CHAPTER 2

LITERATURE REVIEW

The relationship between language and music has been discussed in two main aspects, namely their similarities and interactions. Several researchers (Lerdahl & Jackendoff 1983, Gilbers 1987, Gilbers and Schreuder 2002) observe that language and music share many features in common and thus some musical and linguistic theories are mutually applicable. In this chapter, I will first present the previous findings of the similar characteristics of language and music. Next, I will focus on the application of optimality theory in language-music connection, which is the topic of this thesis. In addition, the relevant knowledge of metrical phonology will also be introduced at the end of this chapter.

2.1 Similarities of Language and Music

2.1.1 Cognitive Processing

Borchgrevink (1982) points out that both prosodic and musical rhythm are organized in the speech hemisphere. Therefore, the connection between speech and

musical function is close. The musical and linguistic reliance on the same psychological mechanism results in many similar sub-functions. Lerdahl and Jackendoff (1983) regard the overlap of language and music as a result of a more general cognitive ability that deal with musical and linguistic symbols rather than coincidence. Besson and Schön (2001) hold the opinion that language and music are composed by components in different processing levels. They also found that general cognitive principles are involved in terms of syntactic processing in language and harmonic processing in music. The parallels between language and music are also confirmed by Levitin & Menon (2003), who state that musical structure is processed in "language" area of the brain. To investigate the neuroanatomical correlates of musical structure, they use functional magnetic neuroimaging (fMRI) and a stimulus manipulation regarding scrambled music. Through the brain responses toward classical music and the scrambled versions of the same piece, it is found that the area that is closely related to the processing of linguistic structure, Boadmann Area 47, has focal activation along with its right hemisphere homologue. Therefore, the regions processing fine-structured language and music should be the same.

The findings in the processing of language and music support their similarities in many aspects. The general cognitive ability provides a very possible explanation for their features in common. Among the so many similar characteristics, I will focus on

