Chapter 1 Introduction

This introductory chapter sets the research goals of the present study (Section 1.1), provides its theoretical underpinnings (Section 1.2), presents its analytical scheme (Section 1.3), describes the corpus (Section 1.4) and outlines the organization of the following chapters (Section 1.5).

1.1 Research goals

This study aims to achieve two goals: first, to develop a grammar for the modern speaker's scansion of classical Chinese verse lines; second, to formally account for the metrical harmony of verse lines cognized by the verse reader. These two goals are intimately interrelated: specifically, as we will argue in the following chapters, the cognitively oriented notion of metrical harmony can be formally grounded in the grammar developed in reaching the first goal.

In elaborating on the goal of developing the modern speaker's scansion grammar (to be simply referred to as the modern grammar below), two points deserve mentioning to begin with. First, the present study is confined to the modern speaker's scansion of verse lines, which accordingly constitute its analytical domain. Phonological issues at higher levels of verse organization such as couplet and stanza are not covered. Second, this study adopts the unique perspective of examining the modern speaker's scansion of the ancient corpus, which is entailed by the fact that classical Chinese verse enjoys great popularity with the modern speaker.

The vast reservoir of classical Chinese verse, spanning a period of over 2,000 years (ca. 1,000 BC – 1,200 AD), is typically divided into five genres, namely, *Shijing*, *Chuci*, *Guti*, *Jinti* and *Ci* (see e.g. Frankel 1972). The modern speaker's scansion of verse lines is characterized by uniformity and diversity simultaneously. On the one hand, strong consistency is observed in his[1] scansion of lines from different genres, for example, the unambiguous preference for binary feet, among other things. On the other hand, lines from different sub-genres might be scanned in subtly different manners: this is most evident in the scansion of lines of comparable structures from different genres. For example, the following *Shijing* line[2]

(1) *[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'

is scanned as (S)(SS)(SS)(SS), where S stands for the syllable. By comparison, the following *Jinti* line of an identical structure

(2) *[wei4[ta1 ren2]] [zuo4 [jia4 [yi1 shang3]]]*
for other people make wedding garment dress
'(She) makes wedding garments for other girls'

[1] For simplicity sake, 'his' instead of 'his/her' is used as a generic reference pronoun in this study.
[2] The translation of verse lines in this study is respectively based on Legge (1871) (for *Shijing*), Yu (1949) (for *Jiuge*), Bynner (1929) (for *Jinti*) and Watson (1984) (for *Guti* and *Ci*).

is scanned differently as (SS)(SS)(S)(SS).

This dual characteristic suggests that the development of the modern grammar represents a scenario where the façade of superficial variation needs to be penetrated to uncover the underlying uniformity, and at the same time, the variation must be systematically captured. Accordingly, the grammar must crucially entertain a certain degree of core stability and inherent flexibility in order to accommodate both aspects[3]. The scansion of verse lines from each genre is captured by a sub-grammar (following Anttila's (1995) use of the term), whose delicate difference accounts for the different scansions; crucially the coexisting sub-grammars are minimally different so as to be unifiable into one single grammar.

Therefore, the first goal is tantamount to the development of a unified modern grammar, and the exploration of its instantiation into different sub-grammars. Optimality Theory (OT) provides an elegant framework to achieve this purpose and constitutes the analytical framework for this study. In OT terms, the grammar is necessarily a partial ranking on the set of constraints postulated to be operative in the modern speaker's scansion of classical Chinese verse lines. The partial ranking enables the grammar to be instantiated into multiple full rankings on the same set of constraints, which correspond to individual sub-grammars. Analytically, the reverse route is taken: a sub-grammar is first developed for each genre, and subsequently these individual sub-grammars are unified into the overarching grammar. The sub-grammars are developed in an incremental fashion by taking the grammatical structure[4] of the verse line as the input and the modern speaker's prosodic parsing of the line as the optimal output. More is to be said on the theoretical underpinnings and the analytical scheme respectively in Sections 1.2 and 1.3 below.

It is further observed that in scanning and performing verse lines, the native speaker may entertain certain judgments, especially regarding whether a line 'feels' smooth and melodic or rugged and jarring. Such judgments, which are especially unequivocal in the case of the most harmonious lines, actually reflect their cognization of the metrical harmony of verse lines. This observation, together with the belief that the native speaker's judgment of metrical harmony constitutes the 'readily observable abilities of experienced poetry readers' that must be accounted for by an adequate verse grammar (Halle and Keyser 1971:139) provides the motivation for the second research goal. We will argue that for every genre, the cognitively oriented notion of metrical harmony can be formally grounded in the corresponding sub-grammar and accordingly in the unified grammar. In addition to offering a formal account of the less tangible notion of metrical harmony, this also indicates the explanatory adequacy of the verse grammar.

[3] Here, we opt for the postulation of a single grammar (encompassing multiple sub-grammars) as it appears intuitively more appealing to conceive of one single grammar for the modern speaker, especially in light of the afore-mentioned strong uniformity exhibited in their scansion of lines across genres. Theoretically, it is also possible to adopt a multiple grammars model where the scansion of each genre is captured via a grammar. As is to be discussed in Section 1.2.2 below, which model is to be adopted carries little theoretical import.

