

# Chapter 2 *Shijing* Sub-grammar

## 2.1 General description of the raw corpus

*Shijing*, also known as ‘Book of Songs’, is the earliest written record of verse in Chinese history and generally regarded as the origin of Chinese literature. It is an anthology compiled around 600 BC and comprising of 305 poems, which were composed during the *Zhou* Dynasty (1066-771 BC) and the *Spring and Autumn* period (770-476 BC) (see Appendix I for a chronology). These poems fall into three main subgenres: (1) *Feng* (literally ‘Airs’), poems thematizing on life of ordinary people; (2) *Ya* (literally ‘Elegance’), odes exalting life of the nobility and the court; and (3) *Song* (literally ‘Ode’), hymns of the Temple and the Altar sung on religious occasions. Of the 305 *Shijing* poems, 160 belong to *Feng*, 105 to *Ya*, and 40 to *Song*.

Like many ancient, inchoate literary forms, *Shijing* poems were mainly sung to music and accompanied by dance at the time of their composition, although the tunes have since long been lost<sup>1</sup>. One possible exception is the *Feng* poems, which were, according to the historical records (Chen 1994), also recited back then. For modern speakers with no access to the original tunes, reciting is the only mode of performing the ancient verse. Of the three subgenres, *Feng* is also best recited by modern speakers, presumably due to the fact that its themes center around the daily life of ordinary people. On this account, only *Feng* poems are examined here; specifically, of the 160 *Feng* poems, we randomly select the 80 odd-numbered ones totaling 1320 lines as the *Shijing* data for the present research.

## 2.2 Methodological issues and preview of the sub-grammar

### 2.2.1 Methodological issues

As mentioned in Chapter 1, the rest of this chapter consists of two parts: the development of the sub-grammar and the formal account of the metrical harmony judgment. The analytical scheme for developing the sub-grammar was already presented there and will not be repeated. In this section we only draw attention to three methodological issues. First, in developing an OT grammar, we need to motivate the introduction and ranking of each constraint. The ranking can only be determined on two accounts: either by conflict or by transitivity. (For more on the analytical procedure in reaching an OT grammar, see e.g. Kager 1999).

---

<sup>1</sup> Three performance styles have been suggested for classical Chinese verse: singing, chanting and reciting (Chen 1994). Verse is sung to a certain musical tune, and recited to a certain rhythm but without reference to musical tunes. Chanting lies in between: more musical than mere recitation, but less so than singing. As reciting is the only mode of modern speakers’ performance of classical Chinese verse and the sole concern of the present research, the word ‘performance’ is used here to refer to recitation throughout.

Second, evidently not all types of line structures contribute equally to the development of the sub-grammar: some constitute crucial evidence for new constraints and/or new ranking while others may be adequately accounted for by the sub-grammar reached till that point, which for simplicity sake, will be referred to as ‘emergent sub-grammar’, and thus provide no argument for new constraints or ranking. Although analytically such lines have little to offer, we nonetheless choose to devote brief attention to them by including some examples and illustrate their scansion. This practice is out of two considerations. One is to enrich the present study with a descriptive dimension and the other is to present a ‘panoramic’ view of the operation of the emergent sub-grammar.

The third issue bears closely on this: to enhance its readability, Section 2.3 is divided into sub-sections according to the line type in terms of syllable numbers; within each sub-section we examine the grammatical structures that are crucial in developing the sub-grammar. Where a certain line type is not advancing new arguments for the sub-grammar, selective examples are provided and their scansions illustrated. This organizing principle is also adhered to in the corresponding sections of the following chapters on the other four genres.

Finally, regarding notational convention, we basically follow the standard OT practice in using tableaux to illustrate the crucial ranking between conflicting constraints, but occasionally the tableau is also used in a global way to demonstrate the working of the emergent sub-grammar. The usage of solid and dotted lines in the tableau was presented in Section 1.2.2.1 of Chapter 1. The Hasse graph is also used frequently. An extended use of tableau, namely, tableau des tableaux, is used in the formal account of the metrical harmony judgment in Section 2.4 and its working will be articulated there.

## 2.2.2 Preview of the sub-grammar

We now briefly outline the modern sub-grammar for *Shijing*. First, not surprisingly, one of the ‘staples’ of the sub-grammar is the constraint monitoring the size of the prosodic unit above the syllable level, namely, prosodic foot. We argue that binary feet are preferred and that there arises the need to split this requirement into BINMAX and BINMIN, respectively imposing binarity as the maximal and minimal foot size. Second, as stated earlier, we are not interested in those scansions blind to the syntax or semantics of the line; a good scansion necessarily refers to the structure and the meaning of the line. Hence constraints governing the matching between the input, i.e. the grammatical structure of the line, and the output, i.e. its prosodic structure, also play a role in the sub-grammar. Two constraints belong to this constraint family: ANCHOR and ANCHOR- $I_{SB}O_{PhP}$  (with SB and PhP respectively standing for Strongest Boundary and Phonological Phrase), which respectively check the boundary matching at the lowest and highest levels of the hierarchy. Third, the prosodic hierarchy of Chinese is discussed and we show that the distribution of monosyllabic feet in the prosodic constituent of Phonological Phrase (PhP) is subject to the constraint that such feet cannot occur PhP-finally. This constraint is couched into \*PHP-FINAL-MONOFT. Fourth, *Shijing* is characterized by the wide use of interjections which we argue are represented and parsed differently from lexical syllables. This is reflected in the presence of a constraint specifically targeted at the prosodic parsing of

interjections in the *Shijing* sub-grammar, namely, GOODFTINTERJ. Section 2.3 provides a detailed account of how these constraints are motivated and ranked.

## 2.3 *Shijing* sub-grammar

### 2.3.1 Point of departure: 2-syll lines

All 2-syll *Shijing* lines share the grammatical structure of [SS] and are scanned as (SS). Examples are:

- (1)        *shui2 yu3?*<sup>2</sup>  
             Who with  
             ‘With whom (shall I go)?’
- (2)        *du2 xi1*  
             alone rest  
             ‘(I) rest alone’
- (3)        *shi4            wei1*  
             interjection interjection  
             ‘ah --’

Analytically, though, such lines offer little (except perhaps the preference for binary feet), and we just pass and move to the 3-syll lines.

### 2.3.2 BINMAX >> BINMIN and \*IP-FINAL-MONOF: evidence from 3-syll lines

For 3-syll lines, two grammatical structures can be identified, i.e., [SS]S versus S[SS]<sup>3</sup>. The grammatical structure only constitutes the input; the optimal output, i.e. the actual scansion, for all 3-syll lines is (S)(SS), irrespective of the input structure<sup>4</sup>. Below are some examples to illustrate the scansion (with the foot boundary indicated by the round brackets) of verse lines respectively of the above-mentioned grammatical structures (indicated by square brackets):

<sup>2</sup> Chinese (including classical Chinese) is basically a SVO language but with a few exceptions. (1) is one such exception where the object, being a pronoun, is inverted to precede the preposition.

<sup>3</sup> Strictly speaking, the bracketing should respectively be [[SS]S] and [S[SS]]; for simplicity sake, we leave out the outer layer of brackets.

<sup>4</sup> However, this indifference to the input structure is not inconsequential. In cases of those verse lines where the prosodic and the grammatical structures mismatch, the native speaker experiences a sense of tension when such lines are recited. The degree of tension may vary depending on how gross the mismatch is; for example, (4) below is felt to be tenser than, say, (7) and (8), which are cognized as metrically perfectly harmonious. The reason is because structurally, the two ‘shen1’'s in (4) parse together, being the two elements of a reduplication whereas for (7) and (8) the first two syllables are structurally loose too. Thus, it appears that the input structure, and arguably the juncture strength recorded in the coding may well have a delicate effect on the native speaker’s cognization of the line. For the present, our focus remains the development of the constraint ranking; the issue of metrical harmony will be addressed in Section 2.4.

- (4)  $[shen1\ shen1]\ [xi1] \rightarrow (shen1)(shen1\ xi)$   
       many                  interjection  
       ‘(They are) many’
- (5)  $[zhi1\ zi3]\ [gui1] \rightarrow (zhi1)(zi3\ gui1)$   
       this personreturn  
       ‘This person returns’
- (6)  $[yi3\ yan1]\ [zai3] \rightarrow (yi3)(yan1\ zai3)$   
       already finished interjection  
       ‘Ah, (it is) already over’
- (7)  $[jiang1][you3\ zhu4] \rightarrow (jiang1)(you3\ si4)$   
       river have bank  
       ‘The river has banks’
- (8)  $[shen1][ze2\ li4] \rightarrow (shen1)(ze2\ li4)$   
       deep then wade  
       ‘(If the river is) deep, then (I will) wade across it’.

For a 3-syll input, among the theoretically infinite number of potential outputs produced by GEN, the relevant ones are (SSS) and (SS)(S)<sup>5</sup>. However, they are less harmonious than the optimal (S)(SS). Using the symbol ‘ $\succ$ ’ to stand for ‘wins over’, we have

$$(9) \quad (S)(SS) \succ (SSS)$$

$$(10) \quad (S)(SS) \succ (SS)(S).$$

(9) shows that a trisyllabic foot is worse than a monosyllabic one, when there are no other alternatives (now that the possibility of leaving a syllable unparsed is already precluded by the high-ranking PARSE-SYL constraint). This calls for two markedness constraints governing the well-formedness of feet: BINMAX and BINMIN<sup>6</sup>. The former

<sup>5</sup> Other potential forms include, for example, (S), (SS), S, SS, S(SS), (SS)S, and (SSSS) etc. which result from a wild operation of phonological processes such as deletion and addition, parsing and non-parsing of syllables etc. But they bear little relevance to the present discussion and so are not considered here. Their failure to emerge as the optimal form can be easily accounted for by postulating that the responsible constraints (such as PARSE-SYL, MAX, DEP; see Prince and Smolensky 1993) are highly ranked. Furthermore, it is of interest to point out that S(SS) and (SS)S can also be eliminated on the account that the prosodic structure in Chinese observes the Strict Layer Hypothesis (Selkirk 1984; Nespor and Vogel 1986).

<sup>6</sup> We suggest that this binarity requirement can be traced to a deeper and more fundamental origin, namely, the preference for the alternation between strong and weak beats at regular intervals to create a rhythmic effect in both linguistic and non-linguistic domains. This preference is most readily expressed by the binarity of the beat group, and typically a strong beat followed by a weak one. The crucial point here is that this preference is not restricted to the linguistic domain; it is in fact a constraint deeply rooted in the cognitive system of human beings (cf. Chatman 1965; also see Dauer 1983; Dell 1984; Selkirk 1984: 36-37; Hayes 1984: 59), and is also suggested to be attributable to neurophysiological mechanisms (Fussell 1979). Another indication of its non-linguistic nature is the fact that this preference is also pervasive in rhythmical forms other than language, or verse, for that matter, such as music and dancing. Furthermore, this binarity preference is innate and universal, evidenced by its wide

requires a foot to be maximally binary, or, more precisely, to contain maximally two syllables in the present context<sup>7</sup>, whereas the latter stipulates that a foot is minimally binary, i.e. contains minimally two syllables. In the phonological literature, sometimes these two constraints are not distinguished and represented as a cover constraint FT-BIN which merely requires that feet be binary under moraic or syllabic analysis (Prince 1980; Kager 1989; Prince and Smolensky 1993). However, here arises the motivation for each of them. The fact that a monosyllabic foot but not a trisyllabic one appears in the optimal output indicates that a monosyllabic foot is conditionally acceptable and that it is worse to parse the input string into a trisyllabic foot than into a monosyllabic plus a disyllabic one. In terms of OT ranking, this preference is tantamount to BINMAX >> BINMIN, and this ranking is true for both input structures of 3-syll lines under discussion here.

(11) (i)

[SS]S	BINMAX	BINMIN
☞ (S)(SS)		*
(SSS)	*!	

(ii)

S[SS]	BINMAX	BINMIN
☞ (S)(SS)		*
(SSS)	*!	

Now, looking at (10), we find that these two constraints and their ranking fall short of accounting why (SS)(S) loses to (S)(SS). Furthermore, the fact that (S)(SS) is the optimal output in both cases where the inputs differ suggests that a further constraint needs to be motivated which is necessarily an output-oriented markedness one making no reference to the input structure. Compare the optimal form (S)(SS) with its competitor (SS)(S), and it is evident that the latter has a monosyllabic foot at the end of the verse line, which, in this case, contains two feet. As pointed out in Chapter 1, a verse line corresponds to Intonational Phrase (IP) prosodically, largely on account of the fact that when recited, a line falls under a unifying intonational contour, which is argued to be the defining characteristic of an IP (Chafe 1974; Pierrehumbert 1980; also Hayes 1989). Hence the operative constraint here is one that forbids monosyllabic feet from occurring IP-finally. This is formulated as

- (12) \*IP-FINAL-MONOF  
Do not place the monosyllabic foot at the final position of an IP.

---

occurrence as the most fundamental rhythmical pattern in relatively pristine art forms such as nursery rhymes and tribal dance across languages and cultures. In view of all this, we suggest that this BINARITY constraint is, more precisely speaking, a eurhythmical constraint that resides universally in human beings.

<sup>7</sup> Binariness of foot requires that feet be binary under either moraic or syllabic analysis; in the current context, a syllabic analysis is more relevant, as Chinese syllables are, with very few exceptions, heavy, and hence bimoraic. The moraic analysis, though, comes into relevance in the discussion below about the parsing of interjection syllables (Section 2.3.4.1).

This constraint, which is evidently also an output-oriented markedness one, governs the position of the monosyllabic foot in higher-level prosodic units. At first sight, it may appear somewhat language-specific; however, we argue that this constraint is in fact a variant of NONFINALITY, a well-established universal phonological constraint, which is reformulated as RHYTHM in Hung (1994). NONFINALITY requires that no prosodic head be final in the Prosodic Word (Prince and Smolensky 1993); in other words, the Prosodic Word must not end in a head syllable of a foot. Duanmu (1999, 2000) convincingly argues that Chinese has both syllabic and moraic feet and that Chinese feet are strictly trochaic at both the syllabic and moraic levels. According to him, the good and bad foot structures in Chinese are:

## (13) (i) Good foot structures in Chinese

x			x			
(S	S)	or	(S	S)		Syllabic foot
(μμ)	(μμ)		(μμ)	(μ)		Moraic foot
x	x		x			

## (ii) Bad foot structures in Chinese

x			x			
(S	S)	or	(S	S)		Syllabic foot
(μ)	(μ)		(μ)	(μμ)		Moraic foot.
				x		

Furthermore, Duanmu assumes that a foot must have (at least) two syllables and adopts the notion of zero syllable proposed in Burzio (1994). In line with this proposal, he considers that a monosyllabic bimoraic syllable constitutes a disyllabic (albeit still bimoraic) foot containing a zero syllable, which is also well-formed. This is represented as (with zero syllable indicated by 0):

(14)	x		
	(S 0)	Syllabic foot	
	(μμ)	Moraic foot.	
	x		

Our position here is to circumvent the need for zero syllable and directly accept monosyllabic bimoraic feet as well-formed, at least in the verse context, on the account that such feet can indeed surface in the optimal scansion of verse lines, even though they violate BINMIN<sup>8</sup>. Thus, we modify Duanmu's inventory of good foot structures presented in (13)(i) by including the monosyllabic, bimoraic foot whilst at the same time discarding the bimoraic, disyllabic foot in (14). Obviously, the single syllable in a monosyllabic foot always carries the stress. Furthermore, following Hammond (1997), only heavy, i.e. bimoraic syllables can carry stress, whilst a light, i.e. monomoraic syllable cannot. Therefore, using H and L to represent the quantitative structure of a syllable, namely, H for heavy, bimoraic and L for light, monomoraic, well-formed feet in Chinese are only of three types: (i) (HH), disyllabic trochee where both syllables are heavy and the first syllable is the head; (ii) (HL),

<sup>8</sup> On the other hand, if we introduce the conventional BINARITY (or FT-BIN) constraint which requires feet to be binary at the syllabic or moraic level (Kager 1999), then monosyllabic feet satisfy this constraint at the moraic level.

disyllabic trochee where only the head syllable is heavy; and (iii) (H), monosyllabic bimoraic trochee bearing its own stress. In contrast, ill-formed feet in Chinese include (i) (LL), disyllabic foot constituted by two monomoraic syllables; (ii) (LH), disyllabic foot where only the head syllable is monomoraic and the non-head is bimoraic; and (iii) (L), monosyllabic monomoraic foot<sup>9</sup>. This is illustrated below (where both the syllabic and moraic stress are indicated by x and syllable weight marked as H or L):

## (15) (i) Good foot structures in Chinese

x	(S	S)	or	x	(S	S)	or	x	(S)	
(μμ)	(μμ)			(μμ)	(μ)			(μμ)		Syllabic foot
x	x			x				x		Moraic foot
H	H			H	L			H		

## (ii) Bad foot structures in Chinese

x	(S	S)	or	x	(S	S)	or	x	(S)	
(μ)	(μ)			(μ)	(μμ)			(μ)		Syllabic foot
										Moraic foot
L	L			L	H			L		

Therefore, by forbidding the head syllable of a foot at the end of a Prosodic Word (PrW), NONFINALITY in effect bans the occurrence of the monosyllabic foot which, to mention in passing, must be bimoraic if the foot is legitimate, at the final position of a Prosodic Word. This ban is precisely in the same spirit as \*IP-FINAL-MONOFT proposed here.

