot (which has stress), Chen introduces a new term for it, 'a minimal rhythmic unit'.

3 It has been shown that languages that appear to lack stress may nevertheless have foot structure. For

other languages, grammatical words in Chinese have less stress than content words. Second, I assume the null hypothesis that, in the absence of counter-evidence, classic Chinese has the same stress rules as other languages, such as English or modern Chinese.

First, consider word stress. Chinese has a very small number of simple disyllabic words, such as *manao* 'amber', and *luobo* 'raddish'. As Lin (1994) points out, in some cases the second syllable can become completely unstressed, such as *luobo* in Standard Chinese. On the other hand, there is no case where the first syllable becomes unstressed. This suggests that main stress in disyllabic words is on the first syllable. This is similar to disyllabic nouns in English, where stress generally falls on the first syllable.[4]

Next consider compound and phrasal stress, where I assume that Chinese has the rules as in English. Chomsky and Halle (1968) propose two rules, the Nuclear Stress Rule, which assigns stress to the right, and the Compound Stress Rule, which assigns stress to the left. Some examples are shown in (35) and (36).

(35) Nuclear Stress Rule: stress on the right

[V O]	buy CARS
[P N]	in SCHOOL
[D N]	the CAR
[N's N]	John's FRIEND
[A N]	red CARS

---

example, both Japanese and Chinese appear to lack stress, but Poser (1990) has argued that Japanese has moraic feet, and Shih (1986) and Chen (2000) have argued that Chinese has syllabic feet. However, Poser does not discuss stress in Japanese, nor does Shih or Chen in Chinese. It is possible that some languages have stressless feet, as Chen (2000) suggests. If so, such languages may have different metrics; for example, their metrical template may be made of equal positions, such as XX/XX/…, instead of S vs. W positions. However, I will argue that Chinese does have stress.

[4] One might wonder if the Iambic-Trochaic Law (Hayes 1985) can help decide stress in Chinese. In particular, if heavy syllables attract stress, we expect the iambic foot. If heavy syllables do not attract stress, we expect the trochaic foot. The case seems to be in between. Heavy syllables do not attract stress in the sense that two adjacent heavy syllables still form one foot, instead of two. On the other hand, light syllables avoid stress, in the sense that they only occur in the second position of a disyllabic word.

(36) Compound Stress Rule: stress on the left

[N N] OIL lamp

There are some problems, however. First, ideally, one would like to see a connection between the rules. Second, some subtle stress differences are not captured. For example, while (35) gives the same stress pattern for VO and AN, where the first word is not stressed and the second is, traditional descriptions give them different stress patterns. For example, according to Kenyon and Knott (1944) and Jones (1950), main stress in VO (such as *take care*) is indeed on O, but in AN (such as *important aid*, *practical study*, *Red Cross*, and *real pleasure*) the two words have equal stress.

Two revisions to the Chomsky-Halle model have been proposed. According to Duanmu (1990a, 2000), compound and phrasal stress goes to the syntactic non-head (Nonhead Stress). According to Cinque (1993), compound and phrasal stress goes to the branch that has a deeper sub-branch. The proposals are similar in three ways. First, both consider compound and phrasal stress to be a single rule. Second, since the syntactic head is an X^0 and a syntactic non-head is an XP, the latter always has a deeper sub-tree. Thus, the two proposals assign stress to the same positions. Third, following Abney (1987), Pollock (1989), and others, both proposals assume that syntactic heads are usually functional elements (particles, affixes, and inflections). In what follows I ignore the difference between Duanmu (1990a, 2000) and Cinque (1993) and call both the Nonhead Stress approach.[5] In (37), I compare the Chomsky-Halle model with the Nonhead Stress model, where the syntactic head is underlined and relative stress among stressed elements is ignored.

[5] According to the theory of Prosodic Hierarchy (Selkirk 1981), phonology does not refer to syntax directly, but only indirectly. According to Hayes (1989: 205), direct reference includes the use of syntactic categories, such as N, V, NP, VP, etc., whereas indirect reference includes the use of syntactic levels, such as X and XP. In this regard, Nonhead Stress does not make direct reference to syntax.

(37)	Syntax	Nonhead Stress	Chomsky and Halle
	[V N]	buy CARS	(same)
	[P N]	in SCHOOL	(same)
	[N N]	WRIST-watch	(same)
	[N [N N]]	GOLD WRIST-watch	(same)
	[[N N] N]	WRIST-watch store	WRIST-watch STORE
	[N [V N]]	COWS eat GRASS	(same)
	[N [A [V N]]	COWS OFTEN eat GRASS	(same)
	[D N]	the CAR	(same)
	[N's N]	JOHN 's FRIEND	John's FRIEND
	[A (F) N]	RED (F) CARS	red CARS

Some remarks are in order. First, in some structures, such as [N [N N]], [N [V N]], and [N [A [V N]], the Chomsky-Halle model assigns stress to more than one word. This is the result of a mechanism known as the 'Stress Subordination Convention' (Chomsky and Halle 1968: 16-17). A similar mechanism is used in Halle and Vergnaud (1987), which is called the 'Stress Equalization Convention'. Second, in the first seven structures, the syntactic heads are as traditionally understood. In [D N], the syntactic head is D, following Abney (1987). In the possessive structure [N's N], the syntactic head is the possessive {*s*}. The noun phrase [A N] is [A (F) N], where the syntactic head is the inflectional element {F}. This follows from the idea that noun phrases are also headed by a functional element (or elements), not always present in English (consider, however, {*ly*} in *lovely day*) but required in some other languages (Ritter 1991, Cinque 1993). For example, the Chinese counter-part to an English [A N] phrase is [A *de* N], where *de* is a functional element (Dai 1992, Duanmu 1998).

The two analyses predict similar stress assignments in most cases. Where they differ, the Nonhead Stress analysis seems to be more accurate. For example, in [N's N] and [A N], both words have stress in the Nonhead Stress analysis, in agreement with the traditional judgment (Kenyon and Knott 1944, Jones 1950). Similarly, as Hayes (1995: 373-382) argues, in structures like [[N N] N] and [[[N N] N] N], no stress should be

assigned after the main stress (here the first N). Finally, the present analysis agrees with the observation that verbs are less likely to be stressed than nouns (Ladd 1980: 90-92, Hayes 1995: 376). This is because verbs often occur as syntactic heads.[6, 7]

In (38) I summarize the stress rules for Chinese (which are basically the same as those for English).

(38) Word stress:

In a disyllabic word, the first syllable has stress

Compound and phrasal stress:

The syntactic non-head has stress (a combination of the Nuclear Stress Rule and the Compound Stress Rule of Chomsky and Halle 1968).[8]

---

[6] Although V is the head of [V N], the head of a sentence is not V but probably an inflectional element (or elements) I, which could be empty. Thus, [N [V N]], such as *cows eat grass*, should be [N [I [V N]], where I and V are both heads. Similarly, [N V], such as *horses run*, should be [N [I V]], where V is now a syntactic nonhead in [I V], and so V has stress. In other words, verbs can get stress, but less often than nouns.

[7] I discussed in section 2 that both [A N] and [N N] are coded as NN, because they are both compound in modern Chinese. A reviewer points out that, if [A N] can be a phrase in English, it could be a phrase in classic Chinese, too. If so, a distinction should perhaps be made between [A N] and [N N]. However, even if classic Chinese is closer to English than to modern Chinese, there is little consequence for the present analysis. The stress pattern of NN is stressed-unstressed, which can fill SW. If [A N] is a phrase, its stress pattern is stressed-stressed, which can also fill SW.

[8] The current definition of Nonhead Stress differs from cyclic stress assignment. For example, consider [V [N N]], such as *make pan-cakes*, analyzed in (i).

(i)	Current definition	Cyclic assignment	
		(x)	Second cycle
	x	(x)	First cycle
	[make [pan-cakes]]	[make [pan-cakes]]	

In the current definition, *pan* gets stress on the first cycle. On the second cycle, the object should have stress, which is already true, and so no new stress is assigned. Under the cyclic assignment, *pan* gets stress on the first cycle. On the second cycle, another stress is assigned to the object, which again falls on *pan*. The difference grows quickly when we analyze longer expressions, such as *sixty thousand herds of cattle*. In general, the current definition produces a flatter metrical structure, in agreement with Gussenhoven (1991). On the other hand, the cyclic assignment agrees with the Faithfulness Condition of Hayes (1995: 380), which requires a one-to-one relation between a pair of parentheses and a stress mark. However, Hayes seems to assume that every pair of syntactic brackets projects a metrical constituent (in addition to other constituents generated by the Continuous Column Constraint), a hypothesis that needs independent support. In any case, as I will show, higher stress levels are not relevant for the metrical analysis of Chinese

There are two additional arguments for the above stress rules. First, they can explain the word order and word length variation in Chinese (Duanmu 2000). Second, in some Chinese dialects, such as Shanghai, compounds and phrases have different tonal domains. The difference can again be explained if Chinese has the stress rules given above (Duanmu 1999a).

Next consider the remaining metrical requirements, shown in (39) and (40), where the stress maximum and its use are similar to those of the Halle-Keyser model.

