Chapter 3 Jiuge sub-grammar

3.1 General description of the raw corpus

The corpus for this chapter comprises of eleven poems of the Jiuge sub-genre, which belongs to the genre collectively referred to as Chuci. A description of the Jiuge poems first entails some background knowledge about the Chuci genre in general.

Chuci, literally meaning ‘Elegies of Chu’, was composed around the Spring and Autumn Period and the Warring States Period around 300 BC when China was divided into seven feudal states, all vying with one another for fief and power (Yang and Yang 1983). Chu was one of the seven states, and Chuci was believed to be mostly composed single-handedly by Qu Yuan (ca.340-278 BC), the greatest poet of this period as well as an under-appreciated court official of Chu, around the time when Chu was on the verge of being annexed by the more powerful state of Qin. Chuci is hailed as the origin of romanticism in Chinese literature in the sense that the poets sought to decry the decline of the country, criticize the impotence of the court, and express their frustration and forlornness by way of romantic allusions and religious allegories, rather than realistic depiction. Consequently, there is a religious theme in many Chuci poems, particularly the Jiuge ones, where shamanism is a prevailing theme. Furthermore, as Chu was located in what is nowadays the south-central part of China, Chuci is also known as ‘Poems of the South’.

The anthology of Chuci that has survived today consists of seventeen chapters, each constituting a unique sub-genre¹. At the same time, the common features they share serve to collectively define the distinct genre known as Chuci. It is characterized by a unique integration of on the one hand, the distinct features of the Chu folk song, in particular, the wide use of the singing element ‘xi’, and on the other hand, the literary style of the northern states, which was mainly inherited from Shijing of the previous literary era. Nonetheless, the seventeen chapters differ in the degree in which they balance these two sources against each other and display variations on the basis of this overall refrain (Chen 1994). For example, Tianwen (‘Heavenly Questions’) and Dazhao (‘Great Summons’) are more heavily influenced by the Shijing tradition and feature more 4-syll lines and relatively few ‘xi’ s whereas all the other fifteen chapters are unified by the pervasive use of ‘xi’. Furthermore, some chapters of the latter group, for example Jiuge (‘Nine Songs’) to be studied here, are characterized by the presence of ‘xi’ in every single line, whereas others, e.g. Lisao (‘On Encountering Trouble’) and Yuanyou (‘Far-off Journey’), have lines containing ‘xi’ interspersed with those bearing more resemblance with Shijing lines.

The omnipresence of ‘xi’, regarded as the hallmark of the Chuci genre, and the virtual absence of transitional lines of the Shijing type in Jiuge renders it a full-blown sub-genre as far as the representation of ‘xi’ is concerned, the distinct characteristic of Chuci. Indeed, Jiuge is also among the most popular Chuci verse for modern

¹ For a full list of the titles and poems of the seventeen chapters, check out the website faculty.virginia.edu/cll/chinese_literature/Chuci/Chucitoc.htm.
speakers. On this account, we opt to select it as the empirical basis for an exploration of the sub-grammar.

With this background about the Chuci genre in general, we now move on to say a few words specifically about Jiuge. First, as is true with all Chuci chapters, Jiuge was intended to be mainly sung at its time of composition, though it was also argued to be recited back then; however, the tunes have long been lost and for modern speakers, the only viable manner for its delivery is recitation. Second, as mentioned earlier, every Jiuge line contains the interjection ‘xi’, and no other function words are used. Furthermore, ‘xi’ only occurs line-medially in Jiuge. In other words, every Jiuge line solely comprises of one (obligatorily present) line-medial ‘xi’ and a number of lexical syllables. Third, the Jiuge chapter contains altogether eleven poems, which are thematically concerned with different deities worshipped in the various parts of Chu. Fourth, the eleven poems display a considerable degree of variation: the number of lines contained in a poem ranges from 5 to 26, either belonging to one long stanza or grouped into several stanzas. The number of lines totals 253, and the overwhelming majority of them (250 out of 253) contain 5, 6 or 7 syllables. Admittedly, this corpus is much smaller than the Shijing corpus; nonetheless, it comprises of all Jiuge lines and as such offers a sufficient basis for the development of a robust scansion Jiuge sub-grammar.

3.2 Methodological issues and preview of the sub-grammar

3.2.1 Methodological issues

The organization of the chapter, and the notational convention are the same as those in the Shijing chapter (Section 2.2.1 of Chapter 2) and will not be repeated here. What differs between the Shijing chapter and this chapter in terms of methodology is that while the Shijing sub-grammar was developed from scratch, the Jiuge sub-grammar builds upon this sub-grammar. More specifically, in developing the Shijing sub-grammar, individual constraints and constraint rankings are motivated solely on the basis of the data; in developing the Jiuge sub-grammar, these constraints and rankings are, in principle, readily available and may be directly imported whenever applicable to the Jiuge data.

Obviously, this practice is enabled by the assumption outlined in Chapter 1 that there is one and only one overarching grammar for the modern speaker’s scansion of classical Chinese verse lines of all the five genres. At the same time, it was also mentioned there that this grammar is necessarily a partial ranking and the sub-grammars for different genres may well differ in their ranking. In view of this, we

2 Quite a few of the seventeen chapters are rather ill-known to modern speakers, largely due to the arcane, and in some cases, even obsolete diction used therein. Furthermore, some chapters, for example, Qijian (‘Seven Remonstrances’), are often featured by long lines with involved structures, which also thwarts its recitability and thus dampens its popularity with the modern speaker. On the other hand, to study the scansion sub-grammar for each of the seventeen sub-genres would be unfeasible for the current research which aims to cover the other four major genres of classical Chinese verse as well.
choose to adopt the relatively weak assumption in our analysis below, namely, that the sub-grammars only share constraints but not necessarily the ranking.

Translated into the specific analytical strategy, this assumption implies the following two points. First, the constraints motivated so far for the *Shijing* sub-grammar are part of the ‘constraint pool’ which contains all the constraints actively involved in the modern speaker’s scansion of classical Chinese verse lines, and we will be freely drawing constraints from this pool in the analysis below. At the same time, it is well conceivable that the *Shijing* constraints do not constitute the whole constraint pool; new constraints motivated by data from other sub-genres will be added to it. Second, in accordance with the above-mentioned assumption, constraint ranking will primarily be motivated on the basis of the *Jiuge* data, although as we will see, the ranking already arrived at in the *Shijing* sub-grammar considerably expedites the analytical process.

### 3.3 Jiuge sub-grammar

Of the 253 lines constituting the *Jiuge* corpus, except for one 8-syll line and two 9-syll lines, the remaining 250 lines range from 5- to 7-syllable long. As mentioned earlier, one of the most distinct features of the *Jiuge* genre is the omnipresence of ‘xi’; in terms of grammatical representation, we treat ‘xi’ as a stand-alone element. The organization of this section follows the same principle as that of Section 2.3 of the last chapter, namely, in the order of line length measured by syllable numbers.

#### 3.3.1 BINMAX, BINMIN and ANCHOR: evidence from 5-syll lines

All 5-syll *Jiuge* lines share the grammatical structure [SS][S][SS] (with the unbracketed middle syllable being ‘xi’) and the optimal scansion of (SS)(xi)(SS). For expository sake, in the following discussion, whenever necessary, we will directly present ‘xi’ in both input and output. Some examples are:

---

3 On a larger scale, this pool of constraints is in turn drawn from the repertoire of universal constraints shared by all human languages.

4 In theory, we could also adopt the strong assumption that the sub-grammars share both the constraints and the ranking. However, in practice, this does not really simplify the analysis and exposition, as in many cases we need first to show the assumed ranking does not work for the current data, and then establish the new ranking. An additional reason for the adoption of the weak assumption and the concomitant analytical strategy laid out here is that we wish to be able to organize the following section in a way identical to that of Section 2.3 of Chapter 2, namely, by covering all line types (in terms of syllable numbers). This is in turn motivated by the desire to present the reader with not only an analytical process to reach the sub-grammar, but also a descriptive picture of the various lines for the current genre (including the ‘non-crucial’ lines that apparently contribute little to the development of the sub-grammar). We wish to emphasize notwithstanding that the choice between these two assumptions is essentially a stylistic matter and as such carries no theoretical significance. This weak assumption and the concomitant analytical strategy will also be adopted in the following chapters.

5 It should be clarified that (SS)(S)(SS) is the optimal parsing for verse scansion, which, as argued in Footnote 11 of Chapter 2, disallows trisyllabic or polysyllabic feet. Indeed, (SSS)(SS) is an acceptable parsing for prose scansion, but crucially, in this parsing, the middle syllable of the trisyllabic foot has to be considerably reduced. The current study is only concerned with the grammar for verse scansion.
(1) \[ [ji2\  ri4]xi1\ [chen2\ liang2] \to\ (ji2\ ri4)(xi1)\ (chen2\ liang2) \]
good day\ xi\ morning\ fine
‘Ah, such a good day and fine morning’.