The next question now is how this constraint should be ranked with the other two constraints proposed so far, i.e. BINMAX and BINMIN. To begin with, BINMAX and \*IP-FINAL-MONOFT do not conflict: indeed, both must be undominated in the constraint hierarchy. The reason is that neither a potential output violating BINMAX, i.e. (SSS) nor one violating \*IP-FINAL-MONOFT, i.e. (SS)(S) can emerge as optimal. Both these two constraints impose a strict, non-negotiable requirement to filter out the sub-optimal forms from the candidate set.

As to the ranking between BINMIN and \*IP-FINAL-MONOFT, again we find that they are not in conflict: a violation of the latter is necessarily accompanied by that of the former although the reverse is not necessarily true. In more concrete terms, if a potential output violates \*IP-FINAL-MONOFT, which means it has an IP-final monosyllabic foot, then due to this monosyllabic foot, it simultaneously violates BINMIN; however, conversely, a potential output can have a monosyllabic foot, hence violating BINMIN but not violating \*IP-FINAL-MONOFT if this foot is not IP-final. It is impossible for a potential form to violate \*IP-FINAL-MONOFT without violating BINMIN; therefore, for candidates violating \*IP-FINAL-MONOFT, a violation mark under BINMIN is not discriminating, and is therefore ‘cancelable’ (Prince and Smolensky 1993). This is illustrated below (since both constraints are purely output-oriented, the input structure is inconsequential and thus not specified here):

<sup>9</sup> The monosyllabic monomoraic foot is also referred to as a ‘degenerate foot’ which contains a single light syllable. Many languages have an absolute ban on such feet, which can also be accounted for by the FT-BIN constraint being undominated in the grammar (Kager 1999: 161).

(16)

SSS	*IP-FINAL-MONOFT	BINMIN
(SS)(S)	*!	*
𠄎 (S)(SS)		*

On the other hand, imagine two candidates, one incurring more than one violation of BINMIN but satisfying \*IP-FINAL-MONOFT, say (S)(S)(S)(SS), while the other incurring only one violation of BINMIN and one violation of \*IP-FINAL-MONOFT, say (SS)(SS)(S). That both are sub-optimal forms indicates that the ranking between the two constraints is immaterial and they do not conflict. This is shown below<sup>10</sup>:

(17)

SSSSS	*IP-FINAL-MONOFT	BINMIN
(S)(S)(S)(SS)		***
(SS)(SS)(S)	*	*

We illustrate the working of the constraint hierarchy arrived at so far with the two grammatical structures for 3-syll lines below. As mentioned earlier, to present the constraint hierarchy in a linear way, the ranking BINMAX >> \*IP-FINAL-MONOFT is trivially assigned to the non-conflicting pair.

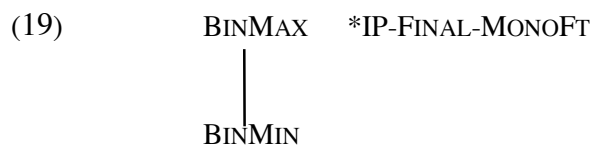
(18) (i)

[SS]S	BINMAX	BINMIN	*IP-FINAL-MONOFT
𠄎 (S)(SS)		*	
(SS)(S)		*	*!
(SSS)	*!		

(ii)

S[SS]	BINMAX	BINMIN	*IP-FINAL-MONOFT
𠄎 (S)(SS)		*	
(SS)(S)		*	*!
(SSS)	*!		

To temporarily summarize, these three constraints, all being markedness ones, suffice to select the optimal output for 3-syll lines. Illustrated in the Hasse graph, the emergent sub-grammar at this stage is:



<sup>10</sup> To illustrate this argument, we have to move ahead of ourselves to present examples from the 5-syll lines, which are optimally parsed into (SS)(S)(SS) or (S)(SS)(SS), depending on the input structure, but never the two given here in (21).



with the line indicating the dominance relation between two constraints. \*IP-FINAL-MONOF<sub>T</sub> stands alone alongside the ranking pair BINMAX and BINMIN, because it conflicts with neither of them.

### 2.3.3 More on the sub-grammar: evidence from 4-syll lines

Moving on to 4-syll lines, we note that a crucial fact is that the lines are invariably scanned into two disyllabic feet, i.e. (SS)(SS), irrespective of their grammatical structures<sup>11</sup>. For example, the following verse lines differ in their grammatical structures, but are scanned the same way<sup>12</sup>:

(20) *[wu3 mei4][qiu2 zhi1] → (wu3 mei4)(qiu2 zhi1)*  
 awake asleep desire her  
 ‘(I) desire her both when I am awake and when I am asleep’

(21) *[yi4 [er3 [zi3 sun1]]] → (yi4er3) (zi3 sun1)*  
 suit you(r) children grandchildren  
 ‘(It) suits your children and grandchildren’

(22) *[[she4 [bi3ju1]] yi3] → (she4 bi3)(ju1yi3)*  
 climb that hill interj  
 ‘Ah, (I) climb that hill’

<sup>11</sup> It deserves mentioning that this is only true for the verse scansion; if read in a prose style, or put in a prose context, 4-syll lines of certain structures may be scanned in ways other than (SS)(SS). For example,

*zai4 [he2 [zhi1 zhou1]]*  
 at river ‘s bank  
 ‘at the river’s bank’

is scanned as *(zai4 he2) (zhi1 zhou1)* when read as a verse line, but when read in the prose context, it is most likely to be scanned as *(zai4) (he2 zhi1 zhou1)* with the first monosyllabic foot considerably lengthened and the middle syllable in the trisyllabic foot reduced both segmentally (into schwa) and tonally (into neutral tone). However, the present study is solely concerned with the scansion of verse lines when they are recited as such (as opposed to recitation in a ‘prose-way’), and it stands to reason that the ‘prose scansion’ can be accounted for by re-ranking some of the constraints proposed here for the ‘verse scansion’ (cf. Golston 1998; Schlepp p.c.). Alternatively, assuming that both the segmental and tonal reduction can be attributed to the deletion of one mora from the originally bimoraic syllable, we could argue that one of the most distinct features between the grammar for verse scansion and that for prose scansion is that the former attaches greater importance to the preservation of syllable weight than the latter, hence reduction of syllable weight is forbidden in the verse scansion. Along this line, we could propose that the constraint Max- $\mu$ , which requires the conservation of the input mora in the output, is ranked very high in verse scansion grammar. In contrast, in prose scansion, this constraint is lowly ranked, at least dominated by some constraints requiring the conservation of the grammatical structure of the input.

<sup>12</sup> As in the case of 3-syll lines, the apparent ignoring of the input structure by the optimal parse is not totally without cost: for example, the 4-syll lines given here are experienced differently by the native speaker. For example, (20) is felt to be metrically very harmonious, in contrast to, say, (22), which is felt to be metrically much tenser. This is going to be discussed in the next section.

(23)  $[[[wo3\ ma3]\ tu2]\ yi3] \rightarrow (wo3\ ma3)\ (tu2yi3)$   
 I/my horse tired interj  
 ‘Ah, my horse is tired’

(24)  $[fei3\ [[wo3si1]\ cun2] \rightarrow (fei3\ wo3)\ (si1\ cun2)$   
 not my thought lie  
 ‘(they are) not where my thoughts lie’

This data is fully accounted for by the emergent sub-grammar. As shown below, (SS)(SS) always incurs the minimal violation and comes out as the winner. The grammatical structure of the line is unspecified, due to its irrelevance:

(25)

SSSS	BINMAX	BINMIN	*IP-FINAL-MONOFt
☞ (SS)(SS)			
(SS)(S)(S)		*!*	*
(S)(SS)(S)		*!*	*
(S)(S)(SS)		*!*	
(SSS)(S)	*!	*	*
(S)(SSS)	*!	*	

The optimal scansion (SS)(SS) incurs no violation under the current constraint hierarchy. This, however, does not contradict the claim that in OT there are no ‘perfect’ winners and that even the optimal output is bound to incur some violation, which is referred to as the ‘minimal violation’ (Prince and Smolensky 1993), as the constraint hierarchy so far is only a subset of the sub-grammar in its final shape when all line types are examined. There, the optimal output (SS)(SS) will predictably violate some constraint(s)<sup>13 14</sup>.

<sup>13</sup> Interestingly, this minimal violation of the constraints achieved when the input is of the structure [SS][SS] may also account for, at least to a certain extent, why in modern Chinese, the overwhelming majority of 4-syll idioms have the grammatical structure of [SS][SS]. More interestingly, even in those cases which do not have this symmetrical structural representation, the established scansion is (SS)(SS); a compelling example is

$[[yi4\ [yi1\ dai4]]\ shui3] \rightarrow (yi4\ yi1)\ (dai4\ shui3)$   
 one clothing belt water  
 ‘(separated by only) the water of the width of a clothing belt’.

It might be of further interest to note how this optimal scansion may sometimes serve in turn to corrupt the interpretation of the input string. In this case, for example, some informants do harbor, next to the correct interpretation, such a ‘corrupted’ interpretation as ‘a piece of clothing brings the water’ which, though not making much sense, corresponds to a ‘back-propagated’ grammatical structure  $[yi4\ yi1]\ [dai4\ shui3]$  from the optimal scansion  $(yi4\ yi1)\ (dai4\ shui3)$ .

<sup>14</sup> Here might be a good point to briefly discuss the Gestalt effect in verse scansion. Although as stated earlier, the present study is confined to the verse line level, in actual verse recitation, the native speaker delivers the whole verse in its entirety rather than just individual lines. Consequently, the scansion of verse line should ideally be perceived in its verse context, in particular in the context of its preceding lines where the prevailing rhythmic pattern has been established in the performer’s mind and is most likely to carry on with momentum to influence how the incoming new lines are to be scanned. This holisticness in the perception of patterns and its subsequent influence on the individual constituting elements are known as the Gestalt effect in cognitively oriented studies on a range of topics. For example, Meyer (1956) is concerned with the Gestalt effect in the perception of music whereas Arnheim (1967) with that in the perception of the visual arts. Tsur (1998) applies it to the perception of poetic rhythm and versification. In verse scansion, the Gestalt effect may be rather strong: the mind

### 2.3.4 ALIGNR (Ft, IP), ANCHORING and GOODFTINTERJ: evidence from 5-syll lines

So far the emergent sub-grammar solely consists of markedness constraints that are only concerned with the well-formedness of the output. The scansion of 5-syll lines exposes the insufficiency of this sub-grammar and provides crucial evidence for inclusion of faithfulness constraints in the sub-grammar.

First, consider lines of the grammatical structure of  $[[SS][SS]]S$ , which are scanned as  $(SS)(S)(SS)$ . For example:

- (26)  $[[ru2\ ci2][liang2\ ren2]]\ he2 \rightarrow (ru2\ ci2)(liang2)\ (ren2\ he2)$   
 like this fine person interj \* (ru2 ci2) (liang2 ren2) (he2)  
 ‘Ah, a fine person like this!’

The analysis under the emergent sub-grammar so far is:

(27)

$[[SS][SS]]S$	BINMAX	BINMIN	*IP-FINAL-MONOF <sub>T</sub>
$\text{E} (SS)(S)(SS)$		*	
$(SS)(SS)(S)$		*	*!
$E (S)(SS)(SS)$		*	
$(SS)(SSS)$	*!		
$(SSS)(SS)$	*!		

As the current constraint hierarchy stands now, the potential (and in fact suboptimal) scansion  $(S)(SS)(SS)$  incurs the same violation as the desired output  $(SS)(S)(SS)$ . (We use E and  $\text{E}$  to respectively mark out the undesired and the true winner.) Obviously, to ensure the latter as the sole winner, some constraint is needed that cashes in on

---

maintains as much of the current pattern as possible, and even tends to perpetuate a metrical or rhythmical pattern initially established. One of the most important ways in which the Gestalt effect works is through the formulation of expectation, on the assumption that ‘the mind is constantly striving toward completeness and stability of shapes (Meyer Ibid.: 87)’. Accordingly, the mind tends to complete what was incomplete, to regularize what was irregular. In the context of verse, this property is reflected in the notion of ‘metrical set’ (Chatman 1965), which produces the assimilative power of rhythm from one line to the next in a poem. The metrical set established by previous lines might be so strong as to assimilate an apparently deviating line into this already established pattern, provided of course, this line occurs significantly infrequently.

With regard to the present analysis, it needs to be emphasized that the Gestalt effect remains viable even though the analytical domain is the line. Here, for instance, lines of the structure  $S[[SS]S]$  (see (24) above for an example) can arguably also be scanned as  $(S)(SS)(S)$ , at least according to some informants. But once considered in the verse context in which it occurs,  $(SS)(SS)$  apparently comes as the only natural scansion whilst  $(S)(SS)(S)$  sounds awkward and out of place. The reason is because the lines preceding this one have already firmly established  $(SS)(SS)$  as the predominant scansion for 4-syll lines, a scansion so deeply entrenched in the reader’s mind that it takes more than one single line to shake off its effect. As stated back in Chapter 1, only the best accepted scansion is considered as the optimal scansion for an input line, and accordingly, our strategy in handling such cases is to opt for the scansion that is good both in isolation and in its verse context. Thus, with this guideline, the optimal scansion for 4235 is  $(SS)(SS)$ , as indicated above. Formally, we might suggest that the Gestalt effect can be expressed in OT terms as an inter-linear faithfulness constraint, which must be highly-ranked (cf. Holtman’s (1996) work on an OT account of rhyme). We leave this issue open here.

what distinguishes (SS)(S)(SS) from (S)(SS)(SS). Compare the two scansions and we notice that in (SS)(S)(SS) the right edges of the feet are better aligned than in (S)(SS)(SS) to the right edge of the IP, which is, as pointed out before, the prosodic correspondent of the verse line. This readily suggests a constraint from the well-attested alignment constraint family. More specifically, to give (SS)(S)(SS) an edge over its competitor (S)(SS)(SS) to finally win, the following constraint needs to be introduced:

- (28) ALIGNR (FT, IP)  
The right edge of every prosodic foot coincides with the right edge of an IP.

As is typical of alignment constraints, ALIGNR (FT, IP) is evaluated gradiently, and specifically in terms of the number of syllables separating the two right edges that are required to coincide but are actually not. For example, (SS)(S)(SS) incurs 5 (= 2+3) violations of this constraint whilst (S)(SS)(SS) incurs 6 (= 2+4)<sup>15</sup>.

The next question is how to rank this newly introduced constraint with the other constraints. For convenience sake, we repeat the tableau in (27) above and indicate the satisfaction/violation of the new constraint ALIGNR (FT, IP) by the same set of candidates:

(29)

[[SS][SS]]S	BINMAX	BINMIN	*IP-FINAL-MONOFT
☞ (SS)(S)(SS)		*	
(SS)(SS)(S)		*	*!
E (S)(SS)(SS)		*	
(SS)(SSS)	*!		
(SSS)(SS)	*!		

[[SS][SS]]S	ALIGNR (FT, IP)
(SS)(S)(SS)	5
(SS)(SS)(S)	4
(S)(SS)(SS)	6
(SS)(SSS)	3
(SSS)(SS)	2

This way, the ranking for ALIGNR (FT, IP) with other constraints becomes explicit. To begin with, BINMAX >> ALIGNR (FT, IP), because if the reverse were true, i.e. ALIGNR (FT, IP) >> BINMAX, then the potential form (SSS)(SS), or more precisely, (SSSSS), for that matter, with zero violations of ALIGNR (FT, IP) would win. This is shown below:

<sup>15</sup> The violation of ALIGNR (Ft, IP) by a candidate is calculated via the degree of misalignment between the right boundary of each foot contained in the IP and that of the IP in terms of the number of intervening syllables, and then adding up the numbers. For example, (SS)(S)(SS) contains three feet, the rightmost foot (SS) has its right boundary perfectly aligned with the right boundary of the IP, the middle foot has two syllables between its right boundary and that of the IP, while the first foot has three syllables between the two right boundaries. Thus, the overall violation of ALIGNR (Ft, IP) by (SS)(S)(SS) is the sum of these two numbers: 2+3 = 5.

(30)

[[SS][SS]]S	BINMAX	ALIGNR (FT, IP)
$\text{☞} \text{ (SS)(S)(SS)}$		5
(SSS)(SS)	*!	2
(SSSSS)	*!	0

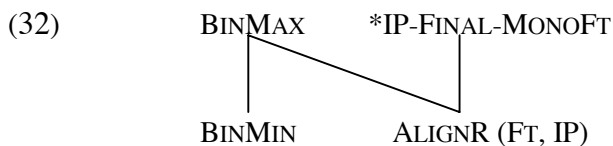
Second, there is no evidence for crucial ranking between BINMIN and ALIGNR (FT, IP). The reason is that they have the same interests and do not conflict. Indeed, given the inviolability of BINMAX, BINMIN is inevitably violated in 5-syll lines.