(39) Stress maximum (preliminary):

A syllable is a stress maximum if it is stressed and is between two unstressed syllables.

(40) Restriction on the stress maximum:

A stress maximum must occur in S.

It should be noted that the stress maximum depends not only on word stress, but also compound and phrasal stress, to be illustrated below.

Next, consider the analysis of the final three syllables, because they seem to pose a particular problem. In most cases, they form a syntactic unit, either [1 2] or [2 1]. The statistics are shown in (41) and (42).

(41) Final three syllables in 5-syllable lines

Structure	Frequency	%
[1 2]	456	58%
[2 1]	278	35%
Others	54	7%
Total	788	100%

---

verse.

(42) Final three syllables in 7-syllable lines

Structure	Frequency	%
[1 2]	388	58%
[2 1]	257	38%
Others	27	4%
Total	672	100%

If the template for the final three syllables is SWS, it is easy to see why it can be filled by [2 1], but it is hard to see why it can be filled by [1 2], which is the most common pattern of all. For illustration, consider the two most frequent final trisyllabic units in five-syllable lines, [V NN], which occurs 245 times, and [NN N], which occurs 134 times. They are analyzed in (43), where x indicates stress.

(43) Two most frequent final trisyllabic units in five-syllable lines

```
 x                    x
[NN N]             [V NN]
 SW S              *S WS
```

In [NN N], the compound NN has trochaic stress. Thus, the second syllable is not a stress maximum, and the unit can fill SWS. In [V NN], the compound NN has trochaic stress. In addition, V does not get phrasal stress. Thus, the second syllable is a stress maximum and it should not fill W.

As a solution, I suggest the following analysis. The final syllable of a line is followed by an empty beat, which maps to the empty beat of the template. Phonetically, the empty beat represents a real unit of time; it is realized either as a pause or as the lengthening of the final syllable (Klatt 1976). Given the empty beat, the final trisyllabic units are analyzed in (44).

(44) Analysis of final trisyllabic units

[2 1]	[1 2]
x x	x x
[NN NØ]	[V [N [NØ]]]
SW SØ	S W SØ

In both structures, the final syllable and the empty beat form a binary unit. Assuming that the empty beat is weak, [NØ] is similar to a disyllabic word or compound and has trochaic stress. Now whether the second syllable has stress or not, it can fill W, because it is followed by a stressed syllable and is thus not a stress maximum. Finally, there is no restriction for the first syllable to fill S. In summary, because of the empty beat, the final syllable is always stressed, and so the penultimate syllable is never a stress maximum, and so the final three syllables can always fill SWS.[9]

Next, consider lines that do not start with a disyllabic word or compound. For illustration, consider the third most common tree in five-syllable lines, whose members are shown in (45), where X is a word whose category is hard to determine.

(45) The third most common tree in five-syllable lines [1 [1 [2 1]]]

Structure	Frequency
[A [V […	40
[N [V […	22
[V [V […	13
[B [V […	4
[N [A […	2
[A [P […	1

[9] Fabb (personal communications) suggests that, if I stipulate that the final syllable has stress, I can account for final trisyllabic units without the empty beat. I prefer the present analysis for several reasons. First, the empty beat seems to be a phonological fact, even though it is not transcribed orthographically. Second, without the empty beat, the idea that the final syllable has stress remains an independent stipulation. Third, without the empty beat, the final foot is monosyllabic, which violates a well-known constraint Foot Binarity.

[M [V […	1
[V [N […	1
[X [V […	1
Total	85

Since the last three syllables are [2 1] and can fill SWSØ, our interest is on the first two syllables. In particular, because the second syllable corresponds to W, it should not be a stress maximum. In most cases, the second syllable is a verb (V) or a preposition (P), followed by its object. Since V and P are syntactic heads and receive no phrasal stress, neither is a stress maximum. The remaining two cases are exemplified below.

(46) [N [A [AA V]]] (2 lines)

(47) [V [N [[V N] V]]] (1 line)

yan shi cai yao qu 249-2
say master gather herb go
'(He) said master went herb-gathering'

In (46), the second syllable precedes a disyllabic compound. Since the compound forms a trochee, the second syllable is not a stress maximum, and so it can occur in W. The line in (47) seems to have a problem. According to Nonhead Stress, the second syllable *shi* 'master' is a syntactic nonhead and so has stress, whereas the Vs around it have no stress. If so, the second syllable *shi* 'master' is a stress maximum and the line is unmetrical. On the other hand, if [V N] forms a compound (and hence a trochee), which happens quite often for disyllabic words in Chinese, then *shi* is not a stress maximum and the line is metrical.

Since the stress maximum is the only serious restriction, many structures are possible. However, there are also many structures that do not occur. Indeed, the stress maximum can explain some subtle distinctions. For example, in five-syllable lines, the second most common structure within [2 [1 2]] is [[V N][V NN]], which occurs 34 times,

yet a very similar structure [[V NN][V N]] is completely absent. Given the fact that Chinese sentences do not require an overt subject, both sentences should be equally good syntactically, as exemplified in (48) and (49).

(48) [[V N][V NN]] 34 examples
deng zhou wang qiu-yue 104-3
embark boat watch autumn-moon
'(I) embark the boat and watch the autumn moon.'

(49) [[V NN][V N]] (No example)
wang qiu-yue si jia
watch autumn-moon miss home
'(I) watch the autumn moon and miss home.'

The difference between the two must lie in metrics. Consider the analysis in (50).

(50)

```
      x      x   x                  x       x
[[V N][V [N [NØ]]]]        [[V NN][V [NØ]]]
  S W   S   W   SØ             S WS   W   SØ
```

In [V N] and [V NN], V has no stress but N or the first N of NN does. In both structures, the second syllable seems to be a stress maximum (between two unstressed syllables), and so it should not occur in W. Why then is [[V N][V NN]] still good? I suggest that the answer lies in syntax. As the underline shows, in [[V NN][V N]], the stress maximum is defined within the same VP. In contrast, in [[V N][V NN]], the stress maximum is defined over two VPs. To capture the difference, I revise the stress maximum in (51).

(51) Stress maximum (final version):
A syllable is a stress maximum if it is stressed and is between two unstressed

syllables within the same cyclic domain.[10]

(52) Cyclic domain:
A cyclic domain includes a branching terminal node and all non-branching terminal nodes that c-command it.[11]

Some examples of cyclic domains and stress maxima are shown in (53), where a vertical line indicates a boundary between cyclic domains.

(53)

```
      x       x  x
[[V N]|[V [N [NØ]]]]      No stress maximum
  S W   S   W  SØ         Example found

     x        x
[[V NN]|[V [NØ]]]         Stress maxima
  S WS   W  SØ            Stress maximum in W;  no example

       x
[V [V [NN | NØ]]          Stress maximum
  S W  SW   SØ            Stress maximum in S;  example found
```

[10] The term 'unstressed' means without word or phrasal stress. Since most Chinese syllables are heavy (Duanmu 2000), most Chinese syllables form a moraic trochee and have some stress.

[11] The cyclic domain defined here is reminiscent of an idea proposed by Liu (1980) and Kaisse (1985), according to which the Mandarin Tone 3 Sandhi rule does not apply across two sub-trees in a sentence. The cyclic domain also has some resemblance to the Clitic Group of Hayes (1989: 208), in particular to the idea of 'category membership'. However, for Hayes each lexical word heads a Clitic Group, whereas in the present analysis a cyclic domain can include several lexical words.

```
    x      x  x
[V [NN | [N [NØ]]]]        First N is stress maximum
  S WS     W  SØ          Stress maximum in W;  no example
```

The cyclic domain condition on the stress maximum is reminiscent of the Bounding Theory of Hayes (1989: 231), according to which the stress maximum (or peak) is defined within a syntactic constituent. However, for Hayes a syllable is a stress peak if it has more stress than 'at least one of its neighbors'. If so, the second syllable in [[V N][V NN]] is a stress peak and the line should be bad. The presence of such lines shows that the present definition is more appropriate.

To conclude this section, I summarize the present analysis in (54) and (55).

(54) Stress rules

a. Word stress:

In a disyllabic word, the first syllable has stress

b. Compound and phrasal stress:

The syntactic non-head has stress (a combination of the Nuclear Stress Rule and the Compound Stress Rule of Chomsky and Halle 1968).

(55) Metrical requirements

a. Templates (where Ø is an empty beat)

Five-syllable lines: SWSWSØ

Seven-syllable lines: SWSWSWSØ

b. A syllable is a stress maximum if it is stressed and is between two unstressed syllables within the same cyclic domain.

c. A stress maximum must occur in S.

4.2. Illustration of the analysis

In this section I show how the current proposal accounts for lines in the corpus. I will refer to the tree patterns by the frequency numbering given in section 2.

4.2.1. Five-syllable lines

In five-syllable lines, there are 10 tree patterns. First, consider the three in (56), where a vertical bar shows the boundary between cyclic domains.