(2) \[ [hu4\ hu4]\ xi1\ [qiu1\ feng1] \to\ (hu4\ hu4)(xi1)(qiu1\ feng1) \]
heavy/reduplication\ xi\ autumn\ wind
‘Ah, how heavy the autumn wind is’.

To begin with, consider the potential candidate form (SSS)(SS), which, although acceptable in prose scansion, is nonetheless suboptimal in verse scansion. This exhibits exactly the same pattern as in the scansion of Shijing lines, where monosyllabic feet are conditionally acceptable but trisyllabic ones always banned. This invites the direct importing of the two binarity constraints BinMAX and BinMIN and their ranking BinMAX >> BinMIN from the constraint pool, which is illustrated as below:

(3) \[
\begin{array}{c|c|c}
[SS]xi[SS] & \text{BinMAX} & \text{BinMIN} \\
\hline
(SS)(xi)(SS) & * & \text{ } \\
(SSxi)(SS) & *! & \text{ } \\
\end{array}
\]

Second, consider the potential but suboptimal candidate (S)(Sxi)(SS)\(^6\). Compare it against the input structure [SS]xi[SS], and it becomes apparent that Anchor from the constraint pool is able to winnow out this candidate by penalizing its poorer boundary matching between the grammatical and the prosodic structures than (SS)(xi)(SS).

As to the ranking of Anchor with BinMAX and BinMIN, the current case provides no crucial argument; the reason is that the optimal winner, (SS)(xi)(SS), perfectly satisfies Anchor, while other suboptimal parsings, including (S)(Sxi)(SS), violate it in one way or another. Therefore, however Anchor is ranked, (SS)(xi)(SS) will always win, as illustrated below:

(4) \[
\begin{array}{c|c|c|c|c}
[SS]xi[SS] & \text{BinMAX} & \text{BinMIN} & \text{Anchor-Io} & \text{Anchor-Oi} \\
\hline
(SS)(xi)(SS) & * & * & * & * \\
(S)(Sxi)(SS) & * & * & * & * \\
(SSxi)(SS) & *! & * & * & * \\
(SS)(xiSS) & *! & * & * & * \\
\end{array}
\]

Thus the emerging sub-grammar developed solely on the basis of 5-syll line is, in Hasse graph:

\(^6\) Note that for some informants, (S)(SS)(SS) is also passable, but the consensus is that this scansion is ‘much less natural than (SS)(S)(SS)’. As befits our decision to address only one optimal candidate spelled out in Chapter 1, we treat (S)(SS)(SS) as sub-optimal.
Before we move on to 6-syll lines, the optimal scansion (SS)(xi)(SS) deserves further attention. Recall that *Shijing* has lines of the same structure [SS][S][SS], where the third syllable is an interjection. However, as shown in Section 2.3.4 of Chapter 2, such lines are best scanned as (S)(SS)(SS). For such *Shijing* lines, (SS)(S)(SS), which is the optimal scansion for *Jiuge* lines of the same structure, is suboptimal. The question now is what contributes to this difference. Apparently, these two lines belong to two different genres, but the real crux, we argue, lies in the fact that while the third syllable in the *Shijing* line is a normal interjection, the third syllable in the *Jiuge* line is ‘xi’, which is unique in many ways. With specific regard to the present case, the contrast between the parsing of 5-syll *Jiuge* lines and that of 5-syll *Shijing* lines of the same structure is attributable to the phonological property of ‘xi’, which is different from that of normal interjection syllables. Specifically, ‘xi’ can constitute a monosyllabic foot on its own whereas a normal interjection syllable can only serve as the weak syllable of a disyllabic foot. Indeed, as mentioned earlier, the omnipresence of ‘xi’ is one of the most distinct characteristics of the *Jiuge* sub-genre and in fact *Chuci* in general (Chen 1994), and it displays some unique phonological behaviors that call for a special treatment. However, since so far we have not demonstrated the full range of the phonological behavior of ‘xi’, we defer the discussion of ‘xi’ till a later point. For now, it is important to bear in mind that unlike normal interjection syllables, ‘xi’ can constitute a monosyllabic foot on its own.

### 3.3.2 **ANCHOR-OI >> BINMIN >> ANCHOR-IO: evidence from 6-syll lines**

6-syll lines constitute more than half of all *Jiuge* lines (128 out of 253); two grammatical structures are identified, namely, [[SS][S][xi][SS]] and [S][S][xi][SS]. Lines of these two structures are respectively scanned as (SS)(Sxi)(SS) and (S)(SS)(xi)(SS). For example,

\[(6) \quad [yu3 \ ns3] mu4 \ xi \ xian2 \ chi2 \rightarrow (yu3 \ ns3) (mu4 \ xi1) (xian2 \ chi2)\]

with you bathe xi place name

‘Ah, (I) bathe with you in Xianchi’.

\[(7) \quad [ling2[huang2 \ huang2]] xi1[ji4 \ jiang4] \rightarrow (ling2) (huang2huang2) \ xi1(ji4 \ jiang4)\]

spirit magnificent/redup. xi already descend

‘Ah, the magnificent spirits have already descended (upon us)’.

\[(8) \quad [fu3 \ chang2 \ jian4] xi1 [yu4 \ er3] \rightarrow (fu3) (chang2 \ jian4) \ (xi1)(yu4 \ er3)\]

stroke long sword xi jade ornament

‘Ah, (I) stroke my long sword and (put on) my jade ornament’.
While the 5-syll lines provide no evidence for the ranking between \textsc{binmin} and \textsc{anchor}, the scansion of lines of the structure \texttt{[SS]S[si][SS]} as \texttt{(SS)(Sxi)(SS)} offers crucial evidence for \textsc{binmin} >> \textsc{anchor}. The scansion \texttt{(SS)(S)(xi)(SS)} which fully conserves the input boundaries is suboptimal, due to its violations of \textsc{binmin}. The ranking argument is shown below:

\begin{verbatim}
(9)
\begin{tabular}{|c|c|c|}
  \hline
  & \textsc{binmax} & \textsc{binmin} \\
  \hline
  \texttt{[SS]S[si][SS]} & * & * \\
  \hline
  \texttt{(SS)(Sxi)(SS)} & * & * \\
  \hline
\end{tabular}
\end{verbatim}

Furthermore, since \textsc{binmax} >> \textsc{binmin}, by transitivity, we have \textsc{binmax} >> \textsc{anchor}. Thus the sub-grammar now is \textsc{binmax} >> \textsc{binmin} >> \textsc{anchor}.

However, this sub-grammar turns out insufficient to account for the scansion of lines of the structure \texttt{[S[SS]]S[si][SS]}: \texttt{(SS)(Sxi)(SS)} would still win, while the real winner is \texttt{(S)(SS)(xi)(SS)}. This is illustrated below:

\begin{verbatim}
(10)
\begin{tabular}{|c|c|c|}
  \hline
  & \textsc{binmax} & \textsc{binmin} \\
  \hline
  \texttt{[S[SS]]S[si][SS]} & * & * \\
  \hline
  \texttt{(SS)(Sxi)(SS)} & * & * \\
  \hline
  \texttt{(S)(SS)(xi)(SS)} & * & * \\
  \hline
\end{tabular}
\end{verbatim}

Under the current sub-grammar, the undesired winner (a) wins because of its maximal satisfaction of the highly ranked \textsc{binmax} and \textsc{binmin}, even though it violates \textsc{anchor}. Indeed, as the sub-grammar stands now, \texttt{(SS)(Sxi)(SS)} is bound to emerge as the winner irrespective of the grammatical structure. The reason is that \texttt{(SS)(Sxi)(SS)}, by evenly chopping up the six syllables in the line into three disyllabic feet, always maximally satisfies the two high-ranking well-formedness constraints, \textsc{binmax} and \textsc{binmin}, even though this chopping violates \textsc{anchor} by ignoring the input boundaries and freely inserting output ones\footnote{Evidently, in the so-called ‘minstrel’ performance style where the syntax and meaning of the verse line are ignored, \texttt{(SS)(Sxi)(SS)} would always win. But as mentioned in Chapter 1, such performance is linguistically uninteresting and not discussed in this study.}. By comparison, the desired winner \texttt{(S)(SS)(xi)(SS)}, whose boundaries fully match the input ones, loses because of the double violations of \textsc{binmin}.