Third, \*IP-FINAL-MONOFT >> ALIGNR (FT, IP), otherwise (SS)(SS)(S) would win. The ranking argument is

(31)

[[SS][SS]]S	*IP-FINAL-MONOFT	ALIGNR (FT, IP)
$\text{☞} \text{ (SS)(S)(SS)}$		5
(SS)(SS)(S)	*!	4

Hence the emergent sub-grammar is, presented in the Hasse graph, as follows:



Now, note that this emergent sub-grammar still consists only of markedness constraints, which implies that (SS)(S)(SS) would always emerge as the optimal scansion independent of the input. However, unlike the 4-syll lines, where the optimal scansion remains the same irrespective of the input structure, the optimal scansion may vary for the 5-syll lines with different grammatical structures. For example, lines of the grammatical structures [S[SS]][SS], S[[SS][SS]] and [S[[SS]S]]S are all optimally scanned as (S)(SS)(SS) rather than (SS)(S)(SS):

(33)  $[san1[zhi1 ri4]] [yu2 lu3] \rightarrow (san1) (zhi1 ri4)(yu24 lu3)$   
 three particle day very heavy  
 ‘On the days of the third (month), (the wind) is very heavy’

(34)  $yuan3 [[fu4 mu3] [xiong1 di4]] \rightarrow (yuan3) (fu4 mu3) (xiong1 di4)$   
 distant fathermother brother brother  
 ‘(She) becomes distant from her parents and brothers’

(35)  $[xing2 [[yu3 zi3] huan2]] xi1 \rightarrow (xing1) (yu3 zi3) (huan2 xi1)$   
 go with you return interj  
 ‘Ah, (I) go and return with you’

As just suggested, given the current constraint hierarchy, (SS)(S)(SS) will always win over other potential forms; (S)(SS)(SS) will always lose due to more violations of ALIGNR (FT, IP). This is illustrated below with S[[SS][SS]]:

(36)

S[[SS][SS]]	BINMAX	BINMIN	*IP-FINAL-MONOFT	ALIGNR (FT, IP)
☛ (SS)(S)(SS)		*		5
(SS)(SS)(S)		*	*!	4
☛ (S)(SS)(SS)		*		6!
(SS)(SSS)	*!			3
(SSS)(SS)	*!			2

It is evident that a mere re-ranking of the current four constraints will not lead (S)(SS)(SS) to become the winner. A new constraint is in order, and crucially this constraint cannot merely be concerned with the output. Now carefully compare both the true winner, (S)(SS)(SS) and the unwanted winner (SS)(S)(SS) against the input structure. The most distinct difference is that the edges of the prosodic feet in (S)(SS)(SS) fully match those of the grammatical constituents in the input but it is not so with (SS)(S)(SS).

This correspondence of edges between the input and the output structures can be further traced to the correspondence of the segments respectively at the edges of the input and the output structures. This brings to bear the ANCHORING constraint proposed in McCarthy (2000:184). There two subcategories of ANCHORING are distinguished, i.e. ANCHOR-POS and ANCHOR-SEG, which respectively require the conservation of a segment's position and a segment per se occurring at the designated edge. We argue that of the two, ANCHOR-SEG is relevant, which requires that segments at the designated peripheries of two representations S1 and S2 correspond to each other, thus in effect requiring the correspondence of edges between S1 and S2 (Kager 1999)<sup>16 17</sup>. For simplicity sake, below we will refer to ANCHOR-SEG as ANCHOR.

Furthermore, McCarthy (Ibid: 183) argues for the 'existence of distinct but symmetric ANCHORING constraints from S1 to S2 and from S2 to S1', a move that 'parallels to an established symmetry in Correspondence Theory'. In other words, ANCHOR is a directional constraint can be broken down into two sub-constraints; McCarthy (Ibid.: 185) expresses them respectively as I-ANCHOR and O-ANCHOR, which are respectively comparable to MAX and DEP. For ease of constraint evaluation, we choose to formulate the two ANCHORING sub-constraints as follows:

<sup>16</sup> Needless to say, this is only true when no syllables are deleted or inserted. This is already taken care of by the highly ranked MAX and DEP mentioned back in Footnote 5.

<sup>17</sup> Indeed, as Kager (1999) points out, the ANCHORING constraint was first proposed in McCarthy and Prince (1995a) to replace ALIGNMENT (McCarthy and Prince 1993a) in those applications concerned with the correspondence between the elements at the designated peripheries (left or right) of S1 and S2, such as Input and Output, Base and Reduplicant etc.. In McCarthy and Prince (1999), it is further argued that ANCHORING should subsume ALIGNMENT in general on the account that edges of constituents can in effect be matched by the correspondence of segments standing at edges, using the ANCHORING format. It is in this extended sense that ANCHORING is being deployed here.

- (37) (i) ANCHOR-IO  
The edge in the input has a correspondent in the output.
- (ii) ANCHOR-OI  
The edge in the output has a correspondent in the input.

The edges in the input refer to the boundary of the grammatical constituent while that in the output refers to the prosodic foot boundary<sup>18</sup>. The two constraints respectively guard against the deletion and insertion of edges and together they are responsible for the boundary matching between the grammatical and the prosodic structures. Below we use ANCHOR when the two ANCHORING sub-constraints are referred to collectively.

As to the evaluation of ANCHOR, it is important to bear in mind that the edge matching is in effect achieved via the correspondence of the segments standing at the designated edges. Therefore, what matters is the presence or absence of the edges, i.e. brackets; neither the number of layers nor the direction of the brackets is relevant to the evaluation. Hence, ANCHOR-IO is satisfied as long as one or more (leftward and/or rightward) edges at a position in the input has a correspondent at the same position in the output, and violated when no such correspondence exists; *mutatis mutandis* for ANCHOR-OI. For example, given the input-output pair of S[[SS][SS]] and (SS)(S)(SS), ANCHOR-IO is violated once due to (and despite) the lack of an output correspondent for the *two* grammatical boundaries between the first and the second syllables in the input and ANCHOR-OI is also violated only once due to (and despite) the lack of input correspondent for the *two* prosodic boundaries between the second and the third syllables in the output.

Having discussed the formulation and evaluation of ANCHOR, we now consider its ranking with the other constraints. Reconsider Tableau (36), which we juxtapose below with a separate column indicating the violation of ANCHOR-IO and ANCHOR-OI by the same (sub-)set of candidates.

- (38) (i)

S[[SS][SS]]	BINMAX	BINMIN	*IP-FINAL-MONOF	ALIGNR (Ft, IP)
☛ (SS)(S)(SS)		*		5
(SS)(SS)(S)		*	*!	4
☛ (S)(SS)(SS)		*		6!
(SS)(SSS)	*!			3
(SSS)(SS)	*!			2

- (ii)

S[[SS][SS]]	ANCHOR-IO	ANCHOR-OI
(SS)(S)(SS)	*	*
(SS)(SS)(S)	**	**
(S)(SS)(SS)		
(SS)(SSS)	**	*
(SSS)(SS)	*	

<sup>18</sup> This formulation also allows us to bypass the need to specify the status of the grammatical constituents involved, which, as argued in Appendix II, are of a rather disparate nature and defy a unified characterization.

It is clear from (ii) that the true winner (S)(SS)(SS) incurs no violation of the two ANCHOR constraints while the undesired winner (SS)(S)(SS) incurs one violation of each of them. This provides the crucial ranking argument for ANCHOR >> ALIGNR (FT, IP), as indicated below. Note that we assume at this point that the two ANCHORING sub-constraints stay together in the hierarchy and are unranked with each other unless evidence arises calling for their ranking apart.

(39)

S[[SS][SS]]	ANCHOR-IO	ANCHOR-OI	ALIGNR(FT, IP)
(SS)(S)(SS)	*!	*	5
𠄎 (S)(SS)(SS)			6

As to the ranking between ANCHOR and the other constraints, the scansion of lines of S[[SS][SS]] falls short of providing crucial evidence: the optimal scansion satisfies ANCHOR, and the suboptimal forms incur various numbers of violations of ANCHOR. Crucial evidence for the ranking comes from the scansion of other coding types. Specifically, the scansion of lines of the structure [[SS][SS]]S (cf. (30) and (31)) provides evidence for the ranking of ANCHOR with BINMAX and \*IP-FINAL-MONOFT: the scansion fully mapping the input structure, i.e. satisfying ANCHOR, loses to that satisfying the output-oriented constraints \*IP-FINAL-MONOFT and BINMAX. Thus, both \*IP-FINAL-MONOFT and BINMAX must dominate ANCHOR. The ranking argument is:

(40) \*IP-FINAL-MONOFT &gt;&gt; ANCHOR

[[SS][SS]]S	*IP-FINAL-MONOFT	ANCHOR-IO	ANCHOR-OI
𠄎 (SS)(S)(SS)		*	*
(SS)(SS)(S)	*!		

(41) BINMAX &gt;&gt; ANCHOR

[[SS][SS]]S	BINMAX	ANCHOR-IO	ANCHOR-OI
𠄎 (SS)(S)(SS)		*	*
(SS)(SS)	*!	*	

As for the ranking between BINMIN and ANCHOR, the evidence comes from verse lines with an even number of syllables, since the potential scansions for those with an odd number of syllables will always contain one monosyllabic foot, thus incurring one violation of BINMIN (unless they contain a trisyllabic foot which violates the highly ranked BINMAX). Specifically, consider lines of the grammatical structure S[S[SS]] which are scanned as (SS)(SS). For example:

(42)      *wo3 [cu2 [dong1 shan1]]* → (*wo3 cu2*) (*dong1 shan1*)  
           I    go    east    mountain  
           ‘I go to the mountain on the east’.

This provides evidence for BINMIN to be ranked higher than ANCHOR:

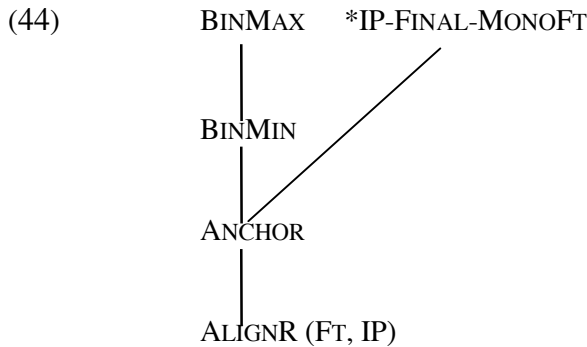


(43) BINMIN >> ANCHOR

S[S[SS]]	BINMIN	ANCHOR-IO	ANCHOR-OI
☞ (SS)(SS)		*	
(S)(S)(SS)	*!*		

The reasoning can also be conducted along this line: if ANCHOR >> BINMIN, then (SS)(SS) would lose and (S)(S)(SS) would win. However, this is empirically not true, hence BINMIN must dominate ANCHOR. In addition, since we have argued for the ranking ANCHOR >> ALIGNR (FT, IP) (cf. (39)), by transitivity, we get BINMIN >> ALIGNR (FT, IP).

To summarize, based on the scansion of the 5-syll *Shijing* lines supplemented by that of the 4-syll ones, we have introduced ANCHOR and ranked it accordingly. The emergent sub-grammar at this stage is presented in the Hasse graph below:



This constraint hierarchy, as it stands now, can adequately account for the scansion of all 5-syll lines, with only one exception, namely, the lines of the structure [SS]S[SS] with a line-medial interjection. This is illustrated below:

(45) [he2 shang4] hu1 [ao2 xiang2] → (he2) (shang4 hu1) (ao2 xiang2)  
 river over interj fly fly  
 ‘(The arrows) fly over the river’

Given the constraint hierarchy reached so far, the optimal scansion is (SS)(S)(SS) rather than the empirically attested (S)(SS)(SS):

(46)

[SS]S[SS]	BINMAX	BINMIN	*IP-FINAL-MONOFT	ANCHOR-IO	ANCHOR-OI	ALIGNR (FT, IP)
☞ (SS)(S)(SS)		*				5
☞ (S)(SS)(SS)		*		*!	*	6
(SS)(SS)(S)		*	*!			4
(SSS)(SS)	*!					2
(SS)(SSS)	*!					3

At first it might appear puzzling why (SS)(S)(SS) fails to be the empirically attested scansion: it outperforms (S)(SS)(SS) by better satisfying both ANCHOR and ALIGNR (FT, IP). What constraint could be motivated to salvage (S)(SS)(SS) over (SS)(S)(SS)? The answer comes from the peculiarity exhibited in the parsing of the interjection. In the optimal scansion de facto, the interjection, which is the third syllable from the left, is parsed as the second syllable of a disyllabic foot. In contrast, in the unwanted winner, the interjection constitutes a monosyllabic foot on its own.

This difference in parsing the interjection syllable between the actual and the unwanted winners suggests that it is preferable to parse an interjection syllable as the second syllable of a disyllabic foot rather than as a monosyllabic foot. Furthermore, it is noteworthy that an interjection syllable cannot occupy the first position of the disyllabic foot: a parsing of the present line into (SS)(SS)(S) in which the second foot starts with an interjection is unquestionably ill-formed (though it happens that in this case, the parsing is suboptimal also on account of its violation of the highly ranked \*IP-FINAL-MONOFT due to the presence of the line-final monosyllabic foot). Therefore, to sum up, suppose we use  $S_1$  to stand for the interjection syllable and S for other syllables (in particular lexical categories), the well-formedness pattern of prosodic feet containing  $S_1$  is as follows:

(47)

Foot type	Well-formedness
( $SS_1$ )	Good
( $S_1$ )	Bad
( $S_1S$ )	Bad

With this peculiarity of the interjection syllable parsing in mind, we re-consider the two candidates (SS)(S)(SS) and (S)(SS)(SS) at issue in (46) where the third syllable is an interjection. It is obvious that the former has a monosyllabic foot constituted by the interjection, which, as shown above, is ill-formed. By comparison, in the latter the interjection is parsed as the second syllable of a disyllabic foot, which is well-formed, although the cost of this parsing is that the edges in the input and the output no longer match each other, namely, ANCHOR is violated.

This suggests two things. First, a new constraint is needed to evaluate the well-formedness of feet containing an interjection syllable. In this light, we propose the constraint GOODFTINTERJ which serves exactly the same purpose as (47) in constraining the shape of the foot containing an interjection syllable. It follows exactly the same well-formedness pattern of such feet delimited in (47) and is formulated as:

(48)

GOODFTINTERJ

The interjection syllable can only be legitimately parsed as the second syllable of a disyllabic foot.

Second, a trade-off relation exists between this new constraint and ANCHOR, or, in other words, this new constraint conflicts with ANCHOR. The ranking argument is provided by the pair of candidates (SS)(S)(SS) and (S)(SS)(SS): (SS)(S)(SS) satisfies ANCHOR, but violates GOODFTINTERJ, and is suboptimal whereas (S)(SS)(SS) satisfies GOODFTINTERJ, but violates ANCHOR, and is optimal. Hence:

## (49) GOODFTINTERJ &gt;&gt; ANCHOR

[SS]S[SS]	GOODFTINTERJ	ANCHOR-IO	ANCHOR-OI
(SS)(S)(SS)	*!		
☞ (S)(SS)(SS)		*	*

Now, we consider the ranking of GOODFTINTERJ with the other constraints. First, since as argued above, ANCHOR >> ALIGNR (Ft, IP), by transitivity, we also have GOODFTINTERJ >> ALIGNR (Ft, IP). Second, GOODFTINTERJ does not conflict with BINMAX; indeed like BINMAX, GOODFTINTERJ is necessarily undominated in the hierarchy. The relevant pair of candidates, (SSS)(SS) and (SS)(S)(SS), which violate BINMAX and GOODFTINTERJ respectively, are both suboptimal, whereas the optimal candidate (S)(SS)(SS) violates neither of the two constraints. This is indicated below:

## (50)

[SS]S[SS]	BINMAX	GOODFTINTERJ
(SSS)(SS)	*!	
(SS)(S)(SS)		*!
☞ (S)(SS)(SS)		

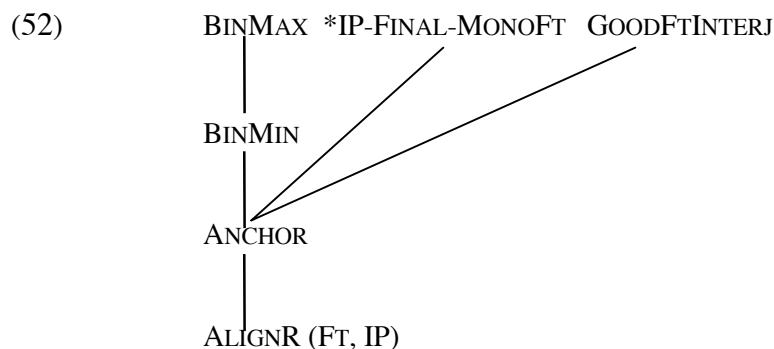
Thirdly, by the same token, GOODFTINTERJ does not conflict with \*IP-FINAL-MONOFT: the relevant candidate pair, (SS)(S)(SS) and (SS)(SS)(S), which respectively violate these two constraints, are both suboptimal whereas the optimal candidate (S)(SS)(SS) violates neither of the two constraints. This non-ranking between these two constraints is illustrated below:

## (51)

[SS]S[SS]	GOODFTINTERJ	*IP-FINAL-MONOFT
(SS)(S)(SS)	*!	
(SS)(SS)(S)	*	*!
☞ (S)(SS)(SS)		

Fourthly, as to the ranking between GOODFTINTERJ and BINMIN, the 5-syll lines provide no evidence. The reason is that BINMIN is not discriminating enough for the two relevant competitors (SS)(S)(SS) and (S)(SS)(SS): both violate BINMIN. In fact, for verse lines containing an odd number of syllables, their scansion inevitably contain either a monosyllabic foot or a trisyllabic one under the premise that all the syllables are parsed. Therefore, all the candidates violate either BINMAX or BINMIN (or both, as in (S)(SSS)(S)). Given the ranking BINMAX >> BINMIN reached above, those forms violating BINMAX are suboptimal whilst of the forms violating BINMIN, the one best satisfying the other constraints in the hierarchy is optimal. Therefore, BINMIN belongs to the kind of constraints that is violated by the winner and some of the losers alike, and at least for 5-syll lines, its ranking with GOODFTINTERJ is immaterial.