(56)	Tree	Frequency	%
1	[2 | [1 2]]	424	53.81%
2	[2 | [2 1]]	193	24.49%
6	[[2 | 2] 1]	12	1.52%

In all cases, there is a cyclic domain boundary after the second syllable. Therefore, whether the first two syllables are a compound or a phrase (such as [V N]), the second syllable cannot be a stress maximum, and the first two syllables can fill SW. In addition, since the last syllable has stress (forming a trochee with an empty beat), the second final syllable cannot be a stress maximum, so the last three syllables can fill SWS.

Next consider [1 [1 [2 1]]], which was discussed before. The final three syllables can fill SWSØ. In addition, except for one case (line 249-2), the second syllable is not a stress maximum, and so the first two syllables can fill SW. Next consider the three patterns in (57).

(57)	Tree	Frequency	%
4	[1 [1 [1 2]]]	32	4.06%
5	[1 [[1 2] 1]]	31	3.93%
7	[[1 [1 2]] 1]	5	0.63%

Here again, if the second syllable is not a stress maximum, the lines are good. In most cases, the second syllable is a syntactic head (V or P), which does not get phrasal stress and so is not a stress maximum. There are five examples where the second syllable is not V or P, listed in (58).

(58)

Tree	Cases
[A [A [B [V N]]]]	106-8, 109-8
[A [A [V NN]]]	100-8, 154-8
[N [A [V [A V]]]	104-5

According to our stress rules, a stress maximum should be a syntactic nonhead that is between two syntactic heads. Here the second syllable is not between syntactic heads, so it is not a stress maximum (because one or both of its neighbors have stress). Finally, consider the remaining patterns, which cover six cases, detailed in (59).

(59)

	Tree	Structure	Cases
8	[[2 1] 2]	[[NN V][A V]]	139-5, 139-6, 166-5, 166-6
9	[[[2 1] 1] 1]	[[[NN N] N] V]	125-7
10	[[1 2] 2]	[[A [V N]][B V]]	127-7

In none of the cases is the second syllable a stress maximum, and therefore all the lines are metrical.

In summary, of the 788 five-syllable lines, all except one (line 249-2) are metrical.

4.2.2. Seven-syllable lines

In seven-syllable lines there are 33 tree patterns. Since the final three syllables can always fill SWS, our focus is on the first four syllables. First, consider the trees in (60),

where cyclic domain boundaries are shown with a vertical line.

(60)

	Tree	Frequency	%
1	[[2 | 2] | [1 2]]	215	31.99%
2	[[2 | 2] | [2 1]]	119	17.71%
4	[2 | [2 | [1 2]]]	47	6.99%
13	[2 | [2 | [2 1]]]	13	1.93%
15	[2 | 2 | [2 1]]	4	0.60%
19	[[[2 | 2] | 2] 1]	2	0.30%
26	[2 | [[2 | 2] 1]]	2	0.30%
29	[[2 | [2 | 2]] 1]	1	0.15%

In the first four syllables, the cyclic domains are all disyllabic, so there is no stress maximum. Therefore all the tees are good. Next consider the trees in (61)

(61)

	Tree	Frequency	%
5	[[1 [1 2]] | [1 2]]	43	6.40%
6	[1 [1 [2 | [1 2]]]]	27	4.02%
7	[1 [1 [2 | [2 1]]]]	20	2.98%
8	[[1 [1 2]] | [2 1]]	19	2.83%
10	[1 [[1 2] | [1 2]]]	18	2.68%
16	[1 [1 [[2 | 2] 1]]]	3	0.45%
20	[1 [[1 2] | [2 1]]]	2	0.30%
32	[1 [[1 [2 | 2]] 1]]	1	0.15%

In all cases, the first four syllables are in the same cyclic domain. What we need to show then is that neither the second nor the fourth syllable is a stress maximum. The fourth syllable cannot be a stress maximum, because it is the last syllable in the domain. The second syllable is not a stress maximum if it is V or P. This is true in most cases, except

five lines, shown in (62).

(62)		
	[[N [N NN]][V NN]]	174-1
	[INT [[N NN][N VV]]]	176-1
	[A [A [[P N][V NN]]]]	206-6
	[[V [N [V N]]][V NN]]	205-3, 205-4

In the first three cases, the second syllable is not a stress maximum, because it is not between syntactic heads. In the fourth case, the second syllable turns out to be a pronoun in both lines. Assuming that pronouns are not assigned phrasal stress, they are not stress maxima either. So all lines are metrical.

Next consider lines with two medial monosyllables, shown in (63).

(63)	Tree	Frequency	%
3	[2 | [1 [1 [2 1]]]]	63	9.38%
11	[2 | [1 [1 [1 2]]]]	15	2.23%
14	[2 | [1 [[1 2] 1]]]	9	1.34%

The trees will be good if the fourth syllable is not a stress maximum. In other words, the fourth syllable should be V or P, or it should not be surrounded by V or P. This is true for all cases. Next consider the examples in (64).

(64)	Tree	Frequency	%
9	[[[2 1] 1] | [1 2]]	18	2.68%
12	[[[2 1] 1] | [2 1]]	13	1.93%
30	[[2 1] | [1 [2 1]]]	1	0.15%
31	[[2 1] | [1 [1 2]]]	1	0.15%

The trees will be good if first two syllables form a compound, so that the second syllable is not a stress maximum. This is true for all cases except one, where the first four

syllables are [[[A V] N] V]. Here the second syllable is a syntactic head and so has no phrasal stress, so it is not a stress maximum either. Thus, all the lines are good. Next, consider the cases in (65).

(65)	Tree	Frequency	%
22	[1 [1 [1 [[1 2] 1]]]]	2	0.30%
23	[1 [1 [1 [1 [1 2]]]]]	2	0.30%
24	[1 [1 [1 [1 [2 1]]]]]	2	0.30%

These will be good if the second and fourth syllables are not stress maxima, which is the case for all of them. Next, consider the trees in (66).

(66)	Tree	Structure	Cases
18	[[[2 | 2] 1] | 2]	[[[NN NN] V][N V]]	2
21	[1 [1 [[[2 1] 1] 1]]]	[A [V [[[NN N] N] N]]]	1
		[A [B [[[NN N] N] V]]]	1
25	[2 | [[[2 1] 1] 1]]	[AA [[[NN N] N] N]]	1
		[NN [[[NN N] N] N]]	1
27	[[[1 2] 1] | [2 1]]	[[[A NN] N][NN V]]	1
33	[2 | [1 [2 | 2]]]	[NN [B [[P N][V N]]]]	1

In none of the cases is the second or fourth syllable a stress maximum. Finally, consider the remaining two trees, which cover three cases, shown in (67).

(67)	Tree	Structure	Cases
17	[[[1 2] 1] | [1 2]]	[[[P NN] N][N VV]]	311-1
		[[[P NN] N][N [V N]]]	311-2
28	[[1 2] | [1 [1 2]]]	[[P NN][V [A NN]]]	222-8

In these three cases, the second syllable is a stress maximum, and so we predict the lines to be unmetrical.

In summary, of the 672 seven-syllable lines, all except three are metrical.

4.3. Metrical tension

Halle and Keyser (1971) argue that a metrical analysis not only can predict which lines are metrical, but also can determine the tension (or complexity) in a metrical line. In addition, it is predicted that there is a correlation between the tension of a line and its frequency of occurrence. Halle and Keyser (1971: 157) state it as follows.

(68) The more complex the line … the less frequent it occurs.

Youmans (1989) argues that there is probably no clear distinction between metrical and unmetrical lines, although lines with less tension are expected to be more frequent. Golston (1998) shares the same view. In addition, he demonstrates with the Middle English verse Cleanness (1812 lines) that there is a quantitative correlation between the tension of a half-line (measured in terms of constraint violations) and its frequency of occurrence.

Given the present corpus, we are able to examine the correlation between the tension of a line and its frequency of occurrence in Chinese verse. In Halle and Keyser (1971), metrical tension is measured in two ways. First, every missing or extra syllable invokes tension. Second, every syllable with inappropriate stress invokes tension. Since our corpus does not contain missing or extra syllables, we focus on syllable stress. Let us assume that, as in English, there is no metrical tension if S is filled with a stressed syllable and W is filled with an unstressed one, and there is metrical tension if S is filled with an unstressed syllable or if W is filled with a stressed one. Now, consider the three most common structures in five-syllable patterns, shown in (69), where x indicates word or phrasal stress, and positions that invoke tension are underlined.

(69) Tension in the three most frequent five-syllable lines

	Tension	Frequency
SW S W SØ x x x [NN [V [N NØ]]]	2	145
SW SW SØ x x x [NN [NN VØ]]	0	41
SW SW SØ x x x [NN [NN NØ]]	0	38

The most frequent structure of all five-syllable lines is [NN [V NN]], which occurs 145 times. However, it has more tension that the second and third most frequent structures, which have no tension. It is also interesting to consider the structure in (70).

(70)

	Tension	Frequency
SW SW SØ x x x [NN [VV NØ]]	0	0

This structure does not occur at all, even though it has no tension and looks very similar to the most frequent structure (both being [S [V O]] sentences). Thus, there seems to be no statistical correlation between tension and frequency in Chinese verse.