The boundary matching between the input and output in the desired winner calls for a proper ranking of \textsc{anchor}. Recall that in (9) only \textsc{binmin} >> \textsc{anchor}-\textsc{io} was arrived at on the basis of crucial evidence, and we got \textsc{binmin} >> \textsc{anchor}-\textsc{oi} thereafter merely by assuming that the two \textsc{anchor} constraints stay together in the hierarchy unless there is evidence to the contrary. Now, the scansion of lines of the structure \texttt{[S[SS]]S[si][SS]} provides exactly such evidence to rank them apart:
specifically, it constitutes the crucial ranking argument for ANCHOR-OI >> BINMIN, and by transitivity, ANCHOR-OI >> ANCHOR-IO. This is shown below:

(11)  
<table>
<thead>
<tr>
<th>[S[SS]]xi[SS]</th>
<th>ANCHOR-OI</th>
<th>BINMIN</th>
<th>ANCHOR-IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>(SS)(Sxi)(SS)</td>
<td>*!</td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>*!(S)(SS)(xi)(SS)</td>
<td></td>
<td>**</td>
<td></td>
</tr>
</tbody>
</table>

As to the ranking between ANCHOR-OI and BINMAX, there is yet no crucial evidence, as neither (SSS)(xi)(SS), which violates BINMAX, nor (SS)(Sxi)(SS), which violates ANCHOR-OI, wins. This shows that both BINMAX and ANCHOR-OI are highly ranked. Hence, the emergent sub-grammar is:

(12)  
<table>
<thead>
<tr>
<th>BINMAX</th>
<th>ANCHOR-OI</th>
</tr>
</thead>
<tbody>
<tr>
<td>BINMIN</td>
<td></td>
</tr>
<tr>
<td>ANCHOR-IO</td>
<td></td>
</tr>
</tbody>
</table>

We close this section by drawing attention to a second type of well-formed feet containing ‘xi’ in addition to the monosyllabic foot (xi) mentioned at the end of Section 3.3.1, namely, (Sxi), which is the second foot in the optimal scansion for [[SS]S]xi[SS] illustrated in (9).

3.3.3 *IP-FINAL-MONOFT and GOODFT’Xl’: evidence from 7-syll lines

Compared with 5- and 6-syll ones, 7-syll lines exhibit more diverse patterns in both the grammatical structure and optimal scansion. Altogether six grammatical structures are identified: [[SS]S]xi[[SS]S], [[SS]S]xi[S[SS]], [SS]xi[S[SS]], [SS]xi[[SS]S], [S[SS]]xi[[SS]S], [S[SS]]xi[S[SS]], and [S[SS]]xi[SS]. Lines of these diverse structures have four optimal scansion, i.e. (SS)(Sxi)(SS), (SS)(xi)(SS), (S)(SS)(xi)(SS) and (SS)(SS)(xi)(SS). Below we give examples for each of the six grammatical structures and their scansion.

(13)  
(i)  

[[dong1 feng1] piao1] xi1 [[shen2 ling2] yu3]  
east wind blow xi holy spirit rain  
‘Ah, the east wind blows and the holy spirit rains’

⇒ (dong1 feng1) (piao1 xi1) (shen2 ling2 yu3)
As not all of the cases presented here are crucial in motivating new constraints or rankings, below we will first discuss those crucial ones to further develop the new sub-grammar and then illustrate the operation of this sub-grammar with some 'non-crucial' cases.

To begin with, so far there is yet no evidence for the ranking between BinMAX and Anchor-OI; the scansion of lines of the structure [[SS]xi[[SS]S]] provides crucial evidence for BinMAX >> Anchor-OI. This is illustrated below:

\[
\begin{array}{|c|c|c|}
\hline
\hline
SS)(Sxi)(SS) & * & \\
S)(Sxi)(SSS) & *! & \\
\hline
\end{array}
\]

Thus, the sub-grammar is now: BinMAX >> Anchor-OI >> BinMIN >> Anchor-IO. Now, consider lines of the structure [S[SS]]xi[[SS]S] and optimally parsed into (S)(SS)(xi)(S)(SS), as shown in (13)(iv). The current sub-grammar fails to predict the correct winner:
As the sub-grammar stands now, (S)(SS)(xi)(SS)(S) which better satisfies the highly ranked ANCHOR-OI, emerges as the winner. In contrast, the desired winner (S)(SS)(xi)(S)(SS) loses on account of violation of ANCHOR-OI. Now carefully compare this pair of competitors and it becomes apparent that the desired winner avoids an IP-final monosyllabic foot in order to satisfy some constraint that is presumably more important than ANCHOR-OI. For this purpose, *IP-FINAL-MONOFT from the constraint pool readily presents itself and it has to dominate ANCHOR-OI, as the unwanted winner satisfies ANCHOR-OI but violates *IP-FINAL-MONOFT. This is illustrated below:

Second, as ANCHOR-OI >> BINMIN >> ANCHOR-IO, by transitivity, we have *IP-FINAL-MONOFT >> BINMIN >> ANCHOR-IO. Third, *IP-FINAL-MONOFT does not conflict with BINMAX: both are in fact inviolable as no potential parsings that violate either of them will survive, for example, (SS)(Sxi)(SS)(S) in violation of *IP-FINAL-MONOFT and (SS)(Sxi)(SSS) in violation of BINMAX. Thus the emergent sub-grammar now is

However, under this sub-grammar, (S)(SS)(xi)(S)(SS), the desired winner, still fails to win; this time it loses to (S)(SS)(xiS)(SS), which equally satisfies *IP-FINAL-MONOFT, but incurs less violation of BINMIN. This is shown below:

Evidently, in order for (S)(SS)(xi)(S)(SS) to win over (S)(SS)(xiS)(SS), some new constraint is needed that crucially cashes in on some difference between the two. A careful observation of the pair reveals that the suboptimal candidate contains a foot (xiS) with ‘xi’ at the first position while in the optimal candidate ‘xi’ occurs at the
second position in the foot (Sxi). Given that Chinese feet are trochaic (Section 2.3.2 of Chapter 2), this difference can be rephrased as ‘xi’ occurring in the head versus non-head position of a foot.

Clearly, the foot (xiS) is offensive in (S)(SS)(xiS)(SS) and this restricted parsing of ‘xi’ is reminiscent of that of normal interjection syllables discussed in Chapter 2; however, the constraint proposed there, i.e. GOODFTINTERJ cannot be directly imported. The reason is because the well-formedness pattern of feet containing (normal) interjections slightly yet crucially differs from that of feet containing ‘xi’. Specifically, a normal interjection syllable can only occur at the non-head, i.e. second position in a disyllabic foot but cannot constitute a monosyllabic foot on its own. In contrast, ‘xi’ can occur either as the non-head syllable in a disyllabic foot headed by a full syllable, or constitute a monosyllabic foot on its own (as was shown towards the end of Section 3.3.1). The only position where ‘xi’ cannot occur is the head position of a disyllabic foot. Therefore, we propose the constraint GOODFT‘XI’ which is formulated as:

(19) GOODFT‘XI’
‘XI’ can only be legitimately parsed as the non-head of a disyllabic foot or as a monosyllabic foot on its own, but not as the head of a disyllabic foot.

The legitimate parsing pattern of ‘xi’ expressed by this constraint can be presented in table form as follows (where S stands for the full lexical syllable):

As for the ranking between GOODFT‘XI’ and the other constraints, first, the two candidates in (18) provide the crucial ranking argument for GOODFT‘XI’ >> BINMIN, illustrated as follows:

Second, as we have argued for BINMIN >> ANCHOR-IO, by transitivity, we get GOODFT‘XI’ >> ANCHOR-IO. Third, 6-syll Jiuge lines provide no crucial evidence for the ranking between GOODFT‘XI’ and ANCHOR-IO. In fact, a careful consideration of all the line types in the Jiuge corpus reveals that none of them offers such crucial data. Thus the emergent sub-grammar now becomes:

---

8 For the moment, we only present this parsing pattern for the purpose of developing the constraint without further explanation as to why the pattern is such; in the section below, we will argue that these patterns can be accounted for by the peculiar phonological property of ‘xi’.
To revisit lines of the structure \([S[SS]]xi[[SS]S]\) in (15), we see that its optimal scansion can now be accounted for:

\[
\begin{array}{cccc}
[S[SS]]xi[[SS]S] & \text{BINMAX} & \text{*IP-FINAL-MONOFT} & \text{GOODFT}'X'i' \\
\text{BINMIN} & \text{ANCHOR-OI} & \text{ANCHOR-IO} & \\
\text{BINMIN} & \text{ANCHOR-OI} & \text{ANCHOR-IO} & \\
\end{array}
\]

![Table](image)

This sub-grammar is also adequate to account for the scansion of 7-syll Jiuge lines of the other grammatical structures in the corpus. For space consideration, below we only illustrate the working of the sub-grammar with the grammatical structures \([SS]xi[S[SS]]\) ((13) (iii)), and \([S[SS]]xi[SS]\) ((13)(vi)):

\[
\begin{array}{cccc}
[SS]xi[S[SS]] & \text{BINMAX} & \text{*IP-FINAL-MONOFT} & \text{GOODFT}'X'i' \\
\text{BINMIN} & \text{ANCHOR-OI} & \text{ANCHOR-IO} & \\
\text{BINMIN} & \text{ANCHOR-OI} & \text{ANCHOR-IO} & \\
\end{array}
\]

![Table](image)

\[
\begin{array}{cccc}
[S[SS]]xi[SS] & \text{BINMAX} & \text{*IP-FINAL-MONOFT} & \text{GOODFT}'X'i' \\
\text{BINMIN} & \text{ANCHOR-OI} & \text{ANCHOR-IO} & \\
\text{BINMIN} & \text{ANCHOR-OI} & \text{ANCHOR-IO} & \\
\end{array}
\]

![Table](image)

\[\text{Note that following our earlier practice, the arbitrary ranking GOODFT}'X'i' >> ANCHOR-OI is assigned to this non-ranking pair. One may of course assign the ranking ANCHOR-OI >> GOODFT}'X'i', which would yield the same optimal candidate.}\]
3.3.3.1 The phonological representation of ‘xi’

In this section, we are going to argue that the well-formedness pattern for feet containing ‘xi’ presented in (20) and expressed in the form of the constraint GOODFT ‘Xi’ is attributable to the unique phonological representation of ‘xi’.