As an interim summary, the emergent sub-grammar now is:



So far, however, we have only mentioned that the newly introduced GOODFTINTERJ constraints the well-formedness of prosodic feet containing an interjection syllable; what remains to be discussed is why the parsing of interjection syllables is subject to this constraint. Below we argue that this parsing pattern of interjection syllables, i.e., the constraint GOODFTINTERJ per se, is attributable to the phonological property of such syllables.

### 2.3.4.1 Phonological property of interjection syllables

Reconsider the pattern in (47), and it is immediately notable that the only legitimate position an interjection syllable can occur in is the second position of a disyllabic prosodic foot. In contrast, a monosyllabic foot constituted by an interjection syllable and a disyllabic foot with an interjection syllable occupying the first position are both ill-formed. We argue that this parsing pattern is attributable to the phonological representation of interjection syllables, namely, they are underlyingly light, i.e. monomoraic. Given the argument that the only good foot type in Chinese is trochee (at both the syllabic and the moraic levels) (cf. Duanmu 1999, 2000; also see (15) above), the only good foot containing an interjection syllable is a disyllabic one where the interjection syllable, being monomoraic, occurs in the second position, i.e. the weak position of a disyllabic trochee. By comparison, a monosyllabic foot solely formed by an interjection syllable, hence also a monomoraic foot, and a disyllabic foot with an interjection syllable occupying the first position, namely, the head of the trochee, are both ill-formed. This is illustrated below together with comments on the well-formedness of each foot type (for clarity sake, we give the moraic and syllabic representations as well as syllable quantification structure next to the pattern of well-formedness presented in (47)):

(53)

Foot type	Prosodic representation	Well-formedness	Comments
(SS <sub>i</sub> )	x (S S) (μμ)(μ) x H L	Good	Syllabic trochee and moraic trochee; heavy syllable carrying stress

(S <sub>l</sub> )	x (S) (μ) L	Bad	Light (monomoraic) syllable unable to carry stress <sup>19</sup>
(S <sub>l</sub> S)	x (S S) (μ) (μμ) x L H	Bad	Light syllable unable to carry stress

This way, the constraint GOODFTINTERJ may be reformulated in terms of the conjunction of well-established universal constraints such as RHTYPE=T (Feet have initial prominence, namely, are trochaic; cf. Kager 1999:172) and STRESS-TO-WEIGHT (If stressed, then heavy; cf. Riad 1992; Myers 1997). However, for simplicity sake, we continue to use GOODFTINTERJ portmanteau in our constraint hierarchy.

Two more points about the prosodic feet containing an interjection syllable call for discussion. First, so far we have only focused on the well-formedness of monosyllabic and disyllabic feet containing interjection syllables, and said nothing about trisyllabic feet. Although trisyllabic feet are expelled as illicit in verse scansion by the highly ranked BINMAX, it is still necessary for the candidates containing such feet (albeit suboptimal) to undergo evaluation by GOODFTINTERJ. As argued above, on the one hand, interjection syllables are underlyingly monomoraic and thus unable to bear stress, and on the other hand, it has been argued that trisyllabic feet in Chinese result from the merging between a disyllabic foot and its neighboring monosyllabic foot, and as such, have the stress pattern of S(trong)-W(eak)-W(eak) (Duanmu 2000: 188). Therefore, an interjection syllable can legitimately occupy the medial or the final position of a trisyllabic foot, and as far as the parsing of the interjection syllable is concerned, (SS<sub>l</sub>S) or (SSS<sub>l</sub>) are both well-formed<sup>20</sup>.

Second, it needs to be pointed out that the phonological property of interjection syllables argued for here is true with all interjections occurring in the *Shijing* genre, including the interjection 'xi'. The reason that 'xi' is singled out is because it also occurs in the Chuci genre, but behaves differently there. Indeed, as we are going to argue in the next chapter, the versatility 'xi' exhibited in different contexts, which refer to both the specific phonological context and the broad context of the genre in which it occurs, is attributable to the presence of an empty mora in its underlying representation.

<sup>19</sup> An alternative way of accounting for the ill-formedness of the monosyllabic, monomoraic foot constituted by the interjection syllable alone is that it fails to meet the BINARITY requirement, which is proposed in the literature as a general foot well-formedness constraint (e.g. Kager 1999), at either the syllabic or the moraic level.

<sup>20</sup> Indeed, this is true in prose scansion where trisyllabic feet are allowed, for example:

(i) fa1 hu1 qing2 → (fa1 hu1 qing2)  
arise interj passion  
'Arise (out of) passion'

(ii) [he2 shang4] hu1 [ao2xiang2] → (he2 shang4 hu1) (ao2 xiang2)  
river above interj fly fly  
'(The arrows) fly over the river'.

### 2.3.5 More on the sub-grammar: the scansion of 6-syll lines

Moving on to 6-syll lines, we note that they are all optimally scanned as (SS)(SS)(SS) irrespective of the grammatical structures. Some examples are:

- (54) (i)  $[[can1 zhi2] [zi3 [zhi1 qv4]]] xi1 \rightarrow (can1 zhi2) (zi3 zhi1) (qie4 xi1)$   
 grasp hold you 's sleeve interj  
 'Ah, (I) hold you by your sleeve'
- (ii)  $[wu3 yue4] [si1 zhong3] [dong4 gu3] \rightarrow (wu3 yue4) (si1 zhong3) (dong4 gu3)$   
 fifth month this cricket move leg  
 'In the fifth month, the crickets move their legs'
- (iii)  $[zheng4 shi4] [yi1 [[pi2 yi4] wo3]] \rightarrow (zheng4 shi4) (yi1 pi2) (yi4 wo3)$   
 political affairs all give offer me  
 'The political affairs are all handed to me'
- (iv)  $[liu4 yue4] [shi2 [yu4 [ji2 yu4]]] \rightarrow (liu4 yue4) (shi2 yu4) (ji2 yu4)$   
 sixth month eat date and plum  
 'In the sixth month, (we) eat dates and plums'
- (v)  $[yu4 [ren2 [zhi1 [jian1 nan2]]]] yi3 \rightarrow (yu4 ren2) (zhi1 jian1) (nan2 yi3)$   
 suffer human 's hardship difficulty interj  
 'Ah, suffering from human's hardships and difficulties'
- (vi)  $dai4 [ji2 [gong1 zi3] [tong2 gui1]] \rightarrow (dai4 ji2) (gong1 zi3) (tong2 gui1)$   
 wait till lord gentleman together return  
 '(I) wait till the gentleman can come back together (with me)'
- (vii)  $wo1 [gu1 [zhuo2 [bi3 [jin1 zun1]]]] \rightarrow (wo1 gu1) (zhuo2 bi3) (jin1 zun1)$   
 I temporarily fill that golden wineglass  
 'I temporarily fill that golden wineglass (with wine).'

Recall it has been shown that for 4-syll lines the optimal scansion is invariably (SS)(SS) irrespective of the input structure. Similar things can be suggested for 6-syll lines: (SS)(SS)(SS) would, given the current sub-grammar, always emerge as optimal. This is true with both lines containing no interjection syllables (e.g. (ii), (iii), (iv), (vi), and (vii) above) and those containing such syllables (e.g. (i) and (v)), as illustrated below:

(55) (i)

[[SS][SS]][SS]	BINMAX	BINMIN	*IP-FINAL-MONOFT	GOODFT INTERJ	ANCHOR-IO	ANCHOR-OI	ALIGNR (FT, IP)
☞ (SS)(SS)(SS)							6
(S)(SS)(S)(SS)		*!*			*	**	10
(S)(SS)(SS)(S)		*!*	*		**	***	9
(SSS)(S)(SS)	*!	*			*	*	5
(SS)(S)(SSS)	*!	*			*	*	7

(ii)

[[SS][S[SS]S]	BINMAX	BINMIN	*IP-FINAL-MONOFt	GOODFt INTERJ	ANCHOR-IO	ANCHOR-OI	ALIGNR (Ft, IP)
☞ (SS)(SS)(SS)					**	*	6
(SS)(S)(SS)(S)		*!*	*				9
(S)(SS)(SS)(S)		*!*	*		*	*	9
(SSS)(S)(SS)	*!	*			*	*	5
(SS)(S)(SSS)	*!	*			*	*	7

(iii)

[[SS][S[SS]]S <sub>I</sub>	BINMAX	BINMIN	*IP-FINAL-MONOFt	GOODFt INTERJ	ANCHOR-IO	ANCHOR-OI	ALIGNR (Ft, IP)
☞ (SS)(SS)(SS) <sub>I</sub>					**	*	6
(SS)(S)(SS)(S) <sub>I</sub>		*!*	*	*			9
(S)(SS)(SS)(S) <sub>I</sub>		*!*	*	*	*	*	9
(SSS)(S)(SS) <sub>I</sub>	*!	*			*	*	5
(SS)(S)(SSS) <sub>I</sub>	*!	*			*	*	7

In all cases, the scansion (SS)(SS)(SS) satisfies the highly ranked constraints in the constraint hierarchy which are all markedness ones, namely, BINMAX, BINMIN, \*IP-FINAL-MONOFt and GOODFtINTERJ. When the line contains no interjections, GOODFtINTERJ is vacuously satisfied; in those lines containing interjections, such syllables always occur line-finally, and as \*IP-FINAL-MONOFt forbids line-final monosyllabic feet, the interjection syllables in the optimal scansion will always adjoin to the preceding syllable to parse into a disyllabic foot, thus satisfying GOODFtINTERJ<sup>21</sup>.

As the sub-grammar stands now, these constraints all dominate ANCHOR, the only constraint that has to refer back to the input. Therefore, (SS)(SS)(SS), by best satisfying all the markedness constraints, outperforms all the other competing candidates which would each incur at least one violation of the highly ranked markedness constraints even though they might better satisfy ANCHOR, whose role in evaluating the candidates is suppressed due to its relatively low ranking.

Thus, even though no new constraints or rankings are motivated, the scansion of 6-syll *Shijing* lines demonstrates the sufficiency of the current sub-grammar.

### 2.3.6 ANCHOR-I<sub>SB</sub>O<sub>PHP</sub> and \*PHP-FINAL-MONOFt: evidence for hierarchicality from 7-syll lines

We now move to 7-syll lines. Some grammatical structures present no challenge to the constraint hierarchy reached so far, for example, [S[SS]][S[S[SS]]] and [S[S[S[S[SS]]]]]S:

- (56) (i) [*san1 [zhi1 ri4]*] [*na4 [yu2 [ling2yin1]]*]  
 third prt day carry to ice shelter  
 ‘In days of the third (month), (we) carry (the ice) to the ice-houses’

<sup>21</sup> As we are going to see in Chapter 8, this distribution of interjection syllables proves significant.

- (ii) [huan2 [yu2 [shou4 [zi3 [zhi1 can4]]]] xi1  
 return I present you particle magnificence interj  
 ‘Ah, (when you) return, I will present (you with) your magnificence’.

→ (huan2 yu2) (shou4 zi3) (zhi1)(can4 xi1)

In both cases, the optimal scansions (SS)(S)(SS)(SS) and (SS)(SS)(S)(SS) are fully predicted by the constraint ranking so far. This is illustrated below:

(57) (i)

[S[SS]][S[S[SS]]]	BINMAX	BINMIN	GOODFT INTERJ	*IP-FINAL- MONOFT	ANCHOR- IO	ANCHOR- OI	ALIGNR (Ft, IP)
☞ (S)(SS)(SS)(SS)		*	*		*		12
(SS)(SS)(S)(SS)		*	*		**!	*	10
(SS)(SSS)(SS)	*!		*		***	*	7
(SSS)(SS)(SS)	*!		*		**!		6

(ii)

[S[S[S[S[SS]]]]]S	BINMAX	BINMIN	GOODFT INTERJ	*IP-FINAL- MONOFT	ANCHOR- IO	ANCHOR- OI	ALIGNR (Ft, IP)
☞ (SS)(SS)(S)(SS)		*			***	*	10
(SS)(S)(SS)(SS)		*			***	*	11!
(S)(SS)(SS)(SS)		*			***	*	12!
(SS)(SSS)(SS)	*!				*		7
(SS)(SS)(SS)(S)		*	*!	*	**		9

But consider lines of the structure [[SS]S][S[S[SS]]] which are, unlike the two examples in (56), optimally scanned as (S)(SS)(SS)(SS):

- (58) [[zhi1 wo3] zhe3] [wei4 [wo3 [xin1 you1]]]  
 know me person say I heart worry  
 ‘Those who know me say I am (just) worrying’.

→ (zhi1) (wo3 zhe3) (wei4 wo3) (xin1 you1)

This proves to be problematic for the current sub-grammar:

(59)

[[SS]S][S[S[SS]]]	BINMAX	BINMIN	GOODFT INTERJ	*IP- FINAL- MONOFT	ANCHOR- IO	ANCHOR- OI	ALIGNR (Ft, IP)
a. ☞ (S)(SS)(SS)(SS)		*			**!	*	12
b. ☞ (SS)(SS)(S)(SS)		*			*		10
c. (SS)(S)(SS)(SS)		*			*		11!
d. (SS)(SSS)(SS)	*!				**		7
e. (SSS)(SS)(SS)	*!				**	*	6

As it stands now, (SS)(SS)(S)(SS) comes out as the winner, whereas empirically the winning candidate is (S)(SS)(SS)(SS). Actually, in terms of constraint violation, the desired winner (a) fares worse than both the unwanted winner (b) and another



suboptimal form (c). Specifically, (a) violates ANCHOR-OI whereas neither (b) nor (c) does and (a) scores the worst with regard to ALIGNR (FT, IP) of the three, which indicates that a mere re-ranking of the current constraints, or ranking of those hitherto unranked constraints, will not solve the problem. Obviously, a new constraint is needed to give credit to some feature of the desired winner which so far goes ignored by the current constraints.

Compare these three competitors and we note that of them, the desired optimal form (S)(SS)(SS)(SS) incurs the most violations of ANCHOR and ALIGNR (FT, IP); but the fact that it still is the real winner instructs us that its more violations of ANCHOR and ALIGNR (FT, IP) must be for a good reason. Specifically along the line of OT-thinking, the only good reason to violate some constraints is to satisfy some more important, i.e. higher-ranking one(s). In other words, in a constraint hierarchy, there always exists a trade-off between the violation of some constraints and the satisfaction of some other ones; which candidate eventually comes out as the winner out of the potentially infinite number of candidates produced by GEN is contingent on the net outcome of the trade-off. Put in more transparent OT parlance, the winner is the most harmonious candidate resulting from the resolution of constraint conflict under the given ranking hierarchy.

Now that the rationale behind the reasoning is clarified, the question is what this forthcoming higher-ranking constraint is. Observe carefully what is unique about the desired optimal form (S)(SS)(SS)(SS) against the other two competitors, with reference to the input structure, and *not* in terms of the constraints already existent in the hierarchy (which we already concluded to be inadequate). Two features are noteworthy. First, one of the most striking differences between (S)(SS)(SS)(SS), the desired winner, and (SS)(SS)(S)(SS), the unwanted winner, is whether the strongest grammatical boundary within the verse line (which occurs between the third and fourth syllables and is represented as a pair of reverse brackets in its grammatical structure) has a correspondent in the output<sup>22</sup>. For the former, it does correspond to the right boundary of the second prosodic foot (which is, more precisely, the PhP boundary, to be argued below), whereas for the latter, it has no correspondent in the output.