On the other hand, there is a syntactic explanation of tree frequency. In particular, [NN [V NN]] is the most common pattern probably because (a) [S [V O]] is the most common sentence and (b) there are more disyllabic nouns (including compound nouns) than disyllabic verbs. In addition, there is a metrical reason why nouns tend to be disyllabic and verbs tend to be monosyllabic: verbs often occur as a syntactic head, which does not get phrasal stress. In contrast, nouns tend to occur as a syntactic nonhead, which

gets phrasal stress, which then should be disyllabic to satisfy foot binarity (Duanmu 1999b).

4.4. Summary

I have completed the analysis of the metricality of all five-syllable and seven-syllable lines. I summarize the results in (71).

(71)

Metrical	1456
Exceptions	4
Total	1460

For reference, the four unmetrical lines are listed in (72).

(72) [V [N [[V N] V]]]

yan shi cai yao qu 249-2

say master gather herb go

'(He) said master went herb-gathering'

[[P NN][V [A NN]]]

wei ta-ren zuo jia yi-shang 222-8

for other-people make wedding clothes

'(I) make wedding clothes for other people'

[[[P NN] N][N VV]]

jin han-shi yu cao qiqi 311-1

near Han-Shi rain grass thick

'Rain near the Han-Shi season; the grass is thick'

[[[P NN] N][N [V N]]]
zhao mai-miao feng liu qing ti 311-2
on wheat-seedling wind willow shine bank
'Wind blows over the wheat field; willow trees shine on the bank'

Since we are dealing with poetry written hundreds of years ago, it is perhaps hard to find proof for the predictions. However, some evidence is available. First, the exceptional lines are rare, in agreement with the general view that Chinese regulated verse is technically strict. Second, what we predict to be exceptional lines have often prompted special comments in traditional literature, although not in metrical terms. For example, Qiu (1976) found line 249-2 to be 'mysterious'. In lines 311-1 and 311-2, the initial P is said to be 'special'. The line 222-8 belongs to what is traditionally called '3-4 sentences', in which the first part has three syllables and the second has four, a pattern to be avoided (Li 1935: 16-17, Wang 1958, Tai 1974, volume III, p.1610). The present analysis has made it explicit why those lines appear to be deviant.

I have also discussed whether there is a correlation between metrical tension and frequency of use. The statistical data provide no evidence for the correlation.

A reviewer observes that most lines start with a disyllabic word or compound. This is indeed true. In five-syllable lines, 47% of the lines start with a disyllabic word or compound, and in seven-syllable lines, 72% of the lines do (see Appendices). In the present analysis, this is expected: the first two syllables correspond to SW in the template, and a disyllabic compound with trochaic stress is a perfect fit for SW.

Hayes (1989) argues that the Prosodic Hierarchy plays a role in English verse. Similarly, Golston (1998) argues that that the Prosodic Hierarchy plays a role Middle English verse. One might wonder, therefore, whether the Prosodic Hierarchy also plays a role in Chinese verse. I did not discuss the Prosodic Hierarchy beyond the foot. There are two reasons. First, given the notion of the stress maximum, most lines can already be accounted for. I do not see an alternative analysis without stress assignment (or the notion of the stress maximum), with or without the Prosodic Hierarchy. Second, since most Chinese words are monosyllabic, it is not obvious how to derive prosodic categories

properly, such as the Prosodic Word or the Clitic Group (see note 11).

5. Further issues

In this section I discuss some further issues. They offer additional support for the present analysis.

5.1. Tonal requirement in regulated verse

Chinese verse is divided into two categories, Old-Style Verse and Recent-Style Verse. Old-Style Verse originated from folk verse and has no special tonal requirements. Recent-Style Verse, or regulated verse, is subject to special tonal requirements, according to which tones are divided into two categories, *Ping* (literally 'even') and *Zhe* (literally 'oblique'). In addition, the two categories should alternate in specific ways.

The phonetic difference between the two tonal categories remains somewhat unclear. Zhou (1948) suggests that the difference lies in length, in that Ping tones are long and Zhe tones are short. On the other hand, Mair and Mei (1991) argue that the tonal categories were invented to imitate heavy vs. light syllables in Sanskrit prosody.

Following Mair and Mei (1991), I denote the tonal categories as A and B. According to Chen, seven-syllable lines have four tonal alternations, shown in (73), along with Chen's foot analysis. There are also requirements on the sequencing of the lines, which need not concern us here.

(73)

Tonal Patterns	Chen's Feet
AABBAAB	(AA)(BB)(AA)(B)
BBAABBA	(BB)(AA)(BB)(A)
AABBBAA	(AA)(BB)(B)(AA)
BBAAABB	(BB)(AA)(A)(BB)

Chen argues that tonal alternation is a reflection of foot structure: syllables in a foot have

the same tones and tones of adjacent feet mostly differ. Both Chen and the present analysis agree that the first four syllables form two feet. The difference lies in (a) the analysis of the last three syllables, and (b) whether stress is on odd-numbered or even-numbered syllables.

It is easy to see that Chen's analysis contains monosyllabic feet and stress clash. In particular, if the Chinese foot is iambic, as Chen assumes, either the final syllable is W, or there is a stress clash, as shown with a seven-syllable line in (74).

(74) Problems with the iambic analysis

a. Final W: (WS)(WS)(WS)W
b. Stress clash: (WS)(WS)(WS)(S)
c. Stress clash: (WS)(WS)(S)(WS)

In the first case, the final syllable is unfooted and W, which contradicts with the fact that it is the rhyming syllable and should be stressed. In the other two cases, there is a monosyllabic foot. In addition, whether the monosyllabic foot is final or medial, it causes stress clash. Therefore, none of the options is an ideal rhythmic template.

Next consider tone. Of interest is a rule known as the '1-3-5' rule, according to which the first, third, and fifth syllables can take any tone, regardless of the tonal requirement for a line. In other words, even-numbered positions have tonal restrictions, but odd-numbered positions do not.[12]

Yip (1980) suggests that it is stressed syllables that have tonal restrictions, and so stress is on even-numbered syllables. Chen (1980: 15) shares the same view. However, there is a problem with the assumption, noted by Duanmu (1990b: 186). In (73) the fifth syllable of the last two lines should have stress, since it is the only syllable of the foot, yet this syllable is free to take any tone. Therefore, it is better to assume that it is stressed syllables that have more choices for tones (except the final syllable, which is the position

[12] As Wang 1958 points out, the 1-3-5 rule is a somewhat simplified statement. For example, a line is bad if there is just one 'even' tone. The details of the rule need not concern us here.

for rhyming).

Next consider the analysis of the last three syllables. We have seen that their syntax is flexible, and Chen often grouped them into one foot. Under this consideration, along with the 1-3-5 rule and the fact that the last syllable is the riming position, Chen's analysis is compared with the present one in (75), where O is a free position, R a riming position,[13] and Ø an empty beat.

(75)

Tonal Patterns	Chen's Analysis	Present Analysis
OAOBOAR	(OA)(OB)(OAR)	(OA)(OB)(OA)(RØ)
OBOAOBR	(OB)(OA)(OBR)	(OB)(OA)(OB)(RØ)
OAOBOAR	(OA)(OB)(OAR)	(OA)(OB)(OA)(RØ)
OBOAOBR	(OB)(OA)(OBR)	(OB)(OA)(OB)(RØ)

As far as required tonal positions are concerned, Chen's analysis is identical to the present one.

According to Mair and Mei (1991), Shen Yue and his followers invented tonal requirements for Chinese regulated verse between 488 and 550, in imitation of Sanskrit prosody. Mair and Mei (1991: 454-455) also conclude that the invention shows that meter and prosody can be transmitted from one language to another. However, if the present analysis is correct, there is perhaps something more fundamental, namely, the core of prosody, such as the rule for compound and phrasal stress and the role of the stress maximum, is probably inherent in all languages. It is true that Shen Yue and his followers considered their invention to be superior to Old-Style Verse, which originated from folk verse. In addition, regulated verse gained prominence in the Tang Dynasty (618-907), especially after it was adopted in official exams. Still, it is possible that it is folk verse that is a better reflection of native intuition. It is worth noting that neither Old-Style Verse nor modern folk verse has tonal requirements. In any case, apart from tonal requirements, regulated verse is still subject to metrical requirements, as I have shown.

---

[13] Riming is required for even-numbered lines but optional for the first line.

5.2. Verse recitation in Southern Min

In support of iambic footing, Chen (1980) cites evidence from Southern Min, based on modern recitation of a traditional ballad and two Tang poems, each read by one person, with a total of 344 seven-syllable lines (Boyce 1980, Dong et al 1967).

For the sake of argument, let us assume that Southern Min and classic Chinese have the same stress rules. Given this, let us consider how the Southern Min data can be reconciled with the present analysis.