For this purpose, it is enlightening to compare the parsing of ‘xi’ with that of normal interjections. The most dramatic difference is that the former can legitimately constitute a monosyllabic foot whereas the latter cannot. This is shown clearly in the following table which presents, side by side, the parsing of ‘xi’ and normal interjection syllables (indicated as SI) as presented in Section 2.3.4 of Chapter 2.

(26)

<table>
<thead>
<tr>
<th>Parsing of ‘xi’</th>
<th>Parsing of normal interjection syllables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foot type</td>
<td>Well-formedness</td>
</tr>
<tr>
<td>(Sxi)</td>
<td>good</td>
</tr>
<tr>
<td>(xi)</td>
<td>good</td>
</tr>
<tr>
<td>(xiS)</td>
<td>bad</td>
</tr>
</tbody>
</table>

As shown above, this difference underlies the different scansions of lines between Shijing and Jiuge lines of the same structure. Superficially, the parsing pattern exhibited by ‘xi’ appears rather paradoxical: on the one hand, if ‘xi’ can constitute a monosyllabic foot, then it should be heavy; on the other hand, crucially it cannot occur as the head of a disyllabic trochee, which shows that it cannot really be heavy. We argue that the formal construct of empty mora offers a ready solution to this apparent dilemma: ‘xi’ has as its phonological representation an empty mora in addition to a filled mora whereas the normal interjection syllable is represented as a monomoraic structure where the mora is filled. Diagrammatically, the phonological representation of ‘xi’ is:\[10\]

(27)

\[
\begin{array}{c}
\sigma \\
\alpha \\
\beta \\
\chi \\
\end{array}
\]

\[\mu \]

\[i\]

\[10\] We assume that all moras, including the empty mora, are parsed into syllables; in other words, PARSE-μ is inviolable. This constraint stems from the general theory of prosodic licensing developed in Itô (1986) within the derivational framework and it requires that all phonological segments be prosodically licensed. Indeed, as a general constraint, PARSE could be understood as a constraint family requiring the prosodic licensing of phonological constructs at all levels of the prosodic structure; for example, PARSE-SYL which requires all syllables to be parsed into feet is another member of this family.
Compare this with the phonological representation of normal interjection syllables, for example, ‘yi’:

\[
\text{(28)}
\]

\[
\begin{array}{c}
\ \ \ \ y \\
\mu \\
\ \ \ \ i
\end{array}
\]

We argue that it is this empty mora in (27) that leads to the greater flexibility of ‘xi’ in its phonological parsing. Specifically, this underlying representation gives rise to two surface representations of ‘xi’ contingent on its position in the foot: (i) as bimoraic when occurring alone or at the second position of a disyllabic (and trochaic) foot, thus rendering these feet licit; (ii) as monomoraic when occurring at the first position of a disyllabic foot, thus rendering the foot illicit. Put differently, this underlying representation enables ‘xi’ to legitimately occur as either a monosyllabic foot or the non-head syllable of a disyllabic foot, but never as the head syllable of a disyllabic foot. Furthermore, we argue that these various surface realizations of ‘xi’ in different environments result from the interaction of several OT constraints. For clarity sake, below we refer to the constraint hierarchy responsible for the selection of the optimal surface realization of ‘xi’ as the ‘‘xi’-grammar’.

Given the postulated presence of an empty mora in the underlying representation of ‘xi’, to start the ball rolling, we need a constraint that demands the conservation of segment-to-mora linkage. McCarthy (2000: 159) proposes two universal faithfulness constraints NO-SPREAD and NO-DELINK which respectively militate against spreading and delinking of autosegmental associations. Specifically, NO-SPREAD requires that any association line that is present in the output have a correspondent in the input, and NO-DELINK requires that any association line present in the input have a correspondent in the output. As such they are respectively analogous to DEP (‘do not insert association lines’) and MAX (‘do not delete association lines’). These two constraints exert opposite influences and apply to each pair of associated autosegmental tiers, such as tone, place, mora and segment. Apparently, what is relevant here is the segment-to-mora association. Following McCarthy’s (Ibid.:159) formal notation, they are respectively as follows:

\[
\text{(29) NO-SPREAD (} \mu, \text{ SEG)}^{11}
\]

Let \(S_1\) and \(S_2\) stand for two related phonological representations, where

\( \mu_1 \) and \( \text{seg}_1 \in S_1 \),
\( \mu_2 \) and \( \text{seg}_2 \in S_2 \),
\( \mu_1 \mathcal{R} \mu_2 \), and
\( \text{seg}_1 \mathcal{R} \text{seg}_2 \),

if \( \mu_2 \) is associated with \( \text{seg}_2 \),

---

11 Two points are worth mentioning here. First, as the mora and segment are the only autosegmental tiers involved here, below we will simply note this constraint as NO-SPREAD. Second, a similar pair of constraints FILLLINK (‘All association relations are part of the input’) and PARSELINK (‘All input association lines are kept’) are proposed in Ito, Mester and Padgett (1995:586) within the older Parse/Fill/Containment framework (Prince and Smolensky 1991, 1993).
then \( \mu_i \) is associated with \( \text{seg}_1 \).

(30) **NO-DELINK** (\( \mu \), \text{SEG})

Let \( S_1 \) and \( S_2 \) stand for two related phonological representations, where

\[
\begin{align*}
\mu_1 \text{ and } \text{seg}_1 & \in S_1, \\
\mu_2 \text{ and } \text{seg}_2 & \in S_2, \\
\mu_1 & \overset{R}{\rightarrow} \mu_2, \text{ and} \\
\text{seg}_1 & \overset{R}{\rightarrow} \text{seg}_2,
\end{align*}
\]

if \( \mu_1 \) is associated with \( \text{seg}_1 \),

then \( \mu_2 \) is associated with \( \text{seg}_2 \).

In the current context, **NO-SPREAD** is tantamount to forbidding the association of the empty mora, unassociated with any segment in the input, with any phonological segment in the output, thus maintaining its ‘emptiness’. **NO-DELINK**, on the other hand, prevents any association lines present in the input from being deleted in the output. We assume that **NO-DELINK** is highly ranked in the ‘\( \cdot\text{x}i\cdot \)’-grammar and hence all input segment-to-mora linkages are preserved in the output; only **NO-SPREAD** is subject to interaction with other constraints, to be discussed below.

In addition, the presence of an empty mora in the syllable structure is highly marked, and a markedness constraint is in order that requires all moras to be linked with phonological segments in the output, thus explicitly banning the occurrence of any empty mora in the surface structure. We refer to it simply as **NOEMPTYMORA**, stated below:

(31) **NOEMPTYMORA**:

A mora must be filled.

Before we proceed, one notational remark is in order. As we assume **PARSE-\( \mu \)** is highly ranked and all moras are parsed into syllables, below for simplicity sake, (27) is simplified into

(32) \[
\mu \quad \mu
\]

where both the syllable node and the onset \( x \) are dropped.

As is evident from (26), ‘\( \cdot\text{x}i\cdot \)’ displays a unique flexibility in its parsing in that it assumes different surface forms depending on its phonological environment: (i) bimoraic when occurring alone, given the legitimacy of \( (\text{x}i) \); (ii) monomoraic when occurring at the first position of a disyllabic foot, given the illegitimacy of \( (\text{x}i\text{S}) \). When occurring at the second position in a disyllabic foot, it can in theory be either bimoraic or monomoraic\(^{12}\).

\(^{12}\) This follows from the inventory of good and bad foot structures in Chinese presented in Chapter 2, and repeated here for convenience sake:
In terms of the preservation or deletion of the empty mora in the underlying representation, this pattern is tantamount to (i) when occurring alone, the underlying empty mora is preserved and filled in the surface representation; (ii) when preceding a full syllable in a disyllabic foot, this empty mora is deleted in the surface form and thus cannot head the disyllabic trochee. As for the scenario where ‘xi’ follows a full syllable in a disyllabic foot, a priori, no inference can be made regarding its surface structure, though as to be seen below in (42), ‘xi’ actually surfaces as bimoraic in this context.