The second feature bears on the comparison between the suboptimal candidate (c), (SS)(S)(SS)(SS), and the desired winner (S)(SS)(SS)(SS). They score equal in that for both of them the strongest boundary in the input has an output correspondent, but differ in the position of the monosyllabic foot. The desired winner (S)(SS)(SS)(SS) avoids having a monosyllabic foot as the second foot from the left boundary of the IP, which is the case with (SS)(S)(SS)(SS). This avoidance is achieved by parsing the first syllable into a monosyllabic foot, thus at the cost of better satisfaction of ANCHOR.

### 2.3.6.1 Introduction of hierarchicality into the sub-grammar

We argue that these two features highlight the hierarchicality of both the input and the output structures and reveal the need for it to be built into the sub-grammar. Up to

---

<sup>22</sup> As is shown in Appendix II, the strongest grammatical boundary in a verse line, which is coded as 4, is of a disparate nature and no straightforward, one-to-one correspondence exists between such a boundary and any single syntactic constituency. More is to be said below.

now, we have been treating both structures, especially the output, as flat and linear. Specifically, in monitoring the boundary mapping between the input and the output, ANCHOR is concerned only with the lowest-level brackets in both structures. Of all the constraints motivated so far, only \*IP-FINAL-MONOFT and ALIGNR(Ft, IP) deploy the higher-level prosodic unit of IP. Other than this, neither the depth of the grammatical boundaries in the input, nor the prosodic units at the various levels of the prosodic hierarchy between the foot and the IP are utilized by the constraints<sup>23</sup>. Below we elaborate on the construct of hierarchicality and its integration in the sub-grammar by respectively analyzing the two features mentioned above.

The first feature, namely, the presence of an output correspondent for the strongest grammatical boundary in the optimal scansion, in fact reflects a requirement similar to ANCHOR-IO, only operative at a higher level of the structural hierarchy. It is important that the strongest boundary here must be a line-medial one in order to become relevant in such an endeavor; obviously, such a boundary is present only in a bi-directionally branching grammatical structure. In terms of bracketing, a line-medial strongest boundary is translatable into a pair of back-to-back brackets. By comparison, line-initial or penultimate strongest boundaries are not relevant here as they respectively correspond to the right-branching or left-branching structure. For example, both lines in (54) (vi) and (vii) have a right-branching structure, while those in (54) (i) and (v) are essentially left-branching with a line-final interjection<sup>24</sup>.

It needs to be pointed out that this strongest grammatical boundary may vary considerably in terms of the corresponding syntactic constituency. For example, it could be the boundary of NP (as in (58)), or VP, as in

- (60)        *qian1 shang3 she4 qin2*  
              lift skirt cross river name  
              ‘(I) lift my skirt and cross the river *Qin*.’

or a clause boundary when the line contains two coordinated clauses:

- (61)        *ren2 she4 ang2 pi3*  
              others cross I not  
              ‘Others cross the river, (but) I do not.’

However, even though the strongest boundary may be generalized as an XP boundary, it is not necessarily true the other way around. Not every maximal projection boundary is the strongest boundary. For example, in

<sup>23</sup> That the prosodic structure is hierarchical is evident from the fact that in an IP containing more than one foot, the degree of lengthening is not all the same at every foot boundary. Rather, a certain foot boundary in the middle of the IP is typically characterized by a bigger prosodic break that is empirically perceptible as a greater lengthening (or, for that matter, longer pause) than that occurring at other foot boundaries (with the exception of the line-final one of course). Given the established correlation between the degree of boundary lengthening (or duration of pause) and the level of prosodic units in the prosodic hierarchy (Beckman and Edwards 1990), this clearly suggests an intermediate prosodic level between foot and IP.

<sup>24</sup> Alternatively, one can analyze the line-final interjection as a sentence level functional category. Under this analysis, the line remains a left-branching structure.

- (62) [ying1 ti2]<sub>IP</sub> [yan4 yu3]<sub>IP</sub>[bao4 [xin1 nian2]<sub>NP</sub>]<sub>VP</sub>  
 nightingale sing swallow speak report new year  
 ‘The nightingales sing and the swallow speak, reporting the new year’

four maximal projections are identified, as indicated here, but only the boundary following the second IP is the strongest. A verse line typically has more than one maximal projection but only the boundary of one of them qualifies as the strongest boundary. Indeed, neither is there any single syntactic constituency whose boundary is always the strongest within a line nor can the strongest grammatical boundary be exclusively reduced to any single syntactic boundary. In view of this, we suggest that, as an analytical expedient, the requirement that the strongest grammatical boundary in the input have a correspondent in the output be straightforwardly expressed as the following constraint which directly borrows the notion of ‘strongest boundary’ and refers to it as SB:

- (63) ANCHOR-I<sub>SB</sub>O  
 The strongest grammatical boundary within the line has a correspondent in the output.

Note that this is only a temporary formulation where the output correspondent of the SB remains undefined. As is to be argued below, this correspondent is Phonological Phrase (PhP), which is a prosodic boundary bigger than the foot boundary.

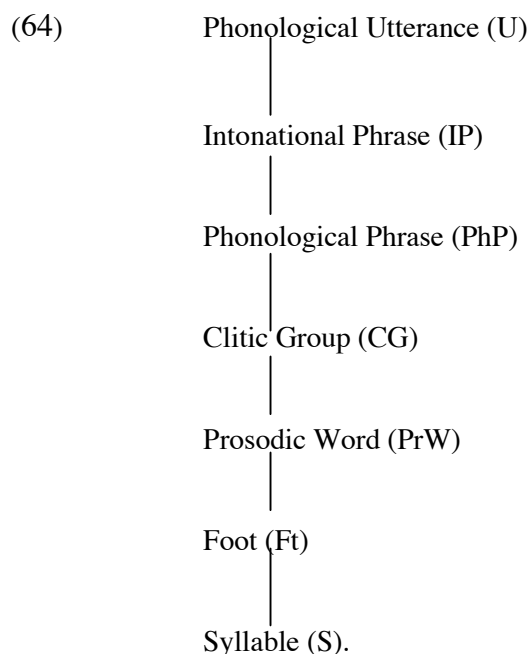
As argued earlier, in gauging the boundary matching between the input and the output, ANCHOR is necessarily bi-directional. In this light, this new constraint, concerned with the preservation of a specific input boundary in the output, is, precisely speaking, the higher-level counterpart of ANCHOR-IO.

We now turn to the second feature mentioned above about the optimal scansion of (58), which is concerned with the position of the monosyllabic foot in the IP. One of the constraints proposed hitherto is already involved with the position of such a foot, i.e. \*IP-FINAL-MONOFT; nonetheless, as is shown in (59), this constraint is not sufficient in winnowing out the suboptimal forms. In the optimal scansion in (59), the monosyllabic foot not only avoids the IP-final position but also moves away from the second position from the left even though this results in more violations of ANCHOR. The crux here is, we propose, that the output is, rather than a linear sequence of prosodic feet, a hierarchical prosodic structure. Accordingly, the ‘second position’ which the monosyllabic foot shuns is in fact the final position of some higher-level prosodic unit, which is, as to be argued below, PhP. To pinpoint this higher-level prosodic unit entails a discussion of the prosodic hierarchy of Chinese.

### 2.3.6.1.1 Prosodic hierarchy of Chinese

That the prosodic organization is hierarchical in nature is far from an original view: it has been firmly established on account of a rich body of phonological data (see for example, Selkirk 1980, 1986; Nespor and Vogel 1986; Hayes 1989; Inkelas and Zec 1990). What has also become indisputable is that the prosodic structure interacts with the syntactic structure in a non-trivial manner. Nonetheless, the various proposals on prosodic phonology do not totally agree on either the exact number of levels of the hierarchy or the characterization of prosodic constituents at different levels. The one

put forward in Nespor and Vogel (1986) has become more or less conventional (and which is also adopted in Hayes' (1989) study of English metrics):

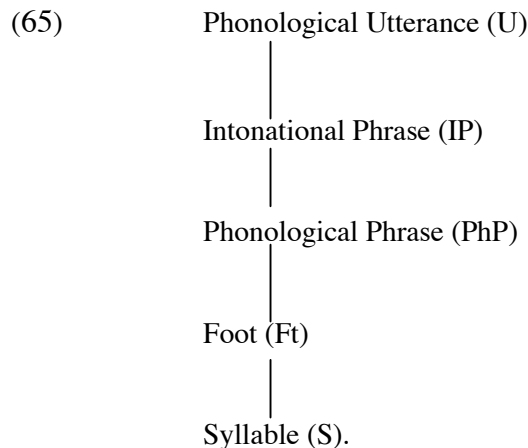


In the current context of Chinese, the prosodic hierarchy, and in particular, its interaction with syntax, has also been a topic of considerable controversy; a wide range of proposals have been put forward, largely on the basis of tone sandhi data (for example, Cheng 1987; Shih 1986; Hung 1987). For practical considerations, these individual proposals are not presented here; instead, we will only mention Chen (1996), which offers a critical overview of the above-cited various proposals and concludes that the prosodic organization in Chinese is ‘of considerable typological significance’ (p. 518) in that the prosodic constituency of Chinese does not fit straightforwardly into any of the prosodic hierarchies proposed so far.

Instead, Chen (Ibid.) proposes a prosodic hierarchy for Chinese that merely consists of two levels: IP and what he refers to as MRU (Minimal Rhythmic Unit). What deserves attention is the highly fluid nature of MRU – basically it could be anything. Indeed, according to Chen, there is neither a uniform form of MRU nor a straightforward way of determining what form MRU may take; rather MRU is determined in an OT fashion as the optimal candidate resulting from the dynamic interaction of a range of constraints on a case-specific basis for every given input. As such, MRU ‘stands apart from the conventional prosodic hierarchy’, ‘is basically a device to group syllables of a wide variety of grammatical ranks and status into rhythmic units, and as such, ‘constitutes a prosodic unit *sui generis*, off-scale, and *hors-series*’ (Ibid.: 518).

We should mention that this proposal of Chen’s is reached on the basis of a wide range of Mandarin tone sandhi data that proves intractable for a prosodic hierarchy along the conventional line as in (64), together with the assumption that the tone sandhi domain is constituted by the prosodic constituents at certain levels. However, the current context of verse scansion is one with a slower speech rate and more rigid prosodic domains, to which the characterization of prosodic constituency is most

sensitive (cf. Jun 1996; Shih 1997). Our position here is that in this current context, the prosodic constituency of Chinese can indeed be fit into a conventional prosodic hierarchy as (64); however, the prosodic hierarchy of Chinese is an impoverished rather than a full-fledged one<sup>25</sup>. Specifically, we adopt the prosodic hierarchy argued for in Shih (1986, 1997), namely, a five-level hierarchical prosodic structure of Chinese comprising of Utterance (U) as the top node, and below that, Intonational Phrase (IP), Phonological Phrase (PhP), prosodic foot (F) and syllable (S). This prosodic hierarchy is illustrated below:



Compare this with (64) and we note that the levels of Clitic Group and Prosodic Word are missing<sup>26</sup>. It is in this sense that we suggest the prosodic hierarchy of Chinese is an impoverished one. We leave open the issue of to what extent this prosodic hierarchy can be reconciled with Chen's two-level hierarchy.

<sup>25</sup> Indeed, the possibility that the full-fledged prosodic hierarchy including all the seven constituents presented in (64) might not be relevant to all languages has been well entertained. For example, Gvozdanovic (1986), in delimiting the higher-level prosodic domains, suggests that the distinction between the Phonological Word and the Phonological Phrase might not be relevant in all languages. Nespors and Vogel (1986, 1989) themselves pointed out that 'there is no *a priori* reason that the phonology of a given language must include all seven units' (1986: 11), for example, at least for some languages there might not be the need for the level of Clitic Group (1989: 113). Also see the next footnote.

<sup>26</sup> Arguably an alternative way to account for the absence of these two levels might be to suggest that they are isomorphous with first, each other, and second, either its upstairs neighbor - Phonological Phrase or its downstairs neighbour - Foot. For the first possible isomorphism, indeed, it has been lively debated whether the Clitic Group needs to be distinguished from the Prosodic Word. For example, Zec (1993) suggests that the Clitic Group is simply the Prosodic Word, as it is accessed in postlexical phonology, whilst Hayes (1989) puts forward arguments for Clitic Groups from English syllabification and argues for the distinction between Prosodic Word and Clitic Group on the basis of English metrics. For Chinese, it has indeed been argued that Prosodic Word is synonymous with Clitic Group; both are defined as a lexical host plus surrounding function words, which behave like prosodically dependent proclitics or enclitics (cf. Nespors and Vogel 1986:145f) (Chen 1996:542 in comparing MRU with the conventional prosodic units such as Prosodic Word and Clitic Group). As to the second possible isomorphism, given the range of characterizations proposed for Phonological Phrase for Chinese (e.g. Beattie 1985; Cheng 1987; Hsiao 1991), it seems more plausible to suggest that Prosodic Word (and thus Clitic Group) is isomorphous with its downstairs neighbor, i.e., Foot and that Phonological Phrase consists of one or more feet. Clearly, in effect this is identical to the prosodic hierarchy proposed by Shih (1997) presented here.

We now proceed to examine the five constituents in this hierarchy. As suggested earlier, a verse line is prosodically an IP, on the account that the whole line falls under a unified intonation contour. Since the line is selected as the domain of analysis for the present study, it follows that IP is the upper bound of the prosodic domain that is relevant<sup>27</sup>. At the lower end, the discussion hitherto has been centering around the parsing of syllables into feet. Indeed, so far we have built IP and Ft into the sub-grammar, so the only prosodic unit remaining to be explored is PhP.

### 2.3.6.1.1.1 *The circumscription of Phonological Phrase (PhP)*

In this section, we argue that PhP is defined, first, by the line-medial SB in a pre-emptive manner when such a boundary is present, and second, when such a boundary is absent, by three constraints whose interaction can be captured in an OT fashion. They are: BINARITY, EVENNESS and LONG-LAST. Below we discuss these constraints and the ranking between the latter three.

#### 2.3.6.1.1.1.1 PhP boundary as the output correspondent of strongest boundary

To begin with, the prosodic hierarchy of Chinese observes the Strict Layer Hypothesis (Selkirk 1984; Nespore and Vogel 1986; Hayes 1989). In Section 2.3.2, for example, we argued for the strong layering view that a PrW (or for that matter, PhP) only consists of one or more feet against the weak layering one which allows an unparsed syllable to be parsed as an immediate daughter of PrW (or given the impoverished prosodic hierarchy presented here, PhP). At higher levels, the Strict Layer Hypothesis holds as well: Shih (1990, 1997) suggests that PhP consists of one or more prosodic feet, whereas on the other hand, an IP consists of one or more PhP. Accordingly, a PhP boundary necessarily coincides with a Foot boundary.

Secondly, recall that it was pointed out earlier that the prosodic hierarchy is influenced by the syntactic structure in important ways. This is already evident in the delimitation of prosodic foot so far, and we suggest that it is also the case with the characterization of PhP. More specifically, the strongest syntactic boundary within the line, which itself corresponds to IP, corresponds to the biggest prosodic boundary within the IP, which is PhP given the prosodic hierarchy proposed above.

In this light, the constraint ANCHOR- $I_{SB}O$  proposed in (63) should, more precisely, be rendered as ANCHOR- $I_{SB}O_{PhP}$ . And the formulation is accordingly refined into:

- (66) ANCHOR- $I_{SB}O_{PhP}$   
The strongest grammatical boundary within the line corresponds to the PhP boundary.

Clearly, this constraint and ANCHOR-IO may be understood as members of a constraint family responsible for the boundary correspondence at various levels of the hierarchical structure between the input and the output.

Now in the light of the PhP boundary delimited as such, reconsider the two features presented earlier for the winner (a) in (59), which is more precisely represented as (S)(SS)|(SS)(SS). It should become evident that they are respectively concerned with

<sup>27</sup> The prosodic status of the verse units above the line level is in itself, though, a fascinating issue worth pursuing. For example, as Schlepp (1980) insightfully points out, the couplet, which is typically constituted by two neighboring lines, corresponds to a Phonological Utterance.

the presence of an output correspondent of a line-medial strongest grammatical boundary and the avoidance of monosyllabic foot PhP-finally. These two distinct features associated with the winner actually calls for the introduction of two constraints, ANCHOR- $I_{SB}O_{PhP}$  and \*PHP-FINAL-MONOFT. The former was already discussed above, and the latter is formulated in analogy to \*IP-FINAL-MONOFT and formulated as:

- (67) \*PHP-FINAL-MONOFT  
Do not place the monosyllabic foot at the final position of a PhP.

It imposes a similar ban on monosyllabic feet at the higher level of PhP and as such is also attributable to the well-attested NON-FINALITY constraint (Kager 1999).

The introduction of \*PHP-FINAL-MONOFT calls for a clear circumscription of the PhP boundary in all candidate forms. The discussion so far about the PhP boundary delimitation is only concerned with lines of a bi-directionally branching structure which contain line-medial SB's. The issue that follows naturally is how to determine the PhP boundary in those lines of a unidirectionally branching grammatical structure where line-medial SB's are absent. This scenario is discussed below before we move further to consider the ranking of the two constraints introduced above.