In Southern Min every full syllable has two tonal patterns. One is used in isolation and certain other positions, usually phrase final (call it A). One is used in nonfinal positions (call it B). In addition, an A tone is usually longer than a B tone. There are eight A tones and eight B tones, some of which overlap (e.g. the A tone of one syllable may be the same as the B tone of another). Boyce (1980:3) reports that the most common pattern is BABABBA, which is found in about 50% of the lines. Chen (1980: 17) points out that if B represents a weak position and A represents a strong position, then a seven-syllable line has three iambs (WS)(WS)(WWS).

However, while there is no question that A is strong, it is not obvious that B is weak. First, both A and B tones fall on full syllables, all of which have some stress, in contrast to weak syllables, which have no stress. Second, the preference for using A in certain positions may be due to the fact that these are usually phrase final positions. The two views make different predictions, shown in (76).

(76) Two predictions

a. A tone is phrase final:

   B can occur in positions 2 and 4

   More B's in position 2 than 4

b. A tone is foot final (for iambic feet):

   B should not occur in positions 2 or 4

   Same number of A's in position 2 and 4

According to (76a), B (nonfinal tone) can occur in positions 2 and 4, if they are not phrase final. In addition, since there is a greater chance to find a major syntactic boundary after the fourth syllable than after the second (Chen 1979), we expect more B's in position 2 (non-phrase final) than in position 4 (phrase final). According to (76b), B should not occur in positions 2 or 4, because these are stress positions. In addition, since both positions 2 and 4 are stress positions, A should be just as frequent in position 2 as in position 4. The facts support (76a): Chen (1980:19) notes that among the 344 lines Boyce cited, B indeed occurs 135 times in positions 2 and 4. In addition, B occurs three times as often in position 2 (101 cases) as in position 4 (34 cases). In other words, there are fewer A's in position 2 than in position 4. If Southern Min has iambs and if (BA) is an iamb, we should not see any B in positions 2 or 4, and we should not expect any difference in the frequency of A between positions 2 and 4.

5.3. Modern judgment

It is often assumed that we should not use modern intuition to judge the metricality of verse written in the past, because stress or metrical rules may have changed. Nevertheless, it is worth noting that the present analysis of Chinese regulated verse matches modern intuition quite well. For example, as Liu (1927) and Schlepp (1980a, b) point out, in tapped recitation, stronger taps fall on odd-numbered syllables, which are S positions in the present analysis. In addition, the lines that are predicted to be unmetrical indeed sound bad to the modern ear, whereas other lines sound good. This would be expected if (a) compound and phrasal stress is universal, as Cinque (1993) and Duanmu (2000) suggest, so that it does not change over time, and (b) the restriction on the stress maximum is also universal.

5.4. Lack of tension in Chinese verse

It is well known that English verse lines are often imperfect: most lines contain

some degree of tension. For example, Golston (1998) argues that each of the correspondent rules of Halle and Keyser (1971) is violated by Shakespeare's poems, and that in the Middle English verse Cleanness, no line is perfectly metrical.

There is a common feeling among Chinese speakers that English verse is indeed less metrical than Chinese verse, in the sense that the rhythm in English verse is less obvious. The present analysis also shows a fairly clear distinction between metrical and unmetrical lines in Chinese verse.

If the difference between English and Chinese is true, an explanation is needed. A possible answer lies in word length. In English, there are many polysyllabic words that inherently trigger tension. For example, *hierarchy* must fill SWS, where the final syllable is unstressed by fills S. In contrast, Chinese words are mostly monosyllabic, and there is little inherent tension. In addition, disyllabic compounds are easily created to fill SW units.

6. Conclusions

I have proposed, based on 1460 lines of Chinese regulated poems, that Chinese verse can be analyzed in the same way as English verse. The analysis is similar to the Halle-Keyser model and is summarized in (77) and (78).

(77) Stress rules

a. Word stress:

In a disyllabic word, the first syllable has stress

b. Compound and phrasal stress:

The syntactic non-head has stress (a combination of the Nuclear Stress Rule and the Compound Stress Rule of Chomsky and Halle 1968).

(78) Metrical requirements

a. Templates (where Ø is an empty beat)

Five-syllable lines: SWSWSØ

Seven-syllable lines: SWSWSWSØ

b. A syllable is a stress maximum if it is stressed and is between two unstressed syllables within the same cyclic domain.

c. A stress maximum must occur in S.

As in English, the restriction on the stress maximum plays a central role. On the other hand, since most Chinese words are monosyllabic, the definition of the stress maximum depends on compound and phrasal stress. Of the 1460 lines, four are ruled unmetrical, in agreement with the fact that they have triggered special remarks in traditional literature. A further result of interest is that, unlike what is often assumed, there is no statistical evidence for a correlation between the metrical tension of a line and its frequency of occurrence in Chinese verse.

I conclude that English and Chinese are fundamentally similar, both in terms of compound and phrasal stress assignment and in terms metrical requirements for verse.

7. References

Abney, Steven. 1987. The English noun phrase in its sentential aspect. Doctoral dissertation, MIT, Cambridge, Mass.

Boyce, Canol D. 1980. Min sandhi in verse recitation. Journal of Chinese Linguistics 8: 1-14.

Burling, Robbins. 1966. The metrics of children's verse: a cross-linguistic study. American Anthropologist 68: 1418-1441.

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

Chen, Matthew Y. 1979. Metrical structure: evidence from Chinese poetry. Linguistic

Inquiry 10: 371-420.

Chen, Matthew. 1980. The primacy of rhythm in verse: A linguistic perspective. Journal of Chinese Linguistics 8:15-41.

Chen, Matthew Y. 2000. Tone Sandhi: Patterns Across Chinese Dialects. Cambridge Studies in Linguistics, No. 92. Cambridge, UK: Cambridge University Press.

Chomsky, Noam, and Morris Halle. 1968. The sound pattern of English. N.Y.: Harper and Row.

Cinque, Guglielmo. 1993. A null theory of phrase and compound stress. Linguistic Inquiry 24.2:239-297.

Dai, John Xiangling. 1992. Chinese morphology and its interface with the syntax. Doctoral dissertation, Ohio State University, Columbus.

Dong, Tonghe, Ronglang Zhao, and Yaxiu Lan. 1967. Ji Taiwan de yizhong Minnanhua [A Southern Min dialect in Taiwan]. Institute of History and Philology Monographs, series A, number 24. Taipei: Academia Sinica.

Duanmu, San. 1990a. A formal study of syllable, tone, stress and domain in Chinese languages. Doctoral dissertation, MIT, Cambridge, Mass.

Duanmu, San. 1990b. Tonal patterns in Chinese regulated verses: Phonological rules or phonological constraints?" in Thomas Green, and S. Uziel. eds.), MIT working papers in linguistics 12: Students Conference in Linguistics 1990, 177-189.

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

Duanmu, San. 1999a. Metrical Structure and Tone: Evidence from Mandarin and Shanghai. Journal of East Asian Linguistics 8.1: 1-38.

Duanmu, San. 1999b. Stress and the development of disyllabic vocabulary in Chinese. Diachronica XVI: 1.1-35.

Duanmu, San. 2000. The Phonology of Standard Chinese. Oxford: Oxford University Press.

Fabb, Nigel. 2002. Language and literary structure: the linguistic analysis of form in

verse and narrative. Cambridge, New York: Cambridge University Press.

Gussenhoven, Carlos. 1991. The English rhythm rule as an accent deletion rule. Phonology 8:1-35.

Halle, Morris, and William Idsardi. 1995. General properties of stress and metrical structure. In The handbook of phonological theory, ed. John Goldsmith, 403-443. Cambridge, MA: Blackwell.

Halle, Morris, and Samuel Jay Keyser. 1971. English stress: its form, its growth, and its role in verse. New York: Harper and Row.

Halle, Morris, and Jean-Roger Vergnaud. 1987. An essay on stress. Cambridge, Mass.: MIT Press.

Hayes, Bruce. 1985. Iambic and trochaic rhythm in stress rules. In Proceedings of the Eleventh Meeting of the Berkeley Linguistics Society, ed. Mary Niepouj et al, 429-446. Berkeley Linguistics Society, California.

Hayes, Bruce. 1995. Metrical stress theory: Principles and case studies. Chicago: University of Chicago Press.

Hyman, Larry. ed. 1977. Studies in stress and accent. In Southern California Occasional Papers in Linguistics 4. Los Angeles: Department of Linguistics, University of Southern California.

Idsardi, William. 1992. The computation of prosody. Doctoral dissertation, MIT, Cambridge, Mass.

Jones, Daniel. 1950. The pronunciation of English. 3rd edition. Cambridge, England: Cambridge University Press.

Kaisse, Ellen. 1985. Connected speech. New York: Academic Press.

Kenyon, John Samuel, and Thomas Albert Knott. 1944. A pronouncing dictionary of American English. Springfield, MA: Merriam.

Kiparsky, Paul. 1975. Metrics, syntax, and meter. Language 51: 576-616.

Kiparsky, Paul. 1977. The rhythmic structure of English verse. Linguistic Inquiry 8: 189-247.

Klatt, Dennis H. 1976. Linguistic uses of segmental duration in English: Acoustic and

perceptual evidence. Journal of the Acoustical Society of America 59.5:1208-1221.