[4] We use the expression 'grammatical structure' of the verse line as a general way to refer to its grammatically parsed structure. This structure is, in most cases, reducible to the syntactic structure of the line, although semantic and even pragmatic considerations may also play a role in certain cases. See Section 1.4 below.

As mentioned earlier, in developing a unified modern verse grammar, we are assuming the unique perspective of exploring how the modern speaker scans the ancient verse lines. Nonetheless, the fact that the poems were all composed in ancient times naturally suggests the relevance of a historical dimension. A separate chapter is devoted to the exploration of the ancient verse grammar where we argue, on the basis of the evidence from the corpus per se, that all the constraints that are deployed in the modern grammar also played a role in the ancient one. However, for the sake of clarity and focus, this discussion on the historical side is necessarily brief.

As a final word, it behooves us to explicitly mention what this study is not about. First, it is not a study of the meter of classical Chinese verse, although our findings offer valuable insight into this issue, which will be briefly addressed in the final chapter. Second, in developing a theory of the native speaker's phonological parsing of verse lines, it is not concerned with the phonetic side of the story, namely, the real-time performance of the verse lines. Third, it is not intended to be a comprehensive study on the phonological system of Chinese in general, either ancient or modern; rather only those issues pertaining to the discussion at hand are addressed, such as the phonological representation of interjection syllables and the prosodic hierarchy in Chinese, especially the delimitation of Phonological Phrase (PhP). Fourth, it is not a literary or functional study on the artistic value or aesthetic effect of the classical Chinese verse.

1.2 Theoretical background

As an OT analysis of verse scansion, this study draws upon the theoretical background from two main sources, i.e., metrics and OT, which are respectively presented below.

1.2.1 Theoretical framework for metrics

Two components comprise the metrics part of the theoretical framework: prosodic metrics and cognitive poetics. As mentioned above, the present study, with its central concern being the development of a formal scansion grammar rather than the investigation of meter or the metrical system of classical Chinese verse, is a phonological rather than a metrical study. Accordingly, these two metrical theories are not adopted in their entirety, and only those aspects pertaining to the present research goals are outlined below.

1.2.1.1 Prosodic metrics

The theory of Prosodic Metrics is developed in Golston and Riad (1994, 1995, 1997a, b) (also cf. Golston 1998; Helsloot 1995). The central thesis of the theory is to ground the study of metrics, including issues such as meter, versification, and metrical tension etc., solely in the formal systems of the ambient language, most notably, the prosodic system. Crucially, in doing so, the need to postulate a separate metrical hierarchy alongside the linguistic system is circumvented, and the analysis only employs universal prosodic constraints without recourse to language-specific or meter-specific constraints or any abstract metrical template. This is different from 'generative metrics' which invokes an independent metrical structure in addition to the prosodic system. Initially proposed in Halle and Keyser (1969, 1971) as an abstract, one-level scheme, this metrical structure is subsequently developed into a hierarchy comprising metrical constituents at different levels such as the metrical position, (metrical) foot, dipod, metron, colon, hemistich, line, couplet, and quatrain (Kiparsky 1975, 1977; Piera 1980; Prince 1989; Hayes 1989; Hayes and MacEachern 1998). According to Kiparsky (1977), the ‘basic metrical patterns’ are produced via ‘some combinatorial processes’ by an independent 'pattern generator' (albeit in an unspecified manner).

In this light, we suggest that prosodic metrics embodies more economy: it eliminates the need for a separate metrical hierarchy, treats poetry as essentially a special form of the ambient language, and espouses the reconciliation of metrics and phonology. As such it is appealing both analytically and conceptually, and is adopted in this study[5].

1.2.1.2 Cognitive poetics

Developed in Tsur (1977, 1992, 1998)[6] and originally referred to as a ‘perception-oriented theory of meter’, the theory of cognitive poetics introduces a cognitive angle into the verse study by granting central importance to the reader's verse performance and his cognitive experience involved therein, in particular the judgment of metrical harmony. Of its main proposals, two bear most closely on the present study.

First, verse performance is granted special emphasis in the present study, whose main goal is to develop a formal grammar for the modern speaker’s scansion. As is to be shown below, the speaker’s performance offers a critical point of departure for the analysis by enabling us to infer, on the assumption about the straightforward relation between the scansion and performance, the speaker’s scansion of the verse line, which constitutes the optimal output of the formal grammar to be developed.

Second, we import from cognitive poetics into the present study its proposal on metrical tension. This proposal holds that metrical tension is, rather than an abstract, isolated property inherent in a verse line, inextricably related to the reader’s dynamic experience of verse scansion and performance of the verse line. The introduction of the cognitive perspective into the current study is obvious in the formal grounding of the metrical harmony cognized by the reader in the grammar.