#### 2.3.6.1.1.1.2 BINARITY, LONG-LAST and EVENNESS: more on the delimitation of Phonological Phrase

In this section we propose three independently motivated general constraints responsible for the PhP boundary delimitation in lines which do not contain line-medial strongest grammatical boundaries, i.e. which have uni-directionally branching grammatical structure (translated into bracketing as either S[S... or ...S]S). They are BINARITY, LONG-LAST and EVENNESS, and their interaction can be captured in an OT fashion.

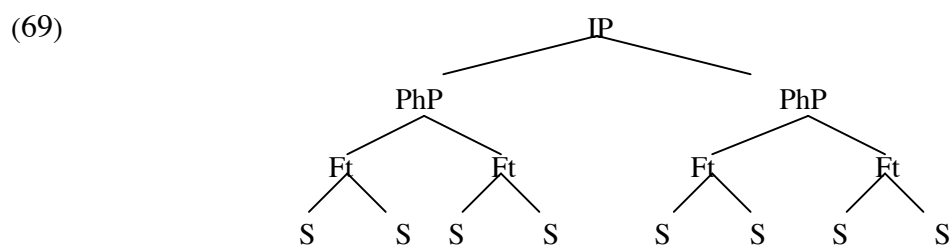
First, we suggest that in the current context of verse scansion, the hierarchical organization of prosodic structure within IP observes the BINARITY constraint. Although it has been argued that prosodic structures are n-ary branching in refutation of earlier claims that they are binary branching (Leben 1982; van der Hulst 1984; Nespor and Vogel 1986), it is also acknowledged that 'the question of binary vs. n-ary branching structures ... is likely to remain a controversial issue for years to come, with respect to phonology as well as to the other components of the grammar' (Nespor and Vogel 1986: 8). One of the main arguments cited therein for a preference of n-ary over binary branching is that the former calls for fewer intermediate levels and hence renders the structure flatter and simpler. As such, this simplicity argument entailed by an n-ary structure is more pertinent in cases where the higher-level prosodic unit such as an IP or Utterance is relatively long.

By comparison, in the verse context, this advantage offered by an n-ary prosodic structure is at best only marginal, as the verse lines are highly restricted in their length. Indeed, of the 3933 lines of verse constituting the present corpus, there are altogether only six 9-syll lines and no lines longer than that. Hence, while leaving open the issue of binarity versus n-arity regarding the prosodic structure in general, we suggest that as far as the current context of verse line scansion is concerned, to postulate an n-ary branching structure offers little particular advantage over a binary one, and that binarity is the basic structuring principle for the prosodic hierarchy (with

IP as the top-level constituent) (cf. similar arguments in Golston 1998). Furthermore, we suggest that the binarity requirement is violable and thus preferably couched into an OT constraint, in the same way that binarity at the foot level has been independently established a violable OT constraint FtBIN. This way, we wish to accommodate the possible cases where the prosodic structure is n-ary rather than binary. We refer to this constraint as BINARITY, formulated as follows:

- (68) BINARITY  
Prosodic structure is binary branching.

BINARITY requires that an IP consists of two PhP's, a PhP two feet and a foot two syllables<sup>28</sup>. Now consider a fully binary prosodic structure (with IP being the top node in the current context):



which is evidently only possible when the IP, i.e. the line here, contains an even number of feet, typically four<sup>29</sup>. This renders it particularly interesting to consider the grouping of feet into PhP's when the IP consists of an odd number of feet, which occurs when the line contains for example five, six, nine, or ten syllables<sup>30</sup>. In such cases, the division of the IP into PhP's necessarily results in the violation of BINARITY, which motivates the second constraint, namely, LONG-LAST, formulated in most general terms as:

- (70) LONG-LAST  
When two constituents in a domain are of different length, place the longer constituents at the right end of the domain.

This constraint is attributable to Hayes (2000) where it was originally proposed for the quatrain structure in English sung verse. Formulated in most general terms as 'in a sequence of groups of unequal length, the longest member should go last' (Ibid.), it expresses a preference for the placement of longer units at the right edge when a balanced, symmetrical grouping is impossible. An unbalanced structure with initial weight is considered bad (see also Hayes 1984: 71 for the result of an experiment in this connection reported in Bolinger (1962)). Furthermore, Hayes suggests that this preference is evidenced at various levels, ranging from the order of two coordinated constituents in terms of ascending phonological weight (e.g. '*soup and sandwich*' ,

<sup>28</sup> Here the BINARITY requirement at the IP and PhP levels are presented together as one constraint in the absence of motivation to separate them. See below for further discussion. The BINARITY requirement at the Foot level, on the other hand, is already specified in BINMAX >> BINMIN.

<sup>29</sup> Note that we deliberately choose to refer to the number of feet instead of that of syllables, on the account that we have argued that a single syllable can form a legitimate monosyllabic foot in verse scansion. Thus, a 7-syll line and an 8-syll one would both contain four feet.

<sup>30</sup> A mathematical abstraction of the syllable numbers in a line containing an odd number of feet (which can be either monosyllabic or disyllabic) is  $(4*n + 1)$  or  $(4*n + 2)$ , i.e., 5, 6, 9 or 10.



'*ladies and gentlemen*') to the 2+3 pattern of feet grouping in a verse line of English pentameter to the arrangement of verse lines in order of increasing length in the *Kalevala* meter of Finnish (Kiparsky 1968).

We argue that this constraint is also operative for prosodic structure: when a perfect symmetry cannot be achieved between the constituting parts, as in the case of an IP consisting of an odd number of feet, it is more preferable to place the longer parts towards the end. The length here is defined in terms of foot numbers. For example, for a 5-syll line containing three feet, the grouping of feet into PhP in observance of LONG-LAST would exhibit a '1+2' pattern, where the first foot and the remaining two feet respectively form one PhP<sup>31</sup>.

The relevance of LONG-LAST to PhP parsing may be supported by the argument that this '1+2' pattern is far from being accidental; rather, it can be given a principled account on independent grounds. Specifically, we argue that this '1+2' pattern is attributable to NON-FINALITY at higher levels of prosodic hierarchy. Recall that we already argued for the relevance of NON-FINALITY in the form of \*IP-FINAL-MONOFT and \*PHP-FINAL-MONOFT. Now bear in mind the prosodic hierarchy of Chinese proposed in (65), and extend this ban of monosyllabic foot at the final position of PhP to one level higher along the hierarchy. As a result, we get a ban of mono-footed PhP at the final position of IP. In other words, when a mono-footed PhP is inevitable, it is preferable to put this mono-footed PhP at the initial position of the IP, whilst leaving the bi-footed PhP at the final position of the IP (which is also the second, assuming that an IP contains two PhP's). This is exactly the '1+2' pattern dictated by LONG-LAST here. Viewed in this light, the LONG-LAST constraint of prosodic organization operative at the PhP level is then an extrapolation of the well-established NON-FINALITY constraint along the vertical dimension of the prosodic hierarchy, and thus well-grounded.

Now consider the PhP parsing of a 5-syll *Shijing* line. As argued in Section 2.3.4, the foot-level scansion is either (SS)(S)(SS) or (S)(SS)(SS), depending on the grammatical structure of the line. As for the PhP parsing, the LONG-LAST constraint gives rise to (SS)|(S)(SS) or (S)|(SS)(SS). What is particularly noteworthy about the latter parsing is that it is a very lop-sided pattern in terms of the distribution of phonological weight across the two PhP's. Note that the phonological weight here is directly measured by moras, and indirectly by syllable counting, assuming that in classical Chinese verse, most syllables are constituted by lexical items and thus bimoraic. Here it is apt to mention that PhP is argued to be 'the lowest prosodic constituent in the prosodic hierarchy that is sensitive to length' (Nespor and Vogel 1986: 185). More specifically, Nespor and Vogel (*Ibid.*) propose that there is a general tendency against forming particularly short (i.e. non-branching) PhP's. The length is stated there in term of branching versus non-branching upon the assumption that non-branching constituents are generally shorter than branching ones. We argue that for Chinese the branching versus non-branching account can be characterized

---

<sup>31</sup> It also deserves mentioning that to group the three feet into two PhP's which are then grouped into one IP is in line with the BINARITY constraint proposed above; in a framework allowing n-ary branching prosodic structures, the three feet can arguably be grouped directly into a ternary PhP, which is then grouped into a monary IP, or alternatively, the three feet can each constitute a monary PhP, and the three PhP's are then grouped into a ternary IP. Hence while entailing a simpler internal structure, the postulation of n-ary branching nonetheless undermines the restrictedness of the analysis.

straightforwardly by the syllable count (see also Duanmu 1997); accordingly the length can directly be gauged by syllable numbers, which is an indicator of phonological weight.

This general tendency to avoid particularly short PhP's actually invites the third constraint for the PhP parsing, which we refer to as Evenness. This constraint requires an even distribution of phonological weight across the PhP's in the IP, and is formulated as<sup>32</sup>:

- (71) EVENNESS  
In a prosodic domain, phonological weight should be evenly distributed across the prosodic units therein.

Therefore, (S)l(SS)(SS), with a 1:4 distribution of phonological weight, is a flagrant violation of the Evenness constraint, while another possible parsing (S)(SS)l(SS) better satisfies this constraint via minimizing the difference of phonological weight.

Indeed, this constraint can be independently motivated on the basis of the stress clash resolution. Following Duanmu's (2000) argument that Chinese is trochaic at syllable, foot and phrase levels, and assuming that the grid, which is the representation of rhythm (cf. among others, Liberman 1975; Liberman and Prince 1977; Hayes 1984), is built on the basis of the prosodic structure of a given string (Nespor and Vogel 1989; but cf. Selkirk 1980; Hayes 1995), we suggest that the parsing (S)l(SS)(SS) is ill-formed because of stress clash, which is illustrated below (moraic trochees are omitted):

- (72)
- |                |   |    |         |
|----------------|---|----|---------|
|                | x |    |         |
| Phrase level   | x | x  | ← Clash |
| Foot level     | x | x  | x       |
| Syllable level | x | xx | xx      |
- (S)l(SS)(SS).

A plausible way to resolve this clash is thus to shift the second grid mark to the next docking site, namely, above the next foot-level grid mark:

- (73)
- |                |   |    |                    |
|----------------|---|----|--------------------|
|                | x |    |                    |
| Phrase level   | x | x  | ← Clash resolution |
| Foot level     | x | x  | x                  |
| Syllable level | x | xx | xx                 |
- (S) (SS)l(SS).

Thus far, we have introduced the three constraints crucial to PhP-level parsing, namely, BINARITY, LONG-LAST, and EVENNESS. They interact and determine the PhP boundary delimitation for lines with a unidirectional branching structure.

<sup>32</sup> It is of interest to point out that the preference for an even distribution of phonological weight across prosodic domains is also evident in other higher-level prosodic constituents such as IP and U (Nespor and Vogel 1986).

We now consider how their interaction may be formally captured in an OT fashion. To begin with, some notes are in order for the evaluation of the constraints. First, as noted in (68), BINARITY is concerned with both the IP-level and PhP-level parsing. Hence (S)(SS)(SS)l incurs two violations of this constraint, due to its monarity at the IP level (one PhP in the IP) and ternarity at the PhP level (three feet in the one PhP) while (S)(SS)l(SS) incurs one violation due to the second PhP which contains only one foot. Second, violation of Evenness is gauged by counting the difference between the two PhP's in terms of syllable numbers. When an IP contains only one PhP, as in the parsing (S)(SS)(SS)l, EVENNESS is vacuously satisfied. Third, similarly, as LONG-LAST entails comparison between two PhP's, it is also vacuously satisfied when an IP contains only one PhP.

As for the ranking, first, (S)(SS)l(SS) provides the crucial argument for the ranking EVENNESS >> LONG-LAST. This is illustrated below:

(74)

(S)(SS)(SS)	EVENNESS	LONG-LAST
(S)l(SS)(SS)	**!*	
☞ (S)(SS)l(SS)	*	*

Second, (S)(SS)l(SS) also provides the ranking argument for BINARITY >> EVENNESS:

(75)

(S)(SS)(SS)	BINARITY	EVENNESS
☞ (S)(SS)l(SS)	*	*
(S)(SS)(SS)l	**!	

Third, by transitivity, we have BINARITY >> LONG-LAST. Indeed, this ranking is also supported by (S)(SS)l(SS). The suboptimal parsing (S)(SS)(SS)l violates BINARITY, but vacuously satisfies LONG-LAST (as well as EVENNESS). This is illustrated below:

(76)

(S)(SS)(SS)	BINARITY	LONG-LAST
☞ (S)(SS)l(SS)	*	*
(S)(SS)(SS)l	**!	

Finally, one legitimate concern remaining to be addressed here is whether BINARITY, which is now operative for both IP- and PhP-level parsing, needs to be stripped apart into, say, BINARITY-PHP and BINARITY-IP. Evidence from PhP parsings of two-footed IP's (i.e. 3- and 4-syll lines) demonstrates that this is not necessary and BINARITY as a cover constraint suffices to select the optimal candidate, as shown below:

(77) (i)

(S)(SS)	BINARITY	EVENNESS	LONG-LAST
(S)l(SS)	**!	*	
☞ (S)(SS)l	*		

(ii)

(SS)(SS)	BINARITY	EVENNESS	LONG-LAST
(SS) (SS)	**!		
四三 (SS)(SS)	*		

In fact, the parsing of such lines reveals an interesting trade-off between the two binarity requirements respectively at the PhP and IP level. Specifically, the fact that (S)(SS)| wins over (S)|(SS), and (SS)(SS)| over (SS)|(SS) suggests that when binarity cannot be achieved at both levels, it is more important to have binary structures at the PhP level than IP level. In other words, it is more important for a PhP to have two feet than for an IP to have two PhP's.

Thus we come up with the ranking hierarchy for the delimitation of PhP boundary for the foot-level scansion of verse line containing no SB:

(78) BINARITY >> EVENNESS >> LONG-LAST<sup>33</sup>.

It is important to point out that in the tableaux here, the optimal foot-level scansion is directly presented as the input solely for simplicity sake and does not suggest a two-level parsing. Rather we suggest that the scansion at both the foot and the PhP levels is carried out in a simultaneous and parallel fashion. For analytical purpose, we refer to the hierarchies respectively in charge of foot-level and PhP-level scansions as the main hierarchy (which has been our main focus so far) and the sub-hierarchy. Accordingly, a theoretically more precise way is to continue taking the grammatical structure of the line as the input and combine the main constraint hierarchy and the sub-hierarchy into one hierarchy. The candidate set, then, would be constituted by all potential foot-level and PhP-level parsings of the input string. For analytical convenience, however, we adopt a flexible practice below, namely, all tableaux will continue to have the grammatical structure of the line as the input, and as for the output forms we will be directly presenting the optimal PhP-level scansion for every potential foot-level scansion (including the optimal and the suboptimal ones), and in so doing tucking away behind the scene the working of the sub-hierarchy in (78) for the delimitation of PhP boundaries once the foot boundaries are given. In particular, the optimal PhP parsings for all the optimal foot-level scansions of *Shijing* lines are presented below as follows, which can then be directly imported into future tableaux under the main hierarchy whenever the PhP boundary needs to be explicitly marked out.

<sup>33</sup> It is notable that in the several tableaux here, the columns under Long-last are all grey, which might set one wondering whether Long-last is superfluous and can be disposed of. The fact is that although it tends to be dormant, Long-last is not superfluous: it turns out crucial in accounting for the metrical harmony in the genres of Jinti and Ci via tableau des tableaux. See discussion in Chapters 5 and 6 respectively.

(79)

Number of syllables in the line/IP	Optimal outputs with both foot and PhP boundaries marked out	Description of the prosodic structure
2	(SS)	1 Ft, 1 PhP, 1 IP
3	(S)(SS)	2 Ft, 1 PhP, 1 IP
4	(SS)(SS)	2 Ft, 1 PhP, 1 IP
5	(S)(SS) (SS) or (SS) (S)(SS)	3 Ft, 2 PhP, 1 IP
6	(SS) (SS)(SS)	3 Ft, 2 PhP, 1 IP
7	(S)(SS) (SS)(SS) or (SS)(SS) (S)(SS)	4 Ft, 2 PhP, 1 IP
8	(SS)(SS) (SS)(SS)	4 Ft, 2 PhP, 1 IP

Three comments about this table are in order. First, only lines that actually occur in the *Shijing* corpus are presented here. Thus, lines longer than 8 syllables are not included<sup>34</sup>. Second, the optimal foot-level parsings presented here are those based on relatively unambiguous empirical evidence of boundary lengthening which has been presented in the tableaux for the main hierarchy. All of them have been discussed so far except the 8-syll lines. Third, this table only includes the PhP-level parsings for those lines which do not contain a line-medial strongest grammatical boundary, since the PhP boundaries of those which do are determined straightforwardly and preemptively by such a grammatical boundary, as dictated by ANCHOR-I<sub>SB</sub>O<sub>PhP</sub> discussed in 2.3.6.1.1.1.1<sup>35</sup>.