This unique flexibility in the parsing of ‘xi’ is partially captured by the postulation of an empty mora in its underlying structure. Furthermore, certain constraints are needed to account for the surface appearance of this mora in some contexts and disappearance in others. On the one hand, the preservation of the underlying empty mora at least in some contexts calls for MAX-µ, a faithfulness constraint requiring all moras that are present in the input to be also present in the output, and this requirement is in force irrespective of whether the mora is empty or not. On the other hand, the underlying empty mora is not always preserved, but only when it is foot-final, as is evidenced from the licitness of (Sxi) and (xi) but illicitness of (xiS). That (xiS) is illicit suggests that the empty mora is deleted and ‘xi’ surfaces as a monomoraic syllable when foot-initial, thus resulting in a bad foot of the quantitative structure (LH).

This instructs us that by imposing a blanket requirement that all underlying moras should be preserved in the surface structure regardless of the environment in which the mora occurs, MAX-µ is too indiscriminating and needs to be supplemented by a ‘fine-tuning’ constraint that is responsible for the preservation of the underlying empty mora only in foot-final positions\(^\text{13}\). This latter constraint can be expressed as a position-specific version of MAX-µ, which is referred to as a ‘positional faithfulness constraint’ (Beckman 1997a, b). Specifically, it requires the preservation of the empty mora only when the syllable occurs foot-finally, and we represent it as MAX-µ\(_F\). The right bracket and the subscript \(F\) following the position-neutral MAX-µ indicates the foot-final environment in which MAX-µ operates\(^\text{14}\). Thus, we have two new constraints from the same family but of different granularity:

(i) Good foot structures in Chinese

\[
\begin{align*}
\text{(i) Good foot structures in Chinese} \\
\text{x} & \quad \text{x} & \quad \text{x} \\
(S\ S) & \quad (S\ S) & \quad (S) \\
(\mu\mu) & \quad (\mu\mu) & \quad (\mu) \\
H & \quad H & \quad L & \quad H
\end{align*}
\]

(ii) Bad foot structures in Chinese

\[
\begin{align*}
\text{(ii) Bad foot structures in Chinese} \\
\text{x} & \quad \text{x} & \quad \text{x} \\
(S\ S) & \quad (S\ S) & \quad (S) \\
(\mu) & \quad (\mu) & \quad (\mu) \\
L & \quad L & \quad H & \quad L
\end{align*}
\]

\(^\text{13}\) Note that in spite of its coarse granularity, MAX-µ is still indispensable and cannot be superseded by the position-specific version of MAX-µ altogether. Both need to be present in the constraint hierarchy and actually as is to be shown immediately below, the power of the analysis lies exactly in the ranking between the general constraint and the position-specific one as well as their ranking with the other two constraints introduced earlier.

\(^\text{14}\) It needs to be pointed out that the proposal of the positional faithfulness constraint MAX-µ\(_F\) bears on weighing between the two alternative views of ‘context markedness’ versus ‘positional faithfulness’ as discussed in Kager (1999:407ff), and the ‘positional faithfulness’ view is adopted here largely in view
(33) \( \text{MAX-}\mu \)
A mora that is present in the input is present in the output.

(34) \( \text{MAX-}\mu|_F \)
A mora that occurs in a foot-final position in the input is present in the output.

Now that we have four constraints at our disposal, the next task is to rank them properly in order to account for the various phonological behaviors of ‘xi’, as presented in (26), which was rephrased in terms of the surface preservation or deletion of the underlying empty mora above.

To begin with, when forming a foot on its own, ‘xi’ is bimoraic where the empty mora is filled. This ban of the empty mora from the surface representation of ‘xi’ shows \( \text{NOEMPTYMORA} \gg \text{NO-SPREAD} \):

(b) \( \mu \mu \)

(35) (xi)

<table>
<thead>
<tr>
<th>( \mu )</th>
<th>( \mu )</th>
<th>( i )</th>
<th>( \text{NOEMPTY MORA} )</th>
<th>( \text{NO-SPREAD} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>b. ( \mu \mu )</td>
<td></td>
<td></td>
<td>( *! )</td>
<td></td>
</tr>
<tr>
<td>c. ( \mu \mu )</td>
<td></td>
<td></td>
<td></td>
<td>( * )</td>
</tr>
</tbody>
</table>

Second, that ‘xi’ surfaces as monomoraic in some cases but bimoraic in other cases calls for the deployment of two faithfulness constraints \( \text{MAX-}\mu \) and \( \text{MAX-}\mu|_F \). For one thing, \( \text{MAX-}\mu \) penalizes the deletion of any underlying mora, and the illegitimacy of \( *(\text{xiS}) \) which testifies to the deletion of the underlying empty mora in this case, offers crucial evidence for \( \text{NO-SPREAD} \gg \text{MAX-}\mu \), and by transitivity, \( \text{NOEMPTYMORA} \gg \text{MAX-}\mu \). This is shown below:

---

of Kager’s (Ibid.) general arguments for it, although he also states there that ‘in most cases it is simply impossible to find evidence for one view or the other’ and that ‘in current literature both views have been adopted’ (see Kager (Ibid.) for details).
A slight twist of mind is needed to interpret this tableau: candidate (a) is the optimal parsing in the sense that ‘xi’ in this position is optimally parsed as monomoraic, even though the foot is illegitimate; indeed, we know that ‘xi’ surfaces as monomoraic in this environment exactly because the foot with ‘xi’ at the head position followed by a full syllable is illegitimate, as it can only be of the (LH) structure.

However, although these three constraints ranked as such succeed in predicting the optimal parsing of ‘xi’ in the environment of (xiS), and in an indirect way, account for why (xiS) is illegitimate, they turn out to be inadequate for the two legitimate parsings of ‘xi’, namely, (xi) and (Sxi). Take the monosyllabic foot constituted by ‘xi’ as an example:

Compare the desired winner (candidate (c)) and the unwanted winner (candidate (a)): in the former the empty mora is deleted with only an inconsequential penalization due to the lowest ranking of MAX-μ. Clearly, some constraint is needed to preserve the empty mora on the part of the desired winner. At the same time, however, this preservation of the empty mora obviously has to be sensitive to the specific position of the mora because the empty mora cannot be preserved when it occurs at the first position of a foot as shown in (36). MAX-μ[f exactly fits in this role by requiring the preservation of the mora only when it is foot-final.
The issue now is how to rank $\text{MAX-} \mu I_F$. First, the pair of candidates (a), the unwanted winner, and (c), the desired winner, provides crucial evidence for the ranking $\text{MAX-} \mu I_F >> \text{NO-SPREAD}$: (a) violates $\text{MAX-} \mu I_F$ but satisfies $\text{NO-SPREAD}$ while (c) violates $\text{NO-SPREAD}$ but satisfies $\text{MAX-} \mu I_F$. This is shown below:

\begin{tabular}{ccc}
 & $\mu$ & $\mu$ \\
\hline
\text{MAX-} \mu I_F & \text{NO-SPREAD} \\
\hline
\text{a.} & $\mu$ & $\mu$ \\
\text{c.} & $\mu$ & $\mu$ \\
\end{tabular}

Second, because $\text{NO-SPREAD} >> \text{MAX-} \mu$, by transitivity, we have $\text{MAX-} \mu I_F >> \text{MAX-} \mu$. Third, there is no crucial evidence for the ranking between $\text{MAX-} \mu I_F$ and $\text{NOEMPTYMORA}$, as candidate (b) satisfies $\text{MAX-} \mu I_F$, but does not outperform the desired winner (c) as it loses on account of its violation of $\text{NOEMPTYMORA}$. Thus, the ‘$\mu i$’-grammar is now:

\begin{equation}
\text{MAX-} \mu I_F, \text{NOEMPTYMORA} >> \text{NO-SPREAD} >> \text{MAX-} \mu
\end{equation}

Under this constraint hierarchy, the parsing of ‘$\mu i$’ when it occurs alone is as follows (compare it with (37)):

\begin{tabular}{cccc}
 & $\mu$ & $\mu$ & $\mu$ \\
\hline
\text{MAX-} \mu I_F & \text{NOEMPTYMORA} & \text{NO-SPREAD} \\
\hline
\text{a.} & $\mu$ & $\mu$ & $\mu$ \\
\text{b.} & $\mu$ & $\mu$ & $\mu$ \\
\text{c.} & $\mu$ & $\mu$ & $\mu$ \\
\end{tabular}

Clearly the introduction of $\text{MAX-} \mu I_F$ dramatically changes the picture and ‘$\mu i$’ surfaces as bimoraic, which accounts for why it can form a legitimate foot on its own.