As in the case of prosodic metrics, both these two points from cognitive poetics reflect its fundamental difference from generative metrics. Specifically, while generative metrics shows more interest in verse texts than in verse performance in the same way generative linguistics is more interested in language form than language use, cognitive poetics shifts more focus to the reader’s cognitive experience in scansion and performance. Indeed, the importance of cognization in verse studies is convincingly argued in Attridge (1982, 1989) where meter is highlighted as a phenomenon in the perceptual domain with its basis in the human neuro-cognitive (as

[5] Here one might argue that in specifying the analytical domain to be the verse line, we are actually using the unit of line, which has been treated as a metrical unit, in our analysis. But at least two reasons might be cited against this: one is that in verse the line is always orthographically delimited, and the other is that the unit of line plays no role in the actual formulation of the grammar. Rather it is the prosodic counterpart of the line, namely, the Intonational Phrase (IP), which is built into the grammar.

[6] As Hayes (1988) points out, poetics is the study of literary works from the structural point of view adopted in linguistics, and as such encompasses metrics. However, apparently, poetics used in Tsur’s work is tantamount to the narrower field of metrics as his work is only concerned with poetry.

well as muscular) system. On the other hand, the emphasis of cognitive poetics on verse performance (and accordingly scansion) in addition to verse texts actually embodies Jakobson's (1960) insightful distinction between verse design and verse delivery, which are concerned with verse texts and the verse performance respectively.

1.2.2 Basics of Optimality Theory and OT approaches to variation

The formal analytical framework of this study is constituted by Optimality Theory (OT) as was first proposed in Prince and Smolensky (1993), McCarthy and Prince (1993a, b) and further developed in a plethora of subsequent literature[7]. Furthermore, as one of the goals of this study is to account for, in a unified way, the various scansions of lines from different genres by the modern speaker, the issue of how to build variation into an OT grammar also becomes foregrounded. Below we first briefly present the basics of OT. (For full discussions, see e.g. Archangeli and Langendoen 1997 and Kager 1999.) Thereafter the theoretical models that have been proposed to deal with variation in OT are outlined and the model adopted in this study specified.

1.2.2.1 OT basics

In the most general terms, OT is a constraint-based and output-oriented grammatical framework which defines Universal Grammar as 'a set of universal constraints and a basic alphabet of linguistic representational categories' (Kager 1999: 4). At its heart is the postulation that the grammar of a language is a set of violable, universal constraints ranked in a specific way, and that any surface form in the language, i.e., output, is the optimal form emerging from the resolution of conflicts between constraints -- optimal in the sense that it incurs minimal violation of the constraint ranking hierarchy that defines the grammar of the particular language in question. As such, the optimal output is also the 'most harmonic' with respect to the set of ranked constraints. As constraints are intrinsically in conflict, a surface form will necessarily fail to satisfy all the constraints of a language, yet still be optimal compared to the others that incur more serious violations.

More specifically, an OT grammar comprises of three components: GEN, CON and EVAL. GEN stands for Generator, which produces a (potentially infinite) set of output candidates for every possible input from the lexicon. These candidates are fed into the constraint hierarchy (CON), and the optimal candidate is selected by Evaluator (EVAL) which evaluates how each output candidate satisfies the ranked constraints in CON. EVAL operates on the principle of parallelism, i.e., optimal satisfaction is computed over the whole hierarchy and the whole candidate set in one single step. Two constraint families are distinguished: faithfulness and markedness, which are inherently competing and which every grammar must reconcile. The former is concerned with the correspondence between the input and the output, while the latter solely focuses on the structural well-formedness of the output candidate.

---

[7] This includes, for example, McCarthy and Prince (1994, 1995a, b, 1999) and the work by others as collected in the Rutgers Optimality Archive (ROA) at http://roa.rutgers.edu.

The operation of an OT grammar is expressed in the tableau form, as shown below. The postulated underlying form is given in the upper left cell of the tableau and output candidates are listed in the left column[8]. Across the top of the tableau are the constraints whose relative importance is indicated by the ranking; the higher a constraint is ranked, the further left it appears in the tableau. Constraints separated by a solid line are strictly ordered: the constraint to the left dominates the one to the right. Constraints separated by a dotted line are unranked with respect to one another due to either lack of evidence or lack of conflict. An asterisk in a cell indicates a violation of the constraint heading that column, and an exclamation mark following an asterisk indicates that this violation is 'fatal' by eliminating any chance for the candidate under consideration to be optimal. The shading of a cell indicates the irrelevance of the satisfaction/violation of the corresponding constraint to the selection of the optimal form. The selected optimal form, i.e., the surface form in the language, is indicated by the pointing finger ☞.

(3)

/input /	Constraint 1	Constraint 2	Constraint 3
Candidate 1	*!		
☞Candidate 2		*	*
Candidate 3		**!	*
Candidate 4		*	**!