Ladd, Dwight Robert. 1980. The Structure of intonational meaning: evidence from English. Bloomington: Indiana University Press.

Li, Jingkang. 1935. Qiyan lüfa juyu [A sample of rules of heptasyllabic regulated verse]. Nanhai Middle School, Canton.

Liberman, Mark. 1975. The intonational system of English. Doctoral dissertation, MIT, Cambridge, Mass.

Lin, Hua. 1994. Against domain-final stress in Mandarin. Paper presented at the Sixth North American Conference on Chinese Linguistics, May 1994. Los Angeles: University of Southern California.

Liu, Dabai. 1927. Zhongguo jiu shipian zhong de shengdiao wenti [Questions concerning the tones of classical Chinese poetry]. Xiaoshuo Yuebao 17, Supplement. Reprinted in Qichao Liang, ed. 1971, Zhongguo wenxue yanjiu [Studies in Chinese literature], 149-208. Kyoto: Zhongwen Chubanshe.

Liu, Feng-hsi. 1980. Mandarin tone sandhi: a case of interaction between syntax and phonology. Paper presented at the LSA Annual Meeting.

Mair, Victor H. and Tsu-Lin Mei. 1991. The Sanskrit origins of recent style prosody. *Harvard Journal of Asiatic Studies* Vol. 51, No. 2. (Dec.): 375-470.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. Linguistic Inquiry 20:365-424.

Poser, William. 1990. Evidence for foot structure in Japanese. Language 66: 78-105.

Prince, Alan, and Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. Ms., Rutgers University and University of Colorado.

Qiu, Xieyou. 1976. Xin yi Tang shi san bai shou [New translation of 300 Tang poems]. Taipei: Samin Shuju.

Ritter, Elizabeth. 1991. Two functional categories in noun phrases: evidence from Modern Hebrew. In Perspectives on Phrase Structure: Heads and Licensing, ed. Susan D. Rothstein, 37-62. Syntax and Semantics 25. San Diego: Academic.

Schlepp, Wayne. 1980a. Tentative remarks on Chinese metrics. Journal of Chinese Linguistics 8: 59-84.

Schlepp, Wayne. 1980b. Some comments on Matthew Chen's system of metrics. Journal of Chinese Linguistics 8 :163-72.

Selkirk, Elisabeth. 1981. On prosodic structure and its relation to syntactic structure. In Nordic Prosody II, ed. Thorstein Fretheim, 111-140. Trondheim, Norway: Tapir.

Selkirk, Elisabeth, and Tong Shen. 1990. Prosodic domains in Shanghai Chinese. In The phonology-syntax connection, ed. Sharon Inkelas and Draga Zec, 313-337. CSLI, Stanford University, Stanford, Calif. Distributed by University of Chicago Press.

Tai, Jingnong. 1974. Baizhong shihua leibian [A classification and compilation of various comments on poetry]. 3 volumes. Taipei: Yiwen Yinshuguan.

Wang, Li. 1958. Hanyu Shilü Xue [Chinese Versification]. Shanghai: Xin Zhishi Chubanshe.

Yip, Moira. 1980. The metrical structure of regulated verse. Journal of Chinese Linguistics 8: 107-125.

Youmans, Gilbert. 1989. Introduction: rhythm and meter. In Rhythm and meter, ed. Paul Kiparsky and Gilbert Youmans, 1-14. San Diego: Academic Press.

Zhou, Fagao. 1948. Shuo ping ze [On even and oblique tones]. Bulletin of the Institute of History and Philology, Academia Sinica 13: 153-162. Reprinted in Fagao Zhou, Zhongguo yuyanxue lunwen ji [Papers on Chinese linguistics], 105-114. Hong Kong: Chongji Publishers, 1968.

Appendix 1: Five-syllable lines sorted by syntax

Style	Poems	Lines
5 Lu	80 (90-169)	640
5 Jue	37 (224-260)	148
Total	117	788

Sample	Syntax	#
91-1	[NN [V NN]]	145
93-1	[NN [NN V]]	41
95-5	[NN [NN N]]	38
91-6	[[V N][V NN]]	34
224-2	[A [V [NN N]]]	32
92-2	[AA [V NN]]	24
90-3	[N [V [NN N]]]	20
92-8	[NN [A [V N]]]	20
100-2	[NN [AA V]]	17
90-2	[AA [NN N]]	13
119-3	[[A V][V NN]]	11
108-1	[NN [A VV]]	11
97-3	[[N V][NN V]]	10
126-4	[V [V [NN N]]]	9
116-5	[[N V][V NN]]	8
113-8	[NN [N NN]]	8
90-8	[A [[P NN] V]]	7
101-3	[N [[P NN] V]]	7
99-8	[A [V [NN V]]]	7
98-7	[NN [A [A V]]]	7
101-1	[VV [NN N]]	7
94-7	[A [V [V NN]]]	6
110-5	[NN [V [N V]]]	6
111-5	[NN [[V N] V]]	6
114-3	[[NN NN] V]	5
159-1	[[N V][N [V N]]]	5
100-5	[[V N][A VV]]	5
123-4	[NN [N [A V]]]	5
104-8	[NN [V AA]]	5
146-6	[[N V][NN N]]	5
127-3	[[V N][NN V]]	5
139-5	[[NN V][A V]]	4
123-3	[[N V][A [A V]]]	4
93-5	[[N V][N [A V]]]	4
106-3	[[V N][N [V N]]]	4
108-2	[AA [N [A V]]]	4
130-2	[NN [V [A V]]]	4
94-8	[NN [V [V N]]]	4
102-7	[[V N][AA V]]	4
102-8	[AA [NN V]]	4
154-5	[NN [[P N] V]]	4
96-7	[NN [[V N] N]]	4
94-3	[[NN [V N]] N]	3
250-4	[B [V [V NN]]]	3
99-3	[N [[V NN] V]]	3
93-4	[V [[P NN] V]]	3
94-1	[A [V [N NN]]]	3
93-8	[W [V [V NN]]]	3
136-8	[[A V][A [V N]]]	3
100-6	[[V N][B [V N]]]	3
90-5	[[V N][V [N V]]]	3
150-5	[[V N][V [V N]]]	3
109-2	[AA [N NN]]	3
96-3	[NN [A [B V]]]	3
100-7	[NN [A [M V]]]	3
97-2	[[V N][NN N]]	3
143-4	[[A [V NN]] V]	2

152-7	[A [[V NN] V]]	2	148-5	[[V N][[A V] V]]	2
229-3	[B [V [NN N]]]	2	145-6	[[V N][[V N] V]]	2
228-2	[M [V [NN N]]]	2	111-7	[NN [[A V] N]]	2
106-8	[A [A [B [V N]]]]	2	226-1	[NN [[A V] V]]	2
100-8	[A [A [V NN]]]	2	136-1	[NN [AA N]]	2
105-4	[B [V [V NN]]]	2	97-7	[NN [WW V]]	2
100-1	[N [V [N NN]]]	2	125-7	[[[NN N] N] V]	1
123-5	[V [V [[N V] N]]]	2	246-4	[[M [V NN]] B]	1
169-4	[[N V][B [V N]]]	2	107-4	[[N [V NN]] V]	1
96-6	[[N V][N [B V]]]	2	118-7	[[V [V NN]] V]	1
169-1	[[V A][A [V N]]]	2	127-7	[[A [V N]][B V]]	1
129-8	[[V N][A [A V]]]	2	112-8	[[[N V][V N]] V]	1
114-8	[[V N][N [A V]]]	2	92-3	[[[P N] VV] N]	1
109-5	[[V N][N VV]]	2	133-1	[[NN VV] N]	1
124-5	[[V V][V NN]]	2	111-6	[[NN VV] V]	1
117-8	[AA [A [V N]]]	2	131-3	[V [[V NN] V]]	1
141-3	[AA [N [V A]]]	2	239-2	[X [[P NN] V]]	1
142-3	[NN [A NN]]	2	105-3	[A [V [A NN]]]	1
104-6	[NN [B [M V]]]	2	165-7	[A [V [A VV]]]	1
224-1	[NN [B [V N]]	2	156-1	[A [V [N [A V]]]]	1
136-7	[NN [Conj NN]]	2	162-2	[A [V [N [V A]]]]	1
163-7	[NN [N [B V]]]	2	260-3	[A [V [N [V N]]]]	1
133-3	[NN [N [V A]]]	2	107-8	[A [V [N WW]]]	1
110-2	[NN [N VV]]	2	142-8	[B [V [A [V N]]]]	1
141-5	[NN [V [B V]]]	2	234-4	[B [V [N [V N]]]]	1
152-4	[NN [V [Conj V]]]	2	104-5	[N [A [V [A V]]]]	1
132-7	[NN [V [V V]]]	2	160-8	[N [V [V [Conj V]]]]	1
123-8	[VV [V NN]]	2	253-2	[N [V [V NN]]]	1
134-5	[[A V][NN V]]	2	161-7	[W [V [A [A V]]]]	1