\footnote{Presumably other constraints such as DEP-$\mu$ are also involved to prevent the free insertion of moras and serve other basic ‘housekeeping’ purposes; however since they are not critical to the present discussion, we leave them out.}
Furthermore, it does not affect the case of *\(xiS\) analyzed earlier in (36), as the final syllable is a full bimoraic one and \(\text{MAX}-\mu\) is vacuously satisfied by all candidates presented there. This is shown below:

\[(41) *\(xiS\)\]

\[
\begin{array}{|c|c|c|c|}
\hline
\mu & \mu & \mu & \mu \\
\hline
\text{MAX}-\mu \} \text{F} & \text{NOEMPTY} & \text{MORA} & \text{NO-SPREAD} & \text{MAX-}\mu \\
\hline
a. \mu & \mu & \mu & \mu \\
\hline
b. \mu & \mu & \mu & \mu \\
\hline
c. \mu & \mu & \mu & \mu \\
\hline
\end{array}
\]

We now proceed to consider the parse of ‘\(xi\)’ in (Sxi) where the first syllable is a heavy, bimoraic one. As is shown by the tableau below, in this environment, ‘\(xi\)’ surfaces as bimoraic, and accordingly the foot has a quantitative structure of (HH), which is well-formed.

\[(42) (Sxi)\]

\[
\begin{array}{|c|c|c|c|}
\hline
\mu & \mu & \mu & \mu \\
\hline
\text{MAX}-\mu \} \text{F} & \text{NOEMPTY} & \text{MORA} & \text{NO-SPREAD} & \text{MAX-}\mu \\
\hline
a. \mu & \mu & \mu & \mu \\
\hline
b. \mu & \mu & \mu & \mu \\
\hline
c. \mu & \mu & \mu & \mu \\
\hline
\end{array}
\]

Thus we have shown that the seemingly perplexing and irregular phonological behavior of ‘\(xi\)’ can be satisfactorily accounted for by the postulation of an underlying
empty mora and the above ‘\textit{xi}’-grammar which crucially makes use of the positional faithfulness constraint MAX-µ$_F$.

However, it needs to be realized that the well-formedness pattern of ‘\textit{xi}’ presented in (26) is only concerned with the foot-level parsing. In fact, ‘\textit{xi}’ is subject to further restrictions at higher levels of the prosodic structure, i.e. PhP and IP. Specifically, while (xi) is a legitimate foot, it cannot head a PhP or IP$^{16}$. For example, the PhP-level parsing for the optimal scansion for $[S[S[SS]]]xi[SS]$ in (25), i.e. (SS)(SS)(xi)(SS), is (SS)(SS)(xi)(SS) rather than (SS)(SS)(xi)(SS). This restriction is evidently insufficiently handled by the ‘\textit{xi}’-grammar reached in (39). We propose that the inability of (xi) to head a PhP can be accounted for by the addition of yet another positional faithfulness constraint NO-SPREAD$_{\text{PhP}}$ to the ‘\textit{xi}’-grammar developed so far as the high-ranking one. Similar to MAX-µ$_F$, this constraint is a position-specific version of a more general faithfulness constraint, i.e. NO-SPREAD, and can be formulated as:

\begin{equation}
\text{NO-SPREAD}_{\text{PhP}}
\end{equation}

Any association line of the PhP-initial segment that is present in the output have a correspondent in the input.

Or more formally as (cf. (29)):

\begin{equation}
\text{Let } S_1 \text{ and } S_2 \text{ stand for two related phonological representations, where } \\
\mu_1 \text{ and } \text{seg}_1 \in S_1, \\
\mu_2 \text{ and } \text{seg}_2 \in S_2, \\
\mu_1 R \mu_2, \text{ and} \\
\text{seg}_1 R \text{seg}_2, \\
\text{if } \mu_2 \text{ is associated with seg}_2, \text{ and} \\
\text{seg}_2 \text{ is PhP-initial,} \\
\text{then } \mu_1 \text{ is associated with seg}_1.
\end{equation}

The ‘\textit{xi}’-grammar is thus:

\begin{equation}
\text{NO-SPREAD}_{\text{PhP}}, \text{ MAX-µ}_F, \text{ NOEMPTYMORA} \gg \text{NO-SPREAD} \gg \text{MAX-µ}.
\end{equation}

Under this revised ‘\textit{xi}’-grammar, the optimal form for a PhP initial ‘\textit{xi}’ is the zero parse ($\phi$), which refers to the non-surfacing of the underlying form (Kager 2000; referred to as ‘null parse’ in Prince and Smolensky 1993). In other words, (xi), albeit legitimate, can never occur PhP-initially. This is illustrated below (cf. (40)):

\begin{raggedright}
$^{16}$Informally, one might suggest that the reason for this has something to do with the fact that although ‘\textit{xi}’, when occurring alone, surfaces as bimoraic, this weight nonetheless results from the linking of the underlying empty mora to the segment i, and as such is ‘acquired’. This acquired weight is somewhat not as strong as the ‘innate’ weight of full lexical syllables which have two filled moras underlingly. Furthermore, given the argument that prosodic units at all levels of the prosodic hierarchy of Chinese are trochaic, the PhP- and IP-initial positions are evidently very strong. Hence the monosyllabic foot (xi), with its ‘acquired’ weight from the empty mora, is somehow not as strong as the monosyllabic foot formed by a full lexical syllable, and as such cannot head a PhP or an IP.
\end{raggedright}
Evidently, the addition of this new positional-specific \( \text{NO-SPREAD}_{\text{PhP}} \) only affects the parsing of PhP-initial ‘\( xi’ \) (including IP-initial one, of course) and has no bearing on the parsing of ‘\( xi’ \) elsewhere.

### 3.3.3.2 More discussion on the ‘\( xi’\)-grammar

The ‘\( xi’\)-grammar developed in (45) calls for more discussion. First, in retrospect, this grammar is configured in such a way that the two position-specific faithfulness constraints \( \text{NO-SPREAD}_{\text{PhP}} \) and \( \text{MAX-} \mu \) respectively dominate their position-neutral counterparts \( \text{NO-SPREAD} \) and \( \text{MAX-} \mu \) with the markedness constraint \( \text{NOEMPTY}_{\text{MORA}} \) ‘sandwiched’ in between. This ranking scheme is compatible with the general pattern formulated in Kager (1999: 408):

\[
\text{IO-Faithfulness (prominent positions)} \gg \text{Markedness} \gg \text{IO-Faithfulness (general)}
\]

which recurs in many OT grammars cross-linguistically, e.g. the grammar for Shona vowel harmony (Beckman 1997a) and that for the distribution of complex codas in Tamil (Beckman 1997b)\(^{17}\).

Second, compare the phonological representations of normal interjection syllables, ‘\( xi’\) and normal lexical syllables, and we notice an intriguing pattern of moraicity gradience: in the above order, the representations are respectively one (filled) mora (as argued in section 2.3.4.1 of Chapter 2), one (filled) mora and an empty one, and two (filled) moras. In terms of syllable weight, this means that in Chinese, whereas normal lexical syllables are heavy and normal interjection syllables are light, ‘\( xi’\) lies in between and may be suggested to be ‘semi-light’ (or ‘semi-heavy’ for that matter).

Third, a natural question is why ‘\( xi’\) is so special with this extra bit of phonological weight. The reason is not completely clear at this moment, but we believe that it may well be, at least partly, traced back to its unique origin: on top of the features associated with interjections in general (such as semantically empty and emotionally laden), ‘\( xi’\) is distinctly characterized by a significant singing element inherited from

\(^{17}\) One might also suggest that the dominance of \( \text{MAX-} \mu \)\(_{\text{f}} \) over \( \text{MAX-} \mu \) exemplifies the ranking between two constraints in a Pāṇinian relationship.
the folk song of the State of Chu. Indeed, some scholars have gone so far as to suggest that ‘xi’ is a linguistic unit specially reserved to simulate the drawl in singing (Legge 1871; Chen 1994) while others treat it as a mere ‘breather’ or ‘breathing sound’, which is an ‘otherwise meaningless sound that originally adapted the lyrics to the melody of the song’ (Field 1986). We suggest this musical element might furnish ‘xi’ with some additional weight, and renders it heavier than normal interjection syllables; however, this acquired weight is not so much as to elevate it on a par with full lexical syllables. As a consequence, it falls in between and cannot head a disyllabic foot. Furthermore, the monosyllabic foot (xi) is not strong enough to head higher-level prosodic units such as a PhP and an IP, as is a monosyllabic foot formulated by a full lexical syllable.