Having thus sketched the most basic concepts and constructs of OT, we wish to add two brief points regarding the presentation used in this study. First, as was mentioned above, a dotted line between two constraints indicates the non-ranking between them. Such constraints are in what Prince and Smolensky (1993: 51) refer to as a 'non-crucial non-ranking' relation with each other. For such a pair of constraints a non-crucial dominance relation can always be arbitrarily and trivially assigned (cf. McCarthy and Prince 1993b: 67) for ease of discussion. Violations of the constraints in such a relation are equally offensive in selecting the optimal form. This is illustrated in the tableau form below:

(4)

/input /	Constraint 1	Constraint 2	Constraint 3
Candidate 1	*!		
Candidate 2		*!	*
☞ Candidate 3			*
Candidate 4		*!	**

Second, besides tableaux presented above, the Hasse graph (cf. McCarthy and Prince 1993b: 56) will also be used to represent the constraint interaction. In such a graph, the constraints are positioned according to their ranking in the hierarchy: the higher a constraint is in the hierarchy, the higher its position in the graph. The presence of a line linking two constraints (which are necessarily not presented at the same level in the graph) indicates the dominance of the constraint at the lower level by that at the higher level, while the absence of a line between two constraints (whose relative position in the graph is non-committal) indicates either the lack of evidence for any

[8] Obviously, given the theoretically infinite number of output candidates, only a subset of them can be presented in a tableau.

crucial ranking between them, or the lack of conflict between them. As such, while tableaux are good for illustrating the local ranking, the Hasse graph is particularly effective in illustrating the global picture of constraint interaction, especially when such interaction defies a linear representation.

1.2.2.2 OT models for variation

As mentioned in Section 1.1, the present study postulates one single unified modern verse grammar, which must crucially be able to accommodate both the uniformity and diversity exhibited in the scansion of classical Chinese verse lines from different genres. This bears on the issue of how to represent variation within the OT framework. On the basis of Anttila (2001), we suggest that in general, four OT models have been proposed to address variation: (i) multiple grammars (the generic model); (ii) partially ordered grammars; (iii) floating constraints; (iv) continuously ranking grammars[9]. Below they are respectively outlined and the one adopted in the present study presented.

To begin with, the multiple grammars model contends that variation arises from the competition of multiple grammars (Kiparsky 1993; Kroch 1994). In the OT framework, this model is known as constraint re-ranking, which holds that variation in surface forms is a function of variation in constraint ranking and that this is true both across and within languages. Re-ranking certain constraints will give rise to a different constraint hierarchy that is able to account for a different set of data and thereby capture attested variation (for application of this model, see McCarthy and Prince 1993: 66 on the grammars for different registers of pre-Classical Latin; also cf. Colina 1995; van Oostendorp 1997; Cassimjee and Kisseberth 1998; Walker 1997b).

It deserves mentioning that the multiple grammars model is a most generic model imposing little restriction on the permissible extent of the difference between the multiple grammars. In this light, both partially ordered grammars and floating constraints constitute special cases of the generic multiple grammars model (Anttila 2001). More specifically, the model of partial ranking[10] (Anttila 1995; Anttila and Cho 1998; Anttila 2000) crucially assumes that an OT grammar is a partial rather than a full ranking on some constraint set. Some constraints in the grammar are underspecified in their ranking with others, and consequently, such a grammar is translatable into more than one fully ranked constraint hierarchy, each selecting one optimal candidate. Thus, the grammar can yield more than one optimal form, all being surface outputs in the language.

On the other hand, the floating constraints model (Reynolds 1994; Nagy and Reynolds 1997) captures variation in terms of the ranking mobility of certain constraints. Similar to the partial ranking model, this model contends that an OT grammar is not necessarily a fully ranked hierarchy and postulates that some

---

[9] Anttila (2001) also mentioned two more models, i.e. tied violations and pseudo-optimality. Both are theoretically the most conservative in the sense that they entail no modification of the standard OT assumptions. However, both are exposed to have a number of conceptual and empirical weaknesses which render them inadequate to accommodate the variation data. As such they are not discussed here.

[10] In Anttila (2001), the partial ranking grammars model proposed for the Finnish genitive plural is renamed as 'stratified grammars model' following the terminology of Tesar and Smolensky (1995). The stratified grammars model, which consists of internally unranked strata of constraints strictly ranked with each other, is a special case of the partial ranking grammar.

constraints may have variable rankings within a certain range, i.e. float. The idea is conceptually simple and intuitively appealing, as depicted in Reynolds (1994: 116):

> … within a given language or dialect, it may be the case that a particular constraint X may be classified only as being ranked somewhere within a certain range lying between two constraints W and Z, without specifying its exact ranking relative to a certain other constraint Y (or constraints Y_1, Y_2, etc.) which also falls between W and Z. A graphic representation of such a variable constraint ordering is as follows:
>
> ……………ConX………………
> ConW >> $ConY_1$ >> $ConY_2$ >> … >> Con Y_n >> ConZ
>
> Here, the constraint (or constraints) which appears on the higher level in the representation is the FC [floating constraint], while those on the lower level are 'hard-ordered' or 'anchored' constraints. The range over which the FCs may extend is defined, not in terms of the constraints (W and Z) which the FC lies between, but rather in terms of the particular subset of fixed or anchored constraints (Y_1, Y_2, … Y_n) with regard to which the FC is considered to be unranked. In other words, the FC may be allowed to fall in any position with respect to its anchored subset – above Y_1, below Y_n, or at any point in between; this is the essence of the FC's relationship with its anchored subset of range.