115-7	[A [P [NN V]]]	1	146-2	[[V A][A [A V]]]	1
108-4	[A [V [[B V] N]]]	1	109-7	[[V N][A [B V]]]	1
124-8	[A [V [[V N] N]]]	1	167-4	[[V N][A [V N]]]	1
251-2	[B [V [AA V]]]	1	169-5	[[V N][B [N V]]]	1
120-1	[B [V [NN N]]]	1	225-2	[[V N][Conj [A V]]]	1
103-7	[B [V [NN V]]]	1	122-8	[[V N][P NN]]	1
156-8	[N [A [AA V]]]	1	146-3	[[V N][V [A V]]]	1
113-5	[N [A [NN V]]]	1	151-7	[[V V][V [N V]]]	1
153-8	[N [V [[V N] N]]]	1	96-4	[[W N][A [VV]]]	1
119-7	[N [V [AA N]]]	1	105-7	[[W N][V NN]]	1
249-2	[V [N [[V N] V]]]	1	131-1	[AA [A [W V]]]	1
124-7	[V [V [[V N] N]]]	1	137-2	[AA [A VV]]	1
95-1	[V [V [NN V]]]	1	95-2	[AA [B [V N]]]	1
239-3	[X [V [NN N]]]	1	254-4	[AA [B VV]]	1
234-3	[X [V [NN V]]]	1	140-5	[AA [N [V N]]]	1
140-2	[[A V][A [A V]]]	1	169-8	[NN [A [N V]]]	1
143-7	[[A V][A NN]]	1	129-1	[NN [A [V V]]]	1
144-7	[[A V][A VV]]	1	166-2	[NN [A [W V]]]	1
232-1	[[A V][B [V N]]]	1	168-7	[NN [B [A V]]]	1
141-7	[[A V][N [V N]]]	1	104-2	[NN [B NN]]	1
166-8	[[A V][N VV]]	1	119-2	[NN [B VV]]	1
126-1	[[A V][V [N V]]]	1	240-4	[NN [M [B V]]]	1
156-2	[[A V][V [V N]]]	1	131-6	[NN [N [N V]]]	1
134-8	[[N V][A [W V]]]	1	163-8	[NN [N [V N]]]	1
258-1	[[N V][N VV]]	1	160-4	[NN [N [W N]]]	1
130-1	[[N V][V [N V]]]	1	160-3	[NN [N [W V]]]	1
103-3	[[P N][A [V N]]]	1	142-5	[NN [V [M V]]]	1
107-6	[[P N][V NN]]	1	226-4	[NN [V B V]]	1
137-8	[[P N][V WW]]	1	131-5	[NN [W [N V]]]	1

157-2	[VV [A [V N]]]	1
140-7	[WW [A VV]]	1
111-1	[[A V][[P N] V]]	1
145-5	[[A V][[V N] V]]	1
115-8	[[A V][NN N]]	1
108-3	[[P N][AA N]]	1
96-1	[AA [[A V] N]]	1
134-1	[AA [[V N] N]]	1
164-8	[AA [[V N] V]]	1
91-7	[AA [AA V]]	1
113-7	[AA [WW V]]	1
166-4	[NN [[A V] A]]	1
244-1	[NN [[N V] V]]	1
228-4	[NN [[V N] B]]	1
90-1	[NN [[W V] N]]	1
166-1	[NN [VV A]]	1
120-4	[NN [WW N]]	1
96-2	[VV [[P N] V]]	1
145-2	[VV [AA V]]	1
169-7	[VV [NN V]]	1

Appendix 2: Seven-syllable lines sorted by syntax

Style	Poems	Lines
7 Lu	54 (170-223)	432
7 Jue	60 (261-320)	240
Total	114	672

Sample	Syntax	#
172-3	[[NN NN][V NN]]	75
170-2	[NN [A [V [NN N]]]]	31
176-3	[[NN NN][NN N]]	18
180-1	[[NN NN][NN V]]	18
193-1	[[NN [V N]][V NN]]	11
170-7	[[NN NN][AA V]]	11
171-1	[[AA NN][V NN]]	10
213-3	[[NN [V N]][NN V]]	10
181-2	[[NN VV][NN V]]	10
183-5	[NN [AA [V NN]]]	10
171-7	[A [V [NN [NN N]]]]	9
192-3	[[A [V NN]][V NN]]	8
196-5	[[NN [V N]][N [A V]]]	8
186-5	[[NN NN][A [V N]]]	8
170-5	[[NN VV][NN N]]	8
213-2	[[[NN N] N][V NN]]	7
173-1	[A [V [NN [V NN]]]]	7
175-1	[[[NN N] N][NN V]]	6
186-6	[[NN [V N]][A [V N]]]	6
170-4	[[NN NN][A VV]]	6
172-1	[[NN [A V]][NN V]]	6
199-5	[[N [V NN]][V NN]]	5
188-7	[[A [V NN]][NN N]]	5
189-5	[[NN VV][V NN]]	5
186-3	[[AA NN][AA V]]	5
184-5	[A [[P NN][V NN]]]	5
212-5	[N [V [NN [N NN]]]]	5
179-1	[NN [A [[P NN] V]]]	5
214-7	[NN [NN [V NN]]]	5
178-8	[NN [VV [NN N]]]	5
273-3	[[[NN N] N][NN N]]	4
212-1	[[N [V NN]][N [V N]]]	4
191-5	[[NN NN][A NN]]	4
189-3	[[NN NN][N VV]]	4
183-2	[A [V [NN [AA V]]]]	4
185-1	[NN [A [V [V NN]]]]	4
220-7	[NN [B [V [NN N]]]]	4
208-5	[NN [V [V [NN N]]]]	4
278-4	[[[NN N] V][V NN]]	3
200-5	[[[NN N] V][NN V]]	3
177-5	[[N [V NN]][N [A V]]]	3
177-1	[[N [V NN]][NN V]]	3
217-6	[[V [V NN]][NN V]]	3
195-2	[[NN [A V]][V NN]]	3
276-1	[[NN [N V]][V NN]]	3
190-5	[[NN [V N]][A [B V]]]	3
171-2	[[NN NN][V [B V]]]	3
190-1	[[VV NN][NN N]]	3
179-7	[V [[V NN][V NN]]]	3
179-8	[B [V [NN [V NN]]]]	3
186-1	[[N V][N V][NN V]]	3
171-4	[[[NN N] N][N [A V]]]	2
279-1	[[[NN N] N][N [V N]]]	2
188-3	[[[NN NN] V][N V]]	2
222-4	[[A [V NN]][A NN]]	2
177-6	[[N [V NN]][N [B V]]]	2
205-3	[[V [N [V N]]][V NN]]	2

281-3	[[V [V NN]][V NN]]	2	293-4	[V [V [NN [NN N]]]]	2
188-2	[[A [V NN]][NN V]]	2	211-7	[NN [A [V [N NN]]]]	2
212-2	[[N [V NN]][NN N]]	2	211-4	[NN [M [V [V NN]]]]	2
290-1	[[[NN [V N]][N [B V]]]	2	216-5	[NN [V [V [N VV]]]]	2
176-2	[[[V N][V N]][V NN]]	2	215-5	[[A V][V [V [NN V]]]]	2
194-7	[[AA NN][A [A V]]]	2	310-3	[[V N][A [V [NN N]]]]	2
182-3	[[AA NN][A VV]]	2	195-5	[[V N][A [V [NN V]]]]	2
182-7	[[NN [B V]][N [A V]]]	2	287-3	[[V N][V [V [NN N]]]]	2
209-5	[[NN [N V]][N [V N]]]	2	174-8	[AA [A [V [NN N]]]]	2
219-3	[[NN [N V]][V [N V]]]	2	198-6	[NN [A [V [AA N]]]]	2
197-5	[[NN [V N]][[A [V N]]]	2	282-2	[[N V][NN [V NN]]]	2
185-5	[[NN [V N]][V [V N]]]	2	203-8	[AA [NN [V NN]]]	2
265-1	[[NN NN][B [V N]]]	2	202-3	[NN [[A V][V NN]]]	2
305-1	[[NN NN][N [B V]]]	2	180-5	[NN [[V N][V NN]]]	2
197-3	[[NN NN][V [V N]]]	2	173-2	[NN [NN [A [V N]]]]	2
216-2	[[NN VV][N VV]]	2	277-1	[NN [NN [Conj NN]]]	2
210-8	[[VV NN][V NN]]	2	267-2	[VV [NN [AA V]]]	2
221-3	[[[N V] NN][NN V]]	2	311-2	[[[P NN] N][N [V N]]]	1
210-5	[[[V N][V N]][NN V]]	2	311-1	[[[P NN] N][N VV]]	1
196-7	[[AA NN][AA N]]	2	307-1	[[[A NN] N][NN V]]	1
193-8	[[AA NN][NN V]]	2	215-1	[[[[A V] N] V][N [A V]]]	1
196-3	[[NN [A V]][[N V] N]]	2	193-7	[[[NN N] N][A VV]]	1
220-3	[[NN [N V]][NN N]]	2	171-3	[[[NN N] N][N [V V]]]	1
195-3	[[NN [N V]][NN V]]	2	182-2	[[[NN N] N][N VV]]	1
216-3	[[NN [V N]][[A V] V]]	2	302-1	[[[NN NN] NN] V]	1
181-5	[A [[V NN][V NN]]]	2	288-4	[[[V N] NN][V N]] B]	1
280-4	[V [V [[NN NN] N]]]	2	218-7	[[A [V NN]][A [V N]]]	1
174-7	[A [V [A [V [NN N]]]]]	2	208-7	[[A [V NN]][M [V N]]]	1
191-7	[A [V [NN [A VV]]]]	2	267-3	[[A [V NN]][N [B V]]]	1