Fourth, as pointed out in Section 2.3.4.1 in Chapter 2, ‘xi’ also occurs in Shijing, but interestingly, there it exhibits no such uniqueness and behaves just like other interjection syllables. The noteworthy point is that modern speakers again are well aware of this subtle yet important difference and treat ‘xi’ in Shijing differently from in Jiuge, as is evidenced in their different scansion of lines containing ‘xi’. We assume that ‘xi’ in Shijing and Jiuge are one and the same lexical entry in the modern speaker’s lexicon and share one underlying representation; the difference only lies in the ‘xi’-grammars respectively responsible for the surface representation of the Shijing ‘xi’ and the Jiuge ‘xi’. More specifically, we argue that the parsing pattern of the Shijing ‘xi’ results from the following ‘xi’-grammar operative for its surface realization:

\[(48) \text{NOEMPTYMORA, NO-SPREAD} \gg \text{MAX-µ}\]

Compare this to the ‘xi’-grammar for the Jiuge ‘xi’ in (45), and it is notable that the positional faithfulness constraints \(\text{NO-SPREAD}_{\text{PhP}}\) and \(\text{MAX-µ}\) are not operative for the Shijing ‘xi’. As a result, the Shijing ‘xi’ always surfaces as monomoraic, lacking the position sensitivity characterizing the parsing of Jiuge ‘xi’. Consequently it can only serve as the non-head of a disyllabic trochee, as shown below:

---

18 Indeed, ‘xi’ is also present in a handful of earlier poems of Guti, the genre following Chuci. And interestingly there, the modern speaker also treats it merely as a normal interjection syllable and parses it accordingly (see Chapter 4). In other words, Chuci is the only genre where ‘xi’ acquires this extra bit of phonological weight; once out of this genre, it is deprived of this phonological privilege and ‘back to normal’. Needless to say, that the modern speaker is able to treat the same lexical item ‘xi’ differently when it occurs in different genres crucially rests upon the fact that she is able to tell which genre a given line is of; largely thanks to the distinct characteristics associated with each genre.

19 As the development of this side-grammar is similar to (though considerably less complicated than) that of the side-grammar for the Jiuge ‘xi’, it is not belabored here.
Fifth, so far we have dealt with three types of feet containing ‘xi’: (Sxi), (xi) and (xiS), the former two being legitimate whereas the latter illegitimate. In view of the weight scalarity mentioned above, we might wonder about the well-formedness of feet comprised of ‘xi’ and a normal interjection syllable. Still using S_i to stand for the interjection syllable, what is at issue here is the legitimacy of (xiS_i) and (S_i xi). Neither of these combinations are present in our corpus, and we suggest that this absence is principled: both turn out to be ill-formed under the ‘xi’-grammar developed above, as shown below (the normal interjection syllable is exemplified with the interjection ‘yi’):
In these two cases, the feet resulting from the surface realization of ‘xi’ respectively have the quantitative structure of (LL) and (LH), and hence are illicit.

Sixth, it is important to point out that similar to the constraint GOODFTINTERJ proposed in the Shijing sub-grammar, GOODFT’XI’ is, in theory, also reformulable in terms of more universal constraints; GOODFT’XI’ is used as a portmanteau constraint merely for convenience sake to tuck away the constraints into which it can be decomposed. Specifically, now that we have argued for the phonological representation of ‘xi’ as a filled mora plus an empty one, what is encapsulated in GOODFT’XI’ is, on the one hand, the ‘xi’-grammar argued for here (which yields the optimal surface forms of ‘xi’), and on the other hand, the two bedrock constraints for good and bad foot structures in Chinese, namely, RhTYPE=TROCHEE and STRESS-TO-WEIGHT (which checks the well-formedness of the foot containing ‘xi’; see Section 2.3.4.1 of Chapter 2). All the four constraints in the sub-grammar, i.e., MAX-μ], NOEMPTYMORA, NO-SPREAD and MAX-μ, as well as the latter two constraints, are all universal ones well-attested cross-linguistically; in addition, as shown in (47), the ranking among the four constraints constituting the ‘xi’-grammar also follows a general pattern of interaction that holds across languages between positional faithfulness constraints, markedness ones, and general faithfulness ones.

Seventh, it needs to be realized that this ‘xi’-grammar is actually ‘the grammar behind GOODFT’XI’”, and GOODFT’XI’ effectively encapsulates it. Expositiorily, similar to the treatment of the sub-hierarchy for the delimitation of the PhP boundary in Chapter 2, the ‘xi’-grammar is also ‘folded away’ in the constraint GOODFT’XI’ in the sub-
grammar. Indeed, in the sense that the phonological parsing of the verse line containing ‘xi’ must conform to the well-formedness of feet containing ‘xi’ rather than the other way around, we might argue that this ‘xi’-grammar, if integrated into the sub-grammar, must be high-ranking, which is reflected in the high-ranking of GOODFT’XI’.

Finally, let us reconsider the optimal parsing in (25). That the PhP-level parsing is (SS)(SS)(xi)(SS) rather than (SS)(SS)(xi)(SS) offers crucial evidence for the dominance of the sub-hierarchy for PhP boundary delimitation, i.e. BINARITY >> EVENNESS >> LONG-LAST, by the sub-grammar responsible for foot-level parsing, where the ‘xi’-grammar is encapsulated in the inviolable GOODFT’XI’\(^{20}\). This is shown below. For simplicity sake, we only show the different PhP parsing of the same foot parsing (SS)(SS)(xi)(SS).

3.3.4 8-syll lines and 9-syll lines

The above discussion on the phonological representation of ‘xi’ is triggered by the parsing of ‘xi’ exhibited in the scansion of 5-, 6- and 7-syll Jiuge lines. The sub-grammar developed so far, as presented in (22), is represented in a linear form as:

\[
\text{BINMAX, *IP-FINAL-MONOFT, GOODFT’XI’ >> ANCHOR-OI >> BINMIN >> ANCHOR-IO.}
\]

As stated earlier, 5-, 6- and 7-syll lines comprise the bulk of the present corpus; still, there remain one 8-syll line and two 9-syll lines in the corpus to be analysed. As is shown below, their scansion can all be adequately handled by the sub-grammar just presented.

3.3.4.1 8-syll lines

The single 8-syll line is of the grammatical structure [S[SS]]xi[[SS][SS]] and optimally scanned as (S)(SS)(xi)(SS)(SS). This is presented below:

\(^{20}\) This argument is based on the assumption that constraints in each of these two hierarchies cluster together and do not intermingle.
The sub-grammar correctly predicts (S)(SS)(xi)(SS)(SS) as the optimal scansion:

<table>
<thead>
<tr>
<th>(S)(SS)(xi)(SS)(SS)</th>
<th>BINMAX</th>
<th>*IP-FINAL-MONOFT</th>
<th>GOODFT’XI’</th>
<th>ANCHOR-OI</th>
<th>BINMIN</th>
<th>ANCHOR-IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (SS)(xi)(SS)(SS)</td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. (SS)(xi)(SS)(SS)</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>c. (SS)(xi)(SS)(SS)</td>
<td>#!</td>
<td>#</td>
<td></td>
<td></td>
<td>#</td>
<td>#</td>
</tr>
<tr>
<td>d. (S)(SS)(SS)(SS)</td>
<td>#!</td>
<td>#</td>
<td></td>
<td></td>
<td>#</td>
<td>#</td>
</tr>
<tr>
<td>e. (SS)(xi)(SS)(SS)</td>
<td>#!</td>
<td>#</td>
<td></td>
<td></td>
<td>#</td>
<td>#</td>
</tr>
<tr>
<td>f. (SS)(xi)(SS)(SS)</td>
<td>#!</td>
<td>#</td>
<td></td>
<td></td>
<td>#</td>
<td>#</td>
</tr>
</tbody>
</table>

3.3.4.2 9-syll lines

Two 9-syll lines appear in our corpus, which are respectively structured as [S][S][S][i][S][S][S][S]] and [S][S][S][i][S][S][S][S]], and optimally scanned as (S)(SS)(xi)(SS)(SS)(SS) and (SS)(SS)(xi)(SS)(SS)(SS). The two lines are given below:

(56)  [qi1 [bu2 xin4]] xi1 [gao4 yu2] [yi3 [bu4 xian2]]
promise not keep xi accuse me because not leisure
‘Ah, the promise is not kept and I was accused of not being leisurely’
⇒ (qi1) (bu2 xin4) (xi) (gao4 yu2) (yi3) (bu4 xian2)

(57)  [yu2 [chu3 [you1 huang2]]] xil [zhong1 [bu2 [jian4tian1]]]
I stay gloomy bamboo xi ever no see sky
‘Ah, I stay in the gloomy bamboo forest, and never see the sky’
⇒ (yu2 chu3) (you1 huang2) (xi) (zhong bu) (jian tian).