Finally, the continuous ranking model (Boersma 1998; Boersma and Hayes 1999, 2001; Hayes 2000) views variation in surface forms as gradient well-formedness and attempts to analyze it in terms of probability under the conception that ranking is a gradient and quantitatively explicit phenomenon. In this model, constraints are assigned numerical ranking values on a continuous numerical scale. The grammar entails a 'stochastic candidate evaluation' (Boersma and Hayes 2001): at evaluation time, a random positive or negative value is temporarily added to the ranking value of each constraint, which results in varying actual ranking values, referred to as 'selection points'. The variable selection points are responsible for the generation of a range of optimal outcomes, therefore capturing the variation (for application, see Hahn 1998; Zuraw 2000; Hayes 2000).

While as acknowledged in Anttila (2001), comparison between these theoretical models has so far proven inconclusive and it remains unclear which of the models is superior, two points should be acknowledged. First, in introducing real-number ranking value, the continuous ranking model reflects some deeper difference from the other three. However, while this added power offers certain descriptive advantage, its necessity and possible consequences in other domains remain to be explored. Second, it might be suggested that the principal difference among models (i), (ii), and (iii) hinges upon the conception of whether an OT grammar is a full or partial constraint ranking. Specifically, the multiple grammars model straightforwardly builds upon the classic OT tenet that an OT grammar is a fully articulated ranking hierarchy of constraints, and in effect handles variation via different grammars, while the partial ranking and floating constraints models hold that an OT grammar may well be a partial ranking where some constraints whose ranking is unspecified. In addition, the multiple grammars model, being generic in nature, is suggested to be over-powerful and some restricted versions of it such as the partial ranking and floating constraints model discussed here appear sufficient.

Due to lack of evidence as to which model is superior, the choice between the models seems somewhat arbitrary. In this study we opt for the floating constraints model, largely for its formal simplicity and analytical facility, to account for the variation in the modern speaker's scansion of different genres of classical Chinese verse. Evidently, in making this choice, we are pronouncing our assumption that an OT grammar is a partial ranking on some constraint set. This implies that the grammar, containing a certain constraint whose ranking is underspecified, can be articulated into several full constraint rankings. This constraint will be argued to float, which provides the grammar with a certain degree of inherent flexibility. The floating constraint may land in different places, which we refer to as 'landing sites', along the ranking hierarchy comprised by the constraints with 'hard-ordered' ranking (Reynolds Ibid.) which we refer to as the 'ranking skeleton'. Its different landings give rise to different full rankings, referred to as sub-grammars, which nonetheless all correspond to one partially ranked grammar.

At the same time, it merits mentioning that while the floating constraints model as originally conceived of in Reynolds (1994) is potentially very powerful, we propose some restrictions on it on the basis of our findings, to be shown in Chapter 9. Briefly, while Reynolds (Ibid.) proposes no limit on the number or nature of floating constraints in the grammar and posits that a floating constraint may 'fall in any position with respect to its anchored subset [of constraints]', we will show that in the grammar developed in this study, the floating constraint is limited in its kind, number, and 'landing sites', thus rendering the floating constraints model more desirable.

1.3 Analytical scheme in the present study

In the floating constraints model, the grammar to be developed consists of a ranking skeleton comprised of constraints with fixed ranking and one floating constraint, whose limited possible landing sites are specified. As such, the grammar is constraining due to the maximally shared core ranking and the limitation on the number and landing sites of the floating constraint. At the same time, the floating constraint offers inherent flexibility to the grammar which is translatable into multiple full ranking hierarchies, i.e. sub-grammars, depending on where the floating constraint lands. Because only one constraint floats and its landing sites are restricted, the sub-grammars are minimally different.

The analytical scheme follows naturally from the floating constraints model. In delimiting the specific components of the sub-grammar, it needs to be borne in mind that the analytic domain in this study is confined to the verse line. Phenomena above the line level, albeit fascinating, lie beyond its focus. The input is constituted by the grammatical structure of the verse line[11], which is attributable to the conspiracy of a range of grammatical factors, the most important being the syntactic constituency.

[11] This is in fact a somewhat expedient way of looking at things: we may alternatively suggest that the grammatically unparsed verse line constitutes the input and that the cluster of constraints responsible for the grammatical parsing of the line are very highly ranked in the overall constraint hierarchy so that other potentially possible grammatical structures are thrown out and never get a chance to surface, leaving the one presented here as the only grammatical structure to be simultaneously evaluated by phonological constraints.

The output candidates are possible ways in which the given verse line may potentially be scanned, and thus are in theory infinite in number.