We conclude this section by mentioning that the sub-hierarchy proposed in (78) equally applies to the PhP boundary delimitation in suboptimal forms. Below we will follow the above-mentioned practice of directly demarcating the optimal PhP boundaries in the candidate forms, bearing in mind that such boundaries are actually the winner among the many possible PhP boundary delimitations for a given foot-level scansion under the sub-hierarchy in (78).

Recall that the discussion in this subsection was triggered by the introduction of ANCHOR-I<sub>SB</sub>O<sub>PhP</sub> and \*PHP-FINAL-MONOFT into the sub-grammar. Now we are ready to explore their ranking in the emergent sub-grammar.

### 2.3.6.2 The ranking of ANCHOR-I<sub>SB</sub>O<sub>PhP</sub> and \*PHP-FINAL-MONOFT in the sub-grammar

The foregoing section (3.3.6.1) discussed the hierarchicality in the grammatical and prosodic structures, especially the PhP boundary delimitation in the verse context. From the viewpoint of developing the sub-grammar, this serves two purposes. First, it motivates the introduction of ANCHOR-I<sub>SB</sub>O<sub>PhP</sub> and \*PHP-FINAL-MONOFT. The former focuses on the output anchoring of the biggest grammatical break in the input,

<sup>34</sup> Both Chuci (the 2<sup>nd</sup> genre) and Ci (the 5<sup>th</sup> genre) have a very tiny number of 9-syll lines, though.

<sup>35</sup> In theory, the two cases may be combined by including ANCHOR-I<sub>SB</sub>O<sub>PhP</sub> in the sub-hierarchy presented in (82) as the top-ranking one to reflect its preemptive role, and subsequently leave it out from the main hierarchy. However, for analytical purpose, we opt to include ANCHOR-I<sub>SB</sub>O<sub>PhP</sub> in the main hierarchy and omit it from the sub-grammar. As a result, it would be superfluous for it to be included again in this sub-hierarchy: after all the sub-hierarchy is also part of the sub-grammar (although there is no evidence for its ranking with the main hierarchy). In other words, for those lines with a line-medial strongest boundary, this grammatical boundary would always emerge, at least in the optimal scansion, as the PhP boundary, and the sub-hierarchy in (78) is only relevant in delimiting PhP boundaries for lines containing no SB.

whereas the latter, by chopping up the IP into intermediate prosodic constituents, imposes a more restricted constraint than \*IP-FINAL-MONOFT on the location of monosyllabic feet within IP. Second, it enables the establishment of a connection between these two newly introduced constraints. Specifically, the output correspondent of the strongest grammatical boundary as required by ANCHOR-I<sub>SB</sub>O<sub>PHP</sub> constitutes the boundary of the PhP, a prosodic unit which cannot end with a monosyllabic foot as required by \*PHP-FINAL-MONOFT. Below we consider their ranking with regard to other constraints in the sub-grammar.

### 2.3.6.2.1 The ranking of ANCHOR-I<sub>SB</sub>O<sub>PHP</sub>

First, the scansion of lines of the grammatical structure [[SS]S][S[S[SS]]] (see (58)) offers crucial evidence for the ranking ANCHOR-I<sub>SB</sub>O<sub>PHP</sub> >> ANCHOR-IO. The scansion of such lines was presented in (59) under the sub-grammar arrived at till then, and exposed its insufficiency. For convenience sake, we repeat (59) below:

(80)

[[SS]S][S[S[SS]]]	BINMAX	BINMIN	GOODFT INTERJ	*IP- FINAL- MONOFT	ANCHOR- IO	ANCHOR- OI	ALIGNR (FT, IP)
ㄟ (S)(SS)(SS)(SS)		*			**!	*	12
ㄟ (SS)(SS)(S)(SS)		*			*		10
(SS)(S)(SS)(SS)		*			*		11!
(SS)(SS)(SS)	*!				**		7
(SSS)(SS)(SS)	*!				**	*	6

Under the sub-grammar presented in this tableau, (SS)(SS)(S)(SS) is the (unwanted) winner, more harmonious than the desired winner (S)(SS)(SS)(SS). However, now with the newly motivated ANCHOR-I<sub>SB</sub>O<sub>PHP</sub>, the picture changes dramatically. The strongest grammatical boundary (SB) in the line lies after the third syllable, and it corresponds to the PhP boundary in the desired winner (S)(SS)|(SS)(SS), but not in (SS)(SS)|(S)(SS). In the old sub-grammar, this failure of the conservation of the SB in the latter goes unpunished and the desired winner loses on account of more violations of ANCHOR-IO. Therefore, ANCHOR-I<sub>SB</sub>O<sub>PHP</sub> must dominate ANCHOR-IO in order for (S)(SS)|(SS)(SS) to beat (SS)(SS)|(S)(SS). The ranking argument is illustrated below:

(81)

[[SS]S][S[S[SS]]]	ANCHOR- I <sub>SB</sub> O <sub>PHP</sub>	ANCHOR-IO
ㄟ (S)(SS) (SS)(SS)		**
(SS)(SS) (S)(SS)	*!	*

This pair of candidates further provides the ranking argument for ANCHOR-I<sub>SB</sub>O<sub>PHP</sub> and the other ANCHOR constraint, i.e. ANCHOR-OI. This is illustrated below:

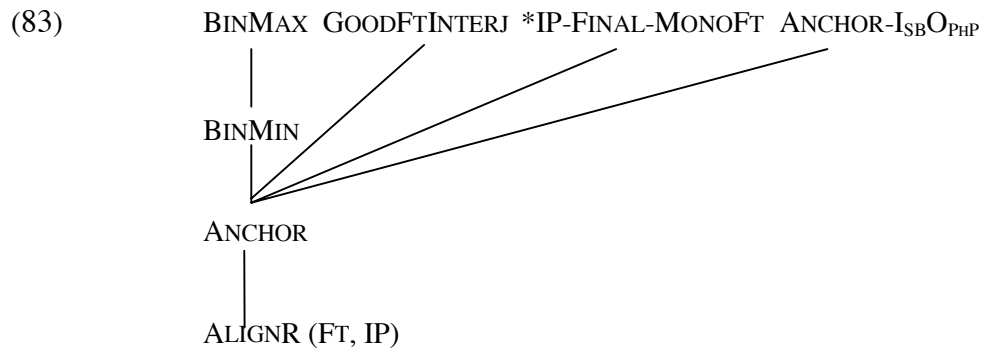
(82)

[[SS]S][S[S[SS]]]	ANCHOR- I <sub>SB</sub> O <sub>PHP</sub>	ANCHOR-OI
ㄟ (S)(SS) (SS)(SS)		*
(SS)(SS) (S)(SS)	*!	

Still using ANCHOR as the cover constraint, we have ANCHOR- $I_{SB}O_{PHP}$   $\gg$  ANCHOR. This ranking alone already suffices to elevate (S)(SS)l(SS)(SS) from a losing candidate to the optimal one.

Next, consider briefly the ranking between ANCHOR- $I_{SB}O_{PHP}$  and the other constraints. To begin with, as ANCHOR  $\gg$  ALIGNR (Ft, IP), by transitivity, we have ANCHOR- $I_{SB}O_{PHP}$   $\gg$  ALIGNR (Ft, IP). As for BINMAX, BINMIN, \*IP-FINAL-MONOFT (to be superseded by the newly proposed \*PHP-FINAL-MONOFT, see immediately below), and GOODFTINTERJ, ANCHOR- $I_{SB}O_{PHP}$  is in conflict with none of them. First, no candidate can survive a violation of BINMAX, or ANCHOR- $I_{SB}O_{PHP}$ , or \*IP-FINAL-MONOFT, which indicates that all three constraints are undominated. Second, as to BINMIN, given that the input line contains an odd number of syllables and that BINMAX is inviolable, BINMIN is inevitably violated in the optimal candidate together with some of its competitors, and accordingly BINMIN is not crucial in selecting the winner and the scansion of 7-syll lines provides no crucial evidence for the ranking between BINMIN and ANCHOR- $I_{SB}O_{PHP}$ . Third, there is no evidence for the ranking between ANCHOR- $I_{SB}O_{PHP}$  and GOODFTINTERJ. Of the *Shijing* lines containing an interjection, only two 5-syll ones have a line-medial interjection (structured as [SS]S<sub>l</sub>[SS] and scanned as (S)(SS<sub>l</sub>)(SS), illustrated in (45)), and all the others have interjections line-finally. Therefore, SB is present only in the two with a line-medial interjection. As shown in (46), such lines are scanned as (S)(SS<sub>l</sub>)l(SS), which satisfies both GOODFTINTERJ and ANCHOR- $I_{SB}O_{PHP}$ .

Therefore, with the introduction of ANCHOR- $I_{SB}O_{PHP}$  and its ranking, the sub-grammar presented in (52) is updated into:



**2.3.6.2.2 The ranking of \*PHP-FINAL-MONOFT**

However, merely adding ANCHOR- $I_{SB}O_{PHP}$  is still insufficient: even though its high ranking succeeds in forcing (SS)(SS)l(S)(SS) out of the competition, the desired winner (S)(SS)l(SS)(SS) still loses to (SS)(S)l(SS)(SS). This is shown below (compare (80)):

(84)

[[SS]S][S[S[SS]]]	BIN MAX	BIN MIN	GOOD FT INTERJ	*IP- FINAL- MONOFT	ANCHOR- $I_{SB}O_{PHP}$	ANCHOR- IO	ANCHOR- OI	ALIGNR (Ft, IP)
☞ (S)(SS)l(SS)(SS)		*				**!	*	12
(SS)(SS)l(S)(SS)		*			*!	*		10
☞ (SS)(S)l(SS)(SS)		*				*		11

(S)(SS)|(SS)(SS) loses on account of more violations of ANCHOR-IO; ANCHOR- $I_{SB}O_{PHP}$  is of no help here, as (SS)(S)|(SS)(SS) does satisfy it, unlike the earlier competitor (SS)(SS)|(S)(SS). Here \*PHP-FINAL-MONOFT becomes critical: (SS)(S)|(SS)(SS), the unwanted winner, fatally violates \*PHP-FINAL-MONOFT due to the monosyllabic foot at the end of the first PhP. By comparison, (S)(SS)|(SS)(SS) wins by refraining from formulating a monosyllabic foot PhP-finally, even though this move results in more violations of ANCHOR-IO. This trade-off implies that it is more important to avoid a PhP-final monosyllabic foot than to preserve all the input boundaries in the output. In other words, \*PHP-FINAL-MONOFT  $\gg$  ANCHOR-IO, illustrated below:

(85)

[[SS]S][S[S[SS]]]	*PHP-FINAL-MONOFT	ANCHOR-IO
☞ (S)(SS) (SS)(SS)		**
(SS)(S) (SS)(SS)	*!	*

This pair of competitors also furnishes the crucial ranking argument for \*PHP-FINAL-MONOFT  $\gg$  ANCHOR-OI, as illustrated below:

(86)

[[SS]S][S[S[SS]]]	*PHP-FINAL-MONOFT	ANCHOR-OI
☞ (S)(SS) (SS)(SS)		*
(SS)(S) (SS)(SS)	*!	

Therefore, we have \*PHP-FINAL-MONOFT  $\gg$  ANCHOR.

We now consider how \*PHP-FINAL-MONOFT should be ranked with the other constraints. For analytical purpose, we temporarily leave out \*IP-FINAL-MONOFT for special discussion below, and consider the remaining constraints, namely, BINMAX, BINMIN, GOODFTINTERJ, ANCHOR- $I_{SB}O_{PHP}$ , and ALIGNR (FT, IP). To begin with, as ANCHOR  $\gg$  ALIGNR (FT, IP), by transitivity, we have \*PHP-FINAL-MONOFT  $\gg$  ALIGNR (FT, IP). Second, as argued above, BINMAX, GOODFTINTERJ and ANCHOR- $I_{SB}O_{PHP}$  are all undominated; \*PHP-FINAL-MONOFT joins their company on the account that no potential output form can emerge as a winner if it violates \*PHP-FINAL-MONOFT by allowing a PhP-final monosyllabic foot, no matter how well it satisfies the other constraints. Third, there is no evidence for the ranking between \*PHP-FINAL-MONOFT and GOODFTINTERJ. As mentioned above in discussing the ranking of ANCHOR- $I_{SB}O_{PHP}$ , all *Shijing* lines containing interjections have interjections line-finally except for two 5-syll lines where the interjection occurs line-medially. For those lines with line-final interjections, GOODFTINTERJ actually forbids the interjection to parse into a monosyllabic foot, therefore working in the same direction as \*PHP-FINAL-MONOFT. For the two 5-syll lines with the line-medial interjection (structured as [SS]S<sub>I</sub>[SS]), again GOODFTINTERJ shares the same interest as \*PHP-FINAL-MONOFT: both encourage the optimal scansion (S)(SS<sub>I</sub>)(SS). Finally, \*PHP-FINAL-MONOFT does not conflict with BINMIN, for the same reason suggested above for the non-ranking between ANCHOR- $I_{SB}O_{PHP}$  and BINMIN, i.e., BINMIN is inevitably violated, hence not discriminating.



Now consider the ranking between \*IP-FINAL-MONOFT and \*PHP-FINAL-MONOFT, which are, as argued earlier, instantiations of NON-FINALITY at different levels. Recall that the former had been argued to be undominated in (52), before the latter was introduced; and when the latter was indeed introduced, we argued above that it is also undominated in the sub-grammar shorn of \*IP-FINAL-MONOFT. Thus, we have two constraints that are from the same family and both inviolable in the sub-grammar. The question now is whether they are both necessary in the sub-grammar, i.e. whether they are both active in selecting the optimal scansion (cf. Prince and Smolensky's (1993: 107) definition of 'active constraints').

In addressing this question, we first need to realize that a subtle relation holds between the satisfaction/violation patterns of these two constraints: a violation of \*IP-FINAL-MONOFT is necessarily accompanied by one of \*PHP-FINAL-MONOFT, as the end of an IP is necessarily the end of a PhP, but the reverse is not necessarily true. This is illustrated below ( $\checkmark$  and \* respectively standing for constraint satisfaction and violation)<sup>36</sup>:

(87)

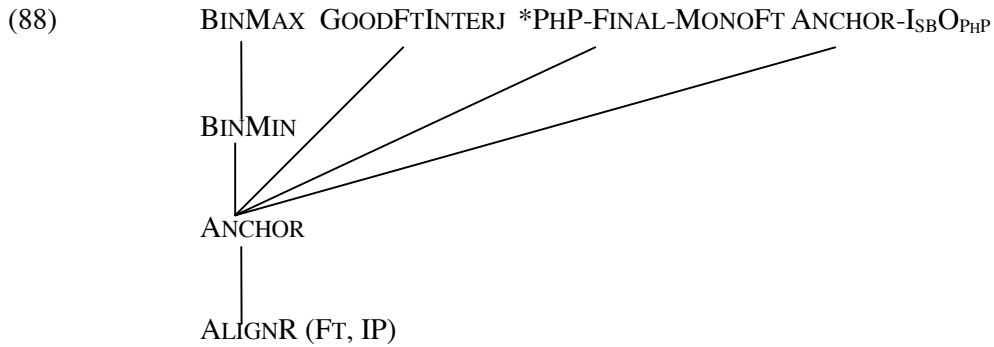
Input	*IP-FINAL-MONOFT	*PHP-FINAL-MONOFT
Candidate a (e.g. (SS) (SS)(S))	*	*
Candidate b (e.g. (SS) (S)(SS))	$\checkmark$	$\checkmark$
Candidate c (e.g. (SS)(S) (SS))	$\checkmark$	*

Notably, the satisfaction/violation pattern for \*IP-FINAL-MONOFT is a true subset of that for \*PHP-FINAL-MONOFT. In other words, \*IP-FINAL-MONOFT is not as discriminating as \*PHP-FINAL-MONOFT. \*PHP-FINAL-MONOFT can filter out those suboptimal forms that \*IP-FINAL-MONOFT cannot. In contrast, any sub-optimal form that is filtered out by \*IP-FINAL-MONOFT is also bound to fail the evaluation of \*PHP-FINAL-MONOFT. Hence these two constraints feature different degrees of granularity: \*IP-FINAL-MONOFT is coarser-grained whilst \*PHP-FINAL-MONOFT is finer-grained. The presence of the finer-grained one in the sub-grammar renders that of the coarser-grained one redundant, because the former can fulfil all the tasks that the latter can perform and in fact even more. Therefore, the introduction of \*PHP-FINAL-MONOFT announces the retirement of \*IP-FINAL-MONOFT from the sub-grammar<sup>37</sup>.

Thus, the sub-grammar is now streamlined into (cf. (83)):

<sup>36</sup> Note (91) is just for illustrative purposes and not intended to be read as a tableau.

<sup>37</sup> It deserves pointing out that under this new sub-grammar, the optimal scansions accounted for by the emergent sub-grammar at earlier stages remain optimal. This backtracking is arguably necessary as the replaced \*IP-FINAL-MONOFT is more liberal than the newly added \*PHP-FINAL-MONOFT, and concern arises as to whether some optimal candidates that won earlier by satisfying, among other constraints, \*IP-FINAL-MONOFT may inadvertently be harboring a PhP-final monosyllabic foot, which was innocuous back then, but fatal now. As it turns out, none of them have such feet.