Again both optimal scissions are well predicted by the sub-grammar, as shown below:

Note that candidate (d) has a trisyllabic foot containing ‘xi’ that needs to undergo evaluation by GOODFT’XI’. So far we have been solely concerned with the well-formedness pattern of monosyllabic or disyllabic feet containing ‘xi’. The sub-grammar leads ‘xi’ to surface as a bimoraic syllable in this context of (SSxi) (the tableau is omitted here, similar to (40). Thus, the trisyllabic foot has a quantitative structure of (HHH), which is well-formed when trisyllabic feet are allowed, i.e. in prose scansion.
Thus, the sub-grammar developed so far is adequate to account for the modern speaker’s scansion of all Jiuge lines. It needs to be pointed out here that although ALIGNR (FT, IP) apparently has no bearing on the scansion, it must be dominated by ANCHOR-IO which so far ranks the lowest:

(60)

Indeed, that ALIGNR (FT, IP) plays no active role in selecting the winner is exactly because of its ranking as such\(^\text{22}\). Thus, with this new addition, the grammar is:

(61)

\(^\text{22}\) For a fully articulated definition of the notion of ‘(in)active’, see Prince and Smolensky (1993: 107).
For simplicity sake, ALIGNR (FT, IP) has been omitted in the tableaux so far due to its inactiveness, but as we will see below, it does become crucial in accounting for the metrical harmony judgment.

### 3.4 Formal grounding of metrical harmony

Similar to the organization of Chapter 2, this section is devoted to a formal account of the metrical harmony. We are going to argue that as in the case of the Shijing lines, the modern speaker’s metrical harmony judgment of Jiuge lines can be formally grounded in the Jiuge sub-grammar. The analytical procedure is identical with that in Section 2.4 of Chapter 2 and will not be repeated here. Furthermore, 8- and 9-syll lines will be omitted due to their negligible percentage (respectively 1 and 2 out of 253 lines) and only 5-, 6- and 7-syll lines will be examined below.

To begin with, as mentioned in Section 3.3.1, all the 5-syll Jiuge lines have the same grammatical structure [SS]xi[SS] and scansion (SS)(xi)(SS). Apparently, the tableau des tableaux would consist of only one candidate parse, which is trivially the optimal one. At the same time, there is converging native judgment on the metrical harmony of such 5-syll lines: they are all experienced to be metrically very harmonious. Thus, in this case, the metrical harmony can be trivially grounded in the sub-grammar.

Alternatively, we can show the formal grounding of the metrical harmony judgment in the grammar as the OT harmony by constructing some hypothetical grammatical structures and then elicit the native metrical harmony judgment on the one hand, and examine the formal harmony on the other. For example, consider the constructed line of the grammatical structure [S]xi[S[SS]] in imitation of line (13) (iii) above:

(62) \[
\begin{align*}
\text{person} & \text{ xi} \text{ have beautiful children} \\
\text{'The person has beautiful children'}
\end{align*}
\]

for which I elicited from my informants the optimal scansion (ren2 xi1) (you3) (mei3 zi3) and the judgment that it is ‘not as harmonious as [SS]xi[SS]). The following tableau des tableaux is therefore constructed:

<table>
<thead>
<tr>
<th>Candidate parses</th>
<th>BIN MAX</th>
<th>*IP- FINAL-MONOFT</th>
<th>GOODFT’ Xi’</th>
<th>ANCHOR-OI</th>
<th>BIN MIN</th>
<th>ANCHOR-IO</th>
<th>ALIGNR (FT, IP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ([S]xi[S[SS]] )</td>
<td>(SS)(xi)(SS)</td>
<td>*</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. ([S]xi[S[SS]] )</td>
<td>(Sxi)(S)(SS)</td>
<td>*</td>
<td>*!</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

23 The interesting question to pursue here is why 5-syll Jiuge lines in the corpus only exhibit this grammatical structure, and all the other grammatical structures fail to appear. Evidently, this question necessarily brings us to look into the ancient side of the picture. As is to be argued in Zuo (in preparation), this is because the ancient verse grammar entertained by the ancient poet encourages the maximal mapping of the boundaries between the grammatical and the prosodic structures. Indeed, compared with other line types where the grammatical structures serving as the input of the sub-optimal parses merely occur with a low frequency, the 5-syll Jiuge lines might be considered an extreme case where the grammatical structures in the sub-optimal parses occur with a zero frequency.

24 Note that ALIGNR (FT, IP) is included in the tableaux des tableaux in this section.
This indeed reveals that the parse corresponding to the grammatical structure that actually occurs in the corpus best satisfies the constraints and as such has the most OT harmony. Actually, it is as good as a 5-syll line can be in that it incurs the ‘minimal violation’ of BINMIN, which is inevitable given that the line contains an odd number of syllables. In contrast, the parse corresponding to the constructed grammatical structure is suboptimal.

Now consider 6-syll lines. As mentioned in Section 3.3.2, there are two grammatical structures, i.e. [[SS]S][xi][SS] and [S][SS][xi][SS], corresponding to two parses. Given their respective optimal scansion being (SS)(Sxi)(SS) and (S)(SS)(xi)(SS), the tableau des tableaux is constructed below:

(64)

<table>
<thead>
<tr>
<th>Candidate parses</th>
<th>BIN MAX</th>
<th>*IP-FINAL-MONOFT</th>
<th>GOODFT' XI'</th>
<th>ANCHOR-OI</th>
<th>BIN MIN</th>
<th>ANCHOR-IO</th>
<th>ALIGNR (FT, IP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ☞ [[SS]S][xi][SS] (SS)(Sxi)(SS)</td>
<td>*</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. ☞ [S][SS][xi][SS] (S)(SS)(xi)(SS)</td>
<td>*!</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Of the two grammatical structures, lines of the [[SS]S][xi][SS] are felt to be metrically more harmonious by the native speaker. As shown here, the parse corresponding to this grammatical structure is optimal under the Jiuge sub-grammar. Thus, again the metrical harmony can be formally grounded in the sub-grammar via the construct of OT harmony.

For 7-syll Jiuge lines, six grammatical structures occur in the corpus, which give rise to six candidate parses, and given the optimal scansion by the ancient speaker for lines of each grammatical structure, the tableau des tableaux is constructed below:

(65)

<table>
<thead>
<tr>
<th>Candidate parses</th>
<th>BIN MAX</th>
<th>*IP-FINAL-MONOFT</th>
<th>GOODFT' XI'</th>
<th>ANCHOR-OI</th>
<th>BINMIN</th>
<th>ANCHOR-IO</th>
<th>ALIGNR (FT, IP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>b. ☞ [[SS]S][xi][S][SS] (SS)(Sxi)(S)(SS)</td>
<td>*</td>
<td>*</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. ☞ [SS][xi][S][SS] (S)(SS)(xi)(SS)</td>
<td>*</td>
<td>*</td>
<td>11!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. ☞ [S][SS][xi][SS] (S)(SS)(xi)(SS)</td>
<td>!</td>
<td>*</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. ☞ [S][SS][xi][SS] (S)(SS)(xi)(SS)</td>
<td>**!</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. ☞ [S][SS][xi][SS] (S)(SS)(xi)(SS)</td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The sub-grammar yields two equi-optimal parses (b) and (f), when the fact of the matter is that only the input structure of the parse (b), i.e. [[SS]S][xi][S][SS] is cognized by the modern speaker as the most harmonious. It is notable that ALIGNR (FT, IP), hitherto inactive, becomes crucial in sifting parse (c) out. To discriminate (f)
from (a) presents a similar scenario to the case of 6-syll, Shijing lines (cf. (95) and (96) in Chapter 2) where parse (f), with its PhP boundary after the first foot, can be winnowed out by the PhP boundary delimitation sub-hierarchy BINARITY >> EVENNESS >> LONG-LAST.

(66)

<table>
<thead>
<tr>
<th>Candidate parses</th>
<th>BINARITY</th>
<th>EVENNESS</th>
<th>LONG-LAST</th>
</tr>
</thead>
<tbody>
<tr>
<td>b.</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>f.</td>
<td><em>!</em></td>
<td>***</td>
<td></td>
</tr>
</tbody>
</table>

This way, parse (b) emerges as the only winner, which coincides with the modern speaker’s metrical harmony judgment. This is illustrated below. As argued in (52), the PhP boundary delimitation hierarchy ranks lower than the foot-level parsing hierarchy.

(67)

<table>
<thead>
<tr>
<th>Candidate parses</th>
<th>BINMAX</th>
<th>*IP-FINAL-</th>
<th>GOODFT-</th>
<th>ALIGNR (FT)</th>
<th>BINARITY</th>
<th>EVENNESS</th>
<th>LONG-LAST</th>
</tr>
</thead>
<tbody>
<tr>
<td>b.</td>
<td></td>
<td>*</td>
<td>*</td>
<td>10</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f.</td>
<td></td>
<td>*</td>
<td>*</td>
<td>10</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To conclude, the metrical harmony judgment for the Jiuge lines can be formally grounded in the corresponding sub-grammar, which necessarily includes both the foot-level and PhP-level parsing hierarchies.