At this point, we wish to clarify our position regarding the dichotomy of scansion versus performance: briefly, scansion represents the prosodic parsing of the line while performance is the actual realization of this parsing into acoustic signals. As such, scansion is abstract and phonological in nature while performance is concrete and phonetic. Or in terms of the grammar versus production dichotomy (cf. Hale and Reiss 2000), scansion is the output of the grammar and performance that of the production system. Crucially, we assume a straightforward sequential relation between scansion and performance, which in effect entails such a relation between phonology and phonetics[12]. Thus, scansion, as the output of the phonological module, is directly fed into the phonetic module, and undergoes a series of presumably trivial, no-frills phonetic operations such as strengthening at the beginning of the phonological unit and lengthening at the end of it ('initial strengthening and final lengthening', cf. Beckman and Edwards 1990; Fougeron and Keating 1997), extra lengthening of monosyllabic feet[13], pitch accent alignment, and intonation interpolation, before being realized as the phonetic form which is the performance of the verse line actually heard.

This linear and straightforward relation between scansion and performance enables us to directly infer the abstract prosodic parsing of a given line from its empirically observable phonetic performance[14], and in view of the fact that the phonetic performance is the actual realization, this corresponding prosodic parsing is thus the optimal parsing. Specifically, the above-mentioned phonetic features serve as important cues for the abstract phonological structure, in particular the foot structure: the unit that is clearly set off by the initial strengthening and final lengthening

---

[12] In stating this assumption, we are fully aware of the ongoing debate about the interplay between phonetics and phonology (e.g. Flemming 1995; Hayes 1996; Steriade 1997; Myers 1997). However, it is not our intention to take position in this debate, and we believe the current assumption, which follows the classical generative assumption of a strict separation between phonology and phonetics (e.g. Chomsky 1981) and is based on the well-founded generalization on the phonetic realization of phonological structures, is sound, especially in the particular context of verse scansion and performance.

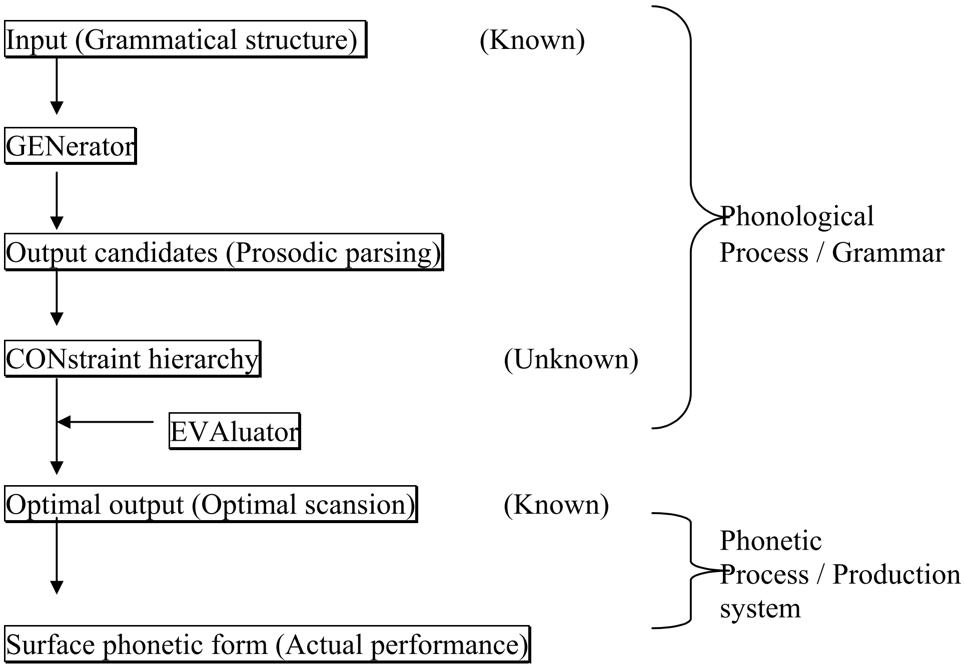
[13] See Hayes (1984) for argument that such lengthening is phonetic in nature. Also for this reason, the phonological construct of 'zero syllable' is dispensed with in our study.

[14] Three points are worth mentioning here. First, there are potentially an infinite number of performance styles (cf. Tsur 1998) for different readers and even for the same reader. However, this need not concern us, as crucially it is not the case that a verse line can be performed in any possible way; rather the many possible performance styles are subject to certain restrictions. Indeed, they may be regarded as the many ways of realizing one and the same phonological structure. Second, we exclude the scansion corresponding to the performance style in which the verse line is mechanically chopped into binary units from left to right, in a manner that is blind to the syntax or semantics of the line. Such a performance style is referred to as the 'minstrel' performance by C. S. Lewis in 'Performance' entry of *The New Princeton Encyclopedia of Poetry and Poetics* (1993) and as one type of 'divergent' performance in Tsur (1998). It is excluded because it is linguistically uninteresting. Third, there are indeed some cases, albeit rare, where the line can be performed in more than one way, each reflecting a different phonological structure, i.e., different scansion. However, in such cases, the different scansions are not equally preferred by the native speaker; one scansion is usually unambiguously more favored than the others, which, though possibly still acceptable, may be somewhat marginal. Our strategy in this study is to assume that for any given line, there is one and only one optimal scansion, which in such cases is constituted by the most preferred one. In this sense, our study is different from Anttila (1995) which deals with multiple surface forms of varying degrees of preferability.

constitutes the smallest prosodic unit above the syllable level, i.e., the phonological foot [15].