We conclude this section by illustrating below how the optimal scansion of the grammatical structure  $[[SS]S][S[S[SS]]]$ , which triggered all this discussion, can now be satisfactorily accounted for. A comparison between the following tableau and its predecessor in (81) reveals the powerfulness of the two newly introduced constraints and the vital significance of introducing hierarchicality into the sub-grammar.

(89)

$[[SS]S][S[S[SS]]]$	BIN MAX	BIN MIN	GOOD FT INTERJ	*PHP- FINAL- MONOFT	ANCHOR- $I_{SB}O_{PHP}$	ANCHOR- IO	ANCHOR- OI	ALIGNR (Ft, IP)
$(S)(SS)(SS)(SS)$		*				**	*	12
$(SS)(SS)(S)(SS)$		*			*!	*		10
$(SS)(S)(SS)(SS)$		*		*!		*		11
$(SS)(SSS)(SS)$	*!					**		7
$(SSS)(SS)(SS)$	*!					**	*	6

### 2.3.7 More on the sub-grammar: scansion of 8-syll lines

After the elaborate discussion on 7-syll lines in the preceding section, the discussion on 8-syll lines appears rather anti-climactic: it offers no evidence for new constraints or new ranking. Hence, this section is a quick illustration of how the sub-grammar can sufficiently deal with the scansion of 8-syll *Shijing* lines.

There are only three 8-syll lines in the *Shijing* corpus and they all share the same grammatical structure  $[[S[SS]]S][S[S[SS]]]$  and are scanned as  $(SS)(SS)(SS)(SS)$ . For example:

- (90)  $[[bu4[zhi1 wo3]]zhe3] [wei4 [wo3 [he2 qiu2]]]$   
 not know me person say I what desire  
 ‘Those who do not know me wonder what I desire’  
 →  $(bu4zhi1) (wo3 zhe3) (wei4 wo3) (he2 qiu2)$ .

This optimal scansion can be fully accounted for by the sub-grammar as follows:

(91)

[[S[SS]]S][S[S[SS]]]	BIN MAX	BIN MIN	GOOD FT INTERJ	*PHP- FINAL- MONOFT	ANCHOR- I <sub>SB</sub> O <sub>PHP</sub>	ANCHOR- IO	ANCHOR- OI	ALIGNR (FT, IP)
☞ (SS)(SS) (SS)(SS)						***	*	12
(S)(SS)(S) (SS)(SS)		**		*!		*		18
(SS)(S) (SS)(S)(SS)		**		*!	*	**	*	16
(SSS)(S) (SS)(SS)	*!	*		*		**		11
(S)(SSS) (SS)(SS)	*!	*				**		13

The optimal scansion, (SS)(SS)(SS)(SS), wins by satisfying all higher-ranking constraints, even though it violates the lower-ranking ANCHOR and ALIGNR (Ft, IP).

So far on the basis of the scansion of all *Shijing* lines in the corpus, we have developed, in an incremental manner, the sub-grammar presented in (88) for the scansion of *Shijing* lines. Below we briefly reflect upon several points in the sub-grammar.

### 2.3.8 Some reflections on the sub-grammar of *Shijing* verse line scansion

Consider the constraints in the sub-grammar and the first thing to notice is that they are all universal, or at least expressible in universal terms. For example, as suggested before, GOODFTINTERJ can be reformulated in terms of RHTYPE=TROCHEE and STRESS-TO-WEIGHT and \*PHP-FINAL-MONOFT in terms of NON-FINALITY.

Second, the sub-grammar only deploys prosodic constraints that can be grounded in the phonological system of Chinese, and the construct of a separate metrical hierarchy or module is dispensed with. This is in conformity with the tenet of prosodic metrics outlined in Chapter 1. These constraints fall into the two categories of faithfulness (ANCHOR-I<sub>SB</sub>O<sub>PHP</sub> and ANCHOR) and markedness constraints (BINMAX, BINMIN, GOODFTINTERJ, \*PHP-FINAL-MONOFT and ALIGNR (Ft, IP)).

Third, the ranking ANCHOR-I<sub>SB</sub>O<sub>PHP</sub> >> ANCHOR-IO testifies the Pāṇinian Ranking Theorem (Prince and Smolensky 1993: 107) and reveals that the rigor of the boundary matching between the input and output varies at different levels of the hierarchical structure. Specifically, it exhibits a pattern of ‘the higher the stricter and the lower the more liberal’, namely, the higher the boundaries are in the hierarchy, the stricter is the boundary matching and the lower the boundaries the looser the matching. A similar pattern of the scalarity of rigor is also observed in Hayes’ (1989) study of the correspondence between the metrical and prosodic hierarchies in English art verse. This renders it interesting to explore to what extent this connection between the strictness of faithfulness constraints involving different levels and the position of the levels in the hierarchical structure can be extrapolated.

Finally, so far we have interpreted \*PHP-FINAL-MONOFT and \*IP-FINAL-MONOFT in terms of NON-FINALITY; in fact, by imposing certain restrictions on the end of a prosodic domain but not the beginning, these two constraints, and accordingly, NON-FINALITY per se, reflect yet another pattern proposed in Hayes (1989), which he refers

to as ‘beginning liberal and ending strict’. Again, it is worth investigating why the end of a phonological domain is susceptible to more restrictions than the beginning.

## 2.4 Formal grounding of metrical harmony

This section takes as its point of departure the observation that when presented with a verse line, the native speaker can usually offer judgment on metrical harmony of the line and that such judgments are especially strong and solid in those lines which are felt to be metrically most harmonious. This section provides a formal account of this metrical harmony experienced by the native speaker by arguing that it can be grounded in the sub-grammar for the corresponding genre.

To begin with, one proviso needs to be mentioned, namely, only the metrically most harmonious lines will be considered. This is in conformity with the above-mentioned observation. As is to be seen below, this also conforms to the postulated working tenet of the ‘tableau des tableaux’, i.e. one and only one optimal candidate will be selected. We will return to this issue below.

The formal mechanism to be employed is the ‘tableau des tableaux’, the nomenclature borrowed from Itô, Mester and Padgett (1995). It is named as such because the tableau is constructed out of the many ‘conventional’ tableaux that are used in developing the sub-grammar: the candidates in the ‘tableau des tableaux’ for a certain line type (say 4-syll lines) are constituted by parses from various grammatical structures for this line type to their respective optimal scansions. As such, the candidates actually represent the ‘parse route’ corresponding to various grammatical structures. The constraint ranking hierarchy for the evaluation of various candidates is constituted by the verse sub-grammar for the genre under discussion. The tableau des tableaux operates on the same principle as the conventional OT tableaux: the candidate that best satisfies the constraints ranked in the given order is the winner and there is only one such candidate. It might be suggested that the (finite) number of candidates in a tableau des tableaux are all optimal and the optimal parse is thus ‘optimal among the optimal’. More specifically, the optimal candidate is the best parse whose ‘parse route’ is the least offensive under the grammar – least offensive in terms of best satisfying the constraints in their given ranking order. Or, if we cite the notion of ‘OT harmony’ which refers to ‘the degree to which a possible analysis of an input satisfies a set of conflicting well-formedness constraints’ (Prince and Smolensky 1993: 3), the optimal candidate enjoys the greatest OT harmony.

Below we will argue that for each line type, the line cognized as metrically most harmonious coincides with the line whose corresponding parse is optimal in the tableau des tableaux for this line type. Put simply, the most harmonious grammatical structure coincides with the grammatical structure of the optimal parse. This consistent, non-trivial correspondence shows that the native speaker’s cognitively oriented metrical harmony judgment can be formally grounded in the grammar via the construct of OT harmony.

First, consider the 3-syll lines. As mentioned in Section 2.3.2, two grammatical structures occur in the corpus: S[SS] and [SS]S. The optimal scansion for lines of

both structures is (S)(SS). With the *Shijing* sub-grammar developed in Section 2.3, the following tableau des tableaux can be constructed<sup>38</sup>:

(92)

Candidate parses	BIN MAX	BIN MIN	GOODFT INTERJ	*PHP-FINAL-MONOF	ANCHOR-I <sub>SB</sub> O <sub>PHP</sub>	ANCHOR-IO	ANCHOR-OI	ALIGNR (Ft, IP)
a. $\begin{matrix} \text{S[SS]} \\ \text{(S)(SS)} \end{matrix}$		*						2
b. $\begin{matrix} \text{[SS]S} \\ \text{(S)(SS)} \end{matrix}$		*				*!	*	2

Corresponding to the two grammatical structures, there are two candidate parses. As shown here, of the two, parse (a) wins. On the side of the native speaker's metrical harmony judgment, 3-syll *Shijing* lines of the structure S[SS] are experienced as metrically most harmonious, as reported by my informants. Evidently, for 3-syll *Shijing* lines, the native speaker's metrical harmony judgment can be formally accounted for by the sub-grammar by virtue of OT harmony.

The other lines types are analyzed similarly. Below I will directly present the tableau des tableaux for 4- and 5-syll lines. For any given line type, the number of candidate parses equals that of the grammatical structure types occurring in the corpus for this line type. However, for simplicity sake, not all the grammatical structures are shown in the tableaux below.

(93) 4-syll lines

Candidate parses	BINMAX	BINMIN	GOODFT INTERJ	*PHP-FINAL-MONOF	ANCHOR-I <sub>SB</sub> O <sub>PHP</sub>	ANCHOR-IO	ANCHOR-OI	ALIGNR (Ft, IP)
$\begin{matrix} \text{[SS][SS]} \\ \text{(SS)(SS)} \end{matrix}$								2
$\begin{matrix} \text{S[S[SS]]} \\ \text{(SS)(SS)} \end{matrix}$						*!		2
$\begin{matrix} \text{[S[SS]S} \\ \text{(SS)(SS)} \end{matrix}$						**	*	2
$\begin{matrix} \text{[[SS]S]S} \\ \text{(SS)(SS)} \end{matrix}$						*!		2
$\begin{matrix} \text{S[[SS]S]} \\ \text{(SS)(SS)} \end{matrix}$						**	*	2

<sup>38</sup> The tableau des tableaux uses the same notation as the conventional tableau. (Cf. Itô, Mester and Padgett's (1995) use of the 'superhand'  $\text{☞}$  to refer to the winner in the tableau des tableaux.)

## (94) 5-syll lines

Candidate parses	BINMAX	BINMIN	GOODFt INTERJ	*PhP- FINAL- MONOFt	ANCHOR- I <sub>SB</sub> O <sub>PhP</sub>	ANCHOR- IO	ANCHOR- OI	ALIGNR (Ft, IP)
a. $\left\{ \begin{array}{l} [[SS][SS]]S \\ (SS)(S)(SS) \end{array} \right.$		*				*!	*	5
b. $\left\{ \begin{array}{l} [SS][S[SS]] \\ (SS)(S)(SS) \end{array} \right.$		*						5
c. $\left\{ \begin{array}{l} [S[SS]][[SS]] \\ (S)(SS)(SS) \end{array} \right.$		*						6!
d. $\left\{ \begin{array}{l} S[[SS]][[SS]] \\ (S)(SS)(SS) \end{array} \right.$		*						6!
e. $\left\{ \begin{array}{l} S[S[SS]]S \\ (SS)(S)(SS) \end{array} \right.$		*				**	*	5

In both cases, the metrically most harmonious lines correspond to the optimal parse, respectively being [SS][SS] and [SS][S[SS]] for these two line types. So far, only the main hierarchy suffices to establish this convergence and thus account for the metrical harmony. However, it turns out insufficient in the case of the 6-syll lines:

## (95) 6-syll lines

Candidate parses	BINMAX	BINMIN	GOODFt INTERJ	*PhP- FINAL- MONOFt	ANCHOR- I <sub>SB</sub> O <sub>PhP</sub>	ANCHOR- IO	ANCHOR- OI	ALIGNR (Ft, IP)
a. $\left\{ \begin{array}{l} [[SS][SS]][[SS]] \\ (SS)(SS)(SS) \end{array} \right.$								6
b. $\left\{ \begin{array}{l} [SS][[SS]][[SS]] \\ (SS)(SS)(SS) \end{array} \right.$								6
c. $\left\{ \begin{array}{l} [SS][S[SS]] \\ (SS)(SS)(SS) \end{array} \right.$						*!		6
d. $\left\{ \begin{array}{l} S[[SS]][[SS]] \\ (SS)(SS)(SS) \end{array} \right.$						*!		6
e. $\left\{ \begin{array}{l} [SS][S[SS]] \\ (SS)(SS)(SS) \end{array} \right.$						*!		6

Parses (a) and (b) emerge as equi-optimal and the main hierarchy alone fails to discriminate between them. It is noteworthy that they differ only in the position of the PhP boundary which reflects their different SB's. This prompts us to supplement the main hierarchy with the sub-hierarchy for PhP boundary delimitation and the tableau des tableaux under this extended hierarchy is as follows. The double line is used between the main hierarchy and the sub-hierarchy due to lack of evidence for their ranking.

(96)

Candidate parses	BINMAX	BINMIN	GOODFt INTERJ	*PHP-FINAL- MONOFt	ANCHOR- Isf O <sub>PHP</sub>	ANCHOR-IO	ANCHOR-OI	ALIGNR (Ft, Ip)	BINARITY	EVENNESS	LONG-LAST
a. $\left\{ \begin{array}{l} [[SS][SS]][SS] \\ (SS)(SS)(SS) \end{array} \right.$								6	*	**	*!
b. $\left\{ \begin{array}{l} [SS][SS][SS] \\ (SS)(SS)(SS) \end{array} \right.$								6	*	**	
c. $\left\{ \begin{array}{l} [SS][S[SS]] \\ (SS)(SS)(SS) \end{array} \right.$						*!		6	*	**	
d. $\left\{ \begin{array}{l} S[[SS][SS]] \\ (SS)(SS)(SS) \end{array} \right.$						*!		6	*	**	
e. $\left\{ \begin{array}{l} [SS][S[SS]] \\ (SS)(SS)(SS) \end{array} \right.$						*!		6	*	**	

We see that the PhP parsing hierarchy plays a crucial role in distinguishing between parse (a) and (b). Indeed, once again, the 6-syll lines of the right-branching structure [SS][SS][SS], which corresponds to the optimal parse, are cognized as metrically the most harmonious<sup>39</sup>. In the sense that both the main hierarchy and the sub-hierarchy are in the *Shijing* sub-grammar, respectively in charge of the foot parsing and PhP parsing, the metrical harmony data from 6-syll lines further supports our claim that metrical harmony can be formally grounded in the verse grammar in the form of OT harmony. Further evidence for this claim is provided by 7- and 8-syll lines the discussion of which is skipped here to avoid repetition.

We wish to conclude this section by further justifying our practice to examine only the most harmonious lines in formally accounting for the metrical harmony, and pay no attention to the degree of metrical harmony and its possible formal correlate. This is, we suggest, not a point of concern; rather it follows from both the empirical and theoretical considerations. Empirically, the native judgment on metrical harmony is converging and solid only regarding the metrically most harmonious lines; passing judgments on the often subtle difference in the degree of metrical harmony among the less harmonious lines is more challenging and typically characterized by less consensus. As Youmans (1989:10) observes, judgment about the degree of metrical tension even by ‘trained ears’ can be inconclusive and unreliable.

The theoretical consideration comes from the relative nature of OT harmony (Prince and Smolensky 1993). An OT grammar is solely concerned with the selection of the one and only one optimal candidate. All other candidates are indiscriminately treated as suboptimal; no second best (the ‘runner-up’), or the third best etc. are distinguished. In other words, the difference in constraint satisfaction/violation among the other suboptimal candidates is irrelevant, and it makes no sense to rank one over another among them. The issue of relativity in reckoning optimality in an OT grammar is explicitly addressed in Prince and Smolensky (Ibid.: 27) as follows:

<sup>39</sup> It is of interest to note that the 6-syll lines respectively of the structures [SS][SS][SS] and [[SS][SS]][SS] are experienced in subtle but distinct ways in terms of their metrical harmony: my informants unanimously reported that the former feels much smoother while the latter somehow feels ‘imbalanced’ and ‘tilted’. This complaint of imbalance, which might be superficially attributable to the position of the SB, is actually captured by LONG-LAST in the PhP parsing sub-hierarchy.

HOF [Harmonic Ordering of Forms] can never determine the absolute number of violations; that is, *count* them. HOF deals not in quantities but in comparisons, and establishes only relative rankings, not positions on any fixed absolute scale.

This implies that formal OT harmony is categorical in nature (McCarthy and Prince 1993a: 88); a form is either optimal or suboptimal and there is no such thing as ‘more/less optimal’. In the tableau des tableaux, only the optimal parse is selected and the satisfaction/violation pattern by the suboptimal parses carries no formal significance. Accordingly, the establishment of correlation is feasible only between the metrically most harmonious lines and the optimal parse. This meshes well with the fact that as far as the metrically most harmonious lines are concerned, the native judgments are strong and converging, but when it comes to less harmonious lines, their native judgments become somewhat equivocal.