185-4	[[A [V NN]][N [V N]]]	1	176-8	[[AA [V N]][B VV]]	1
184-8	[[B [V NN]][A VV]]	1	191-4	[[AA NN][B [V N]]]	1
304-4	[[B [V NN]][V NN]]	1	276-3	[[AA NN][N VV]]	1
174-1	[[N [N NN]][V NN]]	1	173-8	[[AA NN][V VV]]	1
313-4	[[N [V NN]][B [M V]]]	1	195-7	[[AA VV][N [A V]]]	1
298-1	[[N [V NN]][B [V N]]]	1	197-2	[[NN [A V]][A VV]]	1
212-3	[[N [V NN]][N VV]]	1	313-2	[[NN [A V]][B [V N]]]	1
292-2	[[N [V NN]][V [B V]]]	1	206-4	[[NN [A V]][B [V V]]]	1
212-4	[[N [V VV]][N [B V]]]	1	261-3	[[NN [A V]][B VV]]	1
290-2	[[N [V VV]][V NN]]	1	175-8	[[NN [B V]][V [N V]]]	1
222-2	[[V [V NN]][A [N V]]]	1	187-2	[[NN [V N]][A VV]]	1
309-1	[[V [V NN]][B [V N]]]	1	276-4	[[NN [V N]][B [V N]]]	1
181-7	[[A [V [V N]]][AA V]]	1	296-3	[[NN [V N]][Conj [V N]]]	1
300-1	[[A [V NN]][AA V]]	1	221-6	[[NN [V V]][A [V N]]]	1
272-2	[[N [V NN]][AA V]]	1	199-3	[[NN AA][A [A V]]]	1
205-2	[[P [V NN]][NN V]]	1	199-4	[[NN AA][A [V N]]]	1
222-8	[[P NN][V [A NN]]]	1	270-2	[[NN AA][N [B V]]]	1
180-7	[[NN [[P N][V N]]] V]	1	194-6	[[NN AA][V NN]]	1
214-8	[[NN N][V [N [V N]]]]	1	187-5	[[NN NN][A [B V]]]	1
220-1	[[NN N][V [NN N]]]	1	286-2	[[NN NN][A [M V]]]	1
207-2	[[NN V][V [AA N]]]	1	187-6	[[NN NN][B [A V]]]	1
320-3	[[[N V][A V]][A VV]]	1	264-3	[[NN NN][Conj [A V]]]	1
175-2	[[[N V][N V]][N [A V]]]	1	177-8	[[NN NN][N [A V]]]	1
201-1	[[[N V][N V]][V NN]]	1	174-5	[[NN NN][N [M V]]]	1
223-7	[[[N V][V N]][A [B V]]]	1	194-4	[[NN NN][N NN]]	1
198-7	[[[V N] NN][A [B V]]]	1	170-8	[[NN NN][V [N V]]]	1
271-4	[[[V N] NN][A [V N]]]	1	208-2	[[NN VV][A VV]]	1
301-1	[[[V N][A V]][B VV]]	1	180-2	[[V N][V N]][V NN]]	1
184-7	[[[V N][V N]][A VV]]	1	295-4	[[VV NN][A [B V]]]	1

192-2	[[VV NN][A [V N]]]	1	200-7	[A [V [A [[V NN] V]]]]	1
194-8	[[[N V][N V]][NN V]]	1	261-4	[V [V [N [[P NN] V]]]]	1
199-8	[[AA [V N]][AA V]]	1	195-8	[V [N [A [V [V NN]]]]]	1
295-1	[[AA AA][NN N]]	1	199-7	[V [V [V [V [A VV]]]]]	1
191-3	[[AA NN][NN N]]	1	206-6	[A [A [[P N][V NN]]]]	1
174-2	[[AA VV][NN V]]	1	311-3	[A [V [[V N][V XX]]]]	1
206-3	[[NN [A V]][AA V]]	1	271-3	[A [V [NN [A NN]]]]	1
261-2	[[NN [B V]][NN V]]	1	175-7	[A [V [NN [M [V N]]]]]	1
284-4	[[NN [V N]][[V N] V]]	1	182-8	[A [V [NN [N [V N]]]]]	1
209-1	[[NN [V N]][NN N]]	1	296-2	[A [V [NN [V [B V]]]]]	1
200-2	[[NN AA][NN V]]	1	320-4	[B [V [[B N][A [V N]]]]]	1
174-4	[[NN NN][[P N] V]]	1	171-8	[B [V [AA [V NN]]]]	1
302-2	[[NN NN][[V N] V]]	1	219-7	[N [V [[V N][V NN]]]]	1
191-8	[[NN VV][[P N] V]]	1	222-3	[N [V [NN [A NN]]]]	1
293-1	[[VV NN][[V N] V]]	1	299-3	[V [V [NN [A NN]]]]	1
314-4	[A [[V [NN NN]] V]]	1	185-3	[V [V [NN [N [A V]]]]]	1
314-2	[A [[V NN][A VV]]]	1	265-4	[V [V [NN [V NN]]]]	1
222-6	[B [[P NN][V NN]]]	1	222-7	[A [V [AA [NN N]]]]	1
184-6	[B [[V NN][V NN]]]	1	173-7	[B [V [NN [AA N]]]]	1
176-1	[INT [[N NN][N VV]]]	1	309-3	[AA [[[NN N] N] N]]	1
181-8	[M [[P NN][V NN]]]	1	280-3	[NN [[[NN N] N] N]]	1
187-1	[N [[V NN][V NN]]]	1	315-3	[AA [[NN NN] V]]	1
183-7	[V [[P NN][A [A V]]]]	1	207-8	[VV [[AA [B V]] N]]	1
222-5	[V [[P NN][V NN]]]	1	206-2	[AA [A [[P NN] V]]]	1
212-8	[A [[V NN][NN N]]]	1	291-3	[NN [B [[V NN] V]]]	1
317-4	[M [[V NN][NN V]]]	1	178-6	[NN [V [[P NN] V]]]	1
317-3	[A [B [[[NN N] N] V]]]	1	311-4	[NN [X [[P NN] V]]]	1
177-7	[A [V [[[NN N] N] N]]]	1	172-8	[[V N][A [V [V NN]]]]	1
207-7	[A [V [[AA [A V]] N]]]	1	296-1	[[V N][V [V [A [V N]]]]]	1

200-8	[NN [A [A [V NN]]]]	1
301-2	[NN [N [V [N NN]]]]	1
202-1	[[N V][A [V [NN N]]]]	1
178-7	[[N V][M [V [NN N]]]]	1
320-1	[[V N][B [V [AA N]]]]	1
191-1	[AA [A [V [NN V]]]]	1
183-4	[NN [A [V [[P N] V]]]]	1
173-4	[NN [A [V [NN V]]]]	1
198-5	[NN [B [V [AA N]]]]	1
173-3	[NN [B [V [AA V]]]]	1
218-5	[NN [B [V [NN V]]]]	1
266-1	[NN [N [V [NN N]]]]	1
218-6	[NN [N [V [NN V]]]]	1
213-7	[NN [B [[P N][V N]]]]	1
185-2	[[A V][NN [V NN]]]	1
205-8	[[P N][[V N][Conj [V N]]]]	1
204-7	[[P N][NN [V NN]]]	1
183-8	[[P N][VV [V NN]]]	1
288-3	[[V A][[V N][V NN]]]	1
270-4	[[V N][[V N][V NN]]]	1
288-2	[[V N][NN [V NN]]]	1
320-2	[[V N][VV [NN N]]]	1
318-3	[[VV [NN [N [V N]]]]	1
201-8	[AA [[V N][V NN]]]	1
170-3	[NN [[A V][B VV]]]	1
270-1	[NN [[A V][N VV]]]	1
199-2	[NN [[N V][A NN]]]	1
180-8	[NN [AA [A [A V]]]]	1
274-1	[NN [AA [B [V N]]]]	1
302-4	[NN [NN [B [A V]]]]	1
292-4	[NN [NN [V [V N]]]]	1
196-1	[NN [VV [N NN]]]	1
207-1	[VV [[V N][Conj [A V]]]]	1
201-7	[VV [NN [N NN]]]	1
210-7	[AA [[V N][[V N] V]]]	1
308-4	[AA [VV [NN N]]]	1
313-3	[NN [[N V][AA V]]]	1
183-3	[NN [AA [[P N] V]]]	1
199-1	[NN [AA [[V N] V]]]	1
314-1	[VV [NN [NN V]]]	1
273-1	[[N V][N V][N [V N]]]	