Therefore, the optimal foot structure, which is inferable in this fashion, is the winner out of the theoretically infinite number of output forms. So now what is known is the input, i.e., the grammatical structure of a line, and the optimal output, i.e., the optimal scansion by the modern speaker; what is to be developed is the constraint hierarchy, i.e., the sub-grammar, under which this optimal scansion inferred from the empirical performance is indeed selected as the optimal output for the given input. The analytical scheme is represented diagrammatically as follows:

(5)



In developing the sub-grammar for each genre, crucial data are cited to motivate the introduction of constraints and the ranking between them. Thereafter, under the assumption that the modern speaker entertains one overall grammar, the five sub-grammars are unified into one grammar via the constructs of floating constraint and ranking skeleton, as discussed above. The five sub-grammars might then be understood as the five instantiations of the overall grammar resulting from the specific landing sites of the floating constraint.

[15] One might further argue that this 'back-inference' of the foot structure from the surface phonetic structure is rendered particularly feasible on at least two accounts. One is that classical Chinese verse is normally performed at a rather slow rate, which serves to 'magnify' the phonetic cues, in particular the boundary lengthening, rendering it reliably present. The other is that due to the largely monosyllabic nature of classical Chinese, a foot can never cut into a word. This is because the foot, which consists of minimally one syllable and typically two, is co-terminous with one or two words. In the exceptional case of the few disyllabic morphemes, these two syllables invariably form one foot. This is quite unlike the case in English, where a word can be parsed into more than one foot. In this connection, it also needs to be mentioned that only foot structure can be legitimately inferred back from the phonetic performance; phonological structures at other levels such as phonological phrase (PhP) are less tractable to this procedure. As is to be seen in Chapter 4, the PhP boundary is delimited primarily through means other than from the phonetic cues and the phonetic cues only serve as supplementary evidence.

The formal account of the metrical harmony directly builds upon the grammar developed in this way. Specifically, it can be elegantly accounted for via the construct of 'tableau des tableaux' (Itô, Mester and Padgett 1995) which compares, under the modern verse grammar developed so far, the multiple parses, each constituted by the parse from a certain grammatical structure of the lines to the optimal scansion of such lines and selects the optimal parse[16]. We discover that for every line type in each genre, the grammatical structure in the optimal parse all corresponds to that of the lines cognized as metrically most harmonious. To cite the notion of 'OT harmony' gauged in terms of the constraint satisfaction/violation (Smolensky and Prince 1993), this optimal parse, which incurs the least violation of the sub-grammar, enjoys the greatest degree of 'OT harmony'. This shows that the metrical harmony judgment can be formally grounded as OT harmony in the verse grammar; in other words, the grammar can account for such judgments and is therefore explanatorily adequate.

It behooves us to quickly mention the analytical scheme for the ancient grammar. Evidently, once we venture into the historical side of the picture, the performance data is no longer available, and only the ancient corpus lies at our disposal. We will be arguing, by excavating a wide array of evidence from the corpus per se such as rhyming patterns, distribution of disyllabic morphemes, and frequency pattern, that the constraints deployed in the modern grammar also played a role in the ancient one.

As a final note, now that we have presented the analytical scheme in the OT framework, it is of interest to consider how the theoretical underpinnings of metrics discussed in Section 1.2 might chip in. First, as is to be seen in the following chapters, the sub-grammars as well as the grammar solely deploy universal prosodic constraints, which embodies the principle of prosodic metrics. Second, the tenets of cognitive poetics are reflected in the crucial role of the native speaker's cognization of the verse line in establishing the correlation between the metrical harmony cognized by the modern speaker and the OT harmony in terms of the constraint satisfaction/violation. At the same time, verse performance, which is of a central importance in cognitive poetics, also plays a vital role in enabling us to infer the optimal scansion.

1.4 Data of analysis

The data for this study comes from the corpus of verse lines, the native speaker's performance of them and their metrical harmony judgments.

The corpus comprises of 3,933 lines from the five genres of classical Chinese verse. The following table gives an overview of the distribution of the lines across the five genres.

(6)

Genre	Number of lines
Shijing	1320

[16] Such a tableau is called a 'tableau des tableaux' because each candidate form is constituted by a parse from the grammatical structure (which serves as the input in the tableau for developing the sub-grammar) to the corresponding optimal output (which is the optimal output in the tableau for developing the sub-grammar). As such it is actually constructed out of the many tableaux needed in the development of the sub-grammar.

Jiuge	253
Guti	843
Jinti	764
Ci	753
Total	3933

All the poems are randomly selected from ancient anthologies that are still enjoying great popularity with the modern speaker. With the exception of *Jiuge*, which is already a small corpus, for each of the other four genres, the odd-numbered poems are selected from the corresponding anthology. As such, the present corpus constitutes approximately half of the whole corpus of well-recited verse lines contained in the above-mentioned anthologies, and may thus be legitimately considered as significantly large. It also needs to be mentioned that while seeking to strike a balance across the genres, we have included slightly more *Shijing* lines and fewer *Jiuge* ones. This is because *Shijing* poems tend to be irregular in length, and can be quite long, while *Jiuge* is strictly speaking only a sub-genre of the *Chuci* genre which appears between *Shijing* and *Guti*, and the reason that only it is selected is that it best embodies the use of 'xi', the defining feature of the *Chuci* genre[17].

Evidently, this corpus consists of poems by various poets from different periods. It deserves mentioning that as the anthologies are collections of well-recited poems by many popular poets of each literary period, no single poet's work is disproportionately represented. Consequently, the poems in the present corpus are evenly distributed across the poets, and there are no poems whose inclusion, or exclusion, for that matter, would dramatically alter the constitution of the corpus.

The other part of the data is constituted by the native speaker's performance of verse lines and cognization of their metrical harmony. The former is empirically observable and straightforwardly gives rise to the scansion which serves as the optimal output of the verse grammar, and the latter can be directly elicited. Two things need to be mentioned. First, we are solely concerned with the performance (and accordingly scansion) of verse lines in Mandarin; performance (and scansion) in other Chinese dialects is a potentially intriguing topic (cf. Boyce's (1980) study of *Jinti* verse performance in the Min dialect), but not discussed in this study. We will be using Chinese to stand for Mandarin interchangeably throughout the study. Second, the pool of informants for this study comprises five native speakers of Chinese[18], who are all at or above university education level[19]. In virtually all cases, they entertain a strong

[17] Another reason that only *Jiuge* is selected is because of all *Chuci* sub-genres, it is best recited, largely because it contains relatively few arcane dictions.

[18] The following table provides further information on these five informants. For privacy consideration, only initials are used for identification.

Name	Age	Gender	Dialect background	Occupation
BM	30	M	Shenyang	Engineer
HH	32	F	Taiyuan	Chemist
XH	29	M	Suzhou	Architect
ZY	28	M	Bengbu	Physicist
QD	25	F	Zhangshu	Mathematician

[19] Under the educational system of China, the recitation of classical Chinese poetry is an integral part of the education program from the last year of elementary school all the way through the university. The informants have acquired competence for both scanning (as well as performing) the verse line and

consensus regarding both the optimal way to perform (and accordingly scan) a line and the judgment on its metrical harmony, such as whether a line is good or awkward, metrically smooth or rugged.

For the exploration of the ancient verse grammar, this corpus needs to be processed into the 'ripe' corpus. The processing mainly seeks to represent the grammatical structure of the verse lines by encoding, via a coding scheme, the boundary strength between two (surface) adjacent syllables. The boundary strength can be grounded in the formal grammatical parsing, in particular syntax, occasionally supplemented by semantics and lexicon. The encoding renders it easy to further process the corpus: in particular, lines are subsumed into different groups according to their coding profiles, and the frequency of each coding type is calculated. This way, the frequency patterns obscure in the raw corpus become highlighted and are ready to serve as evidence for the exploration of the ancient grammar. As the historical dimension is not the main focus of this study, the coding scheme for processing the corpus is relegated to Appendix II, where the resulting ripe corpus is also presented.

1.5 Structure of the dissertation

Chapters 2 to 6 are respectively devoted to analysis of the five genres and constitute the core of the study. Each chapter consists of two parts: the development of the sub-grammar for the genre under discussion and the formal grounding of the metrical harmony judgment for this genre in the corresponding sub-grammar just developed. Chapter 7 unifies the five sub-grammars into the overarching modern grammar. Chapter 8 briefly considers the historical dimension of the research, mainly discussing the relevance of the modern constraints in the ancient grammar. Additional issues such as the meter of classical Chinese verse are also addressed briefly. Appendix I presents the chronology of Chinese history and II provides the guidelines for corpus processing and the ripe corpus.

passing on judgments regarding its metrical harmony. It is a moot point as to whether the ability to scan and perform verse properly is acquired or learned; on the one hand, it is certainly somewhat different from the ability of speaking one's native language, but on the other hand, to the extent that the native speaker needs first to correctly parse and understand the verse line before scanning it properly, it might be suggested that the competence is, at least partially, innate and an integral part of the grammar after all. Indeed, verse scansion is rather comparable to prose scansion, perhaps except that it is somewhat artificial and calls for, in addition to the involvement of linguistic competence, more active and conscious deployment of one's rhythmical competence, which is arguably innate (Hayes 1989). The issue of 'learned vs. innate' is discussed in Jackendoff's (1989) work on a music grammar, following which we might suggest that in the case of a verse grammar, the constraints are innate, but the ranking needs to be learned via exposure to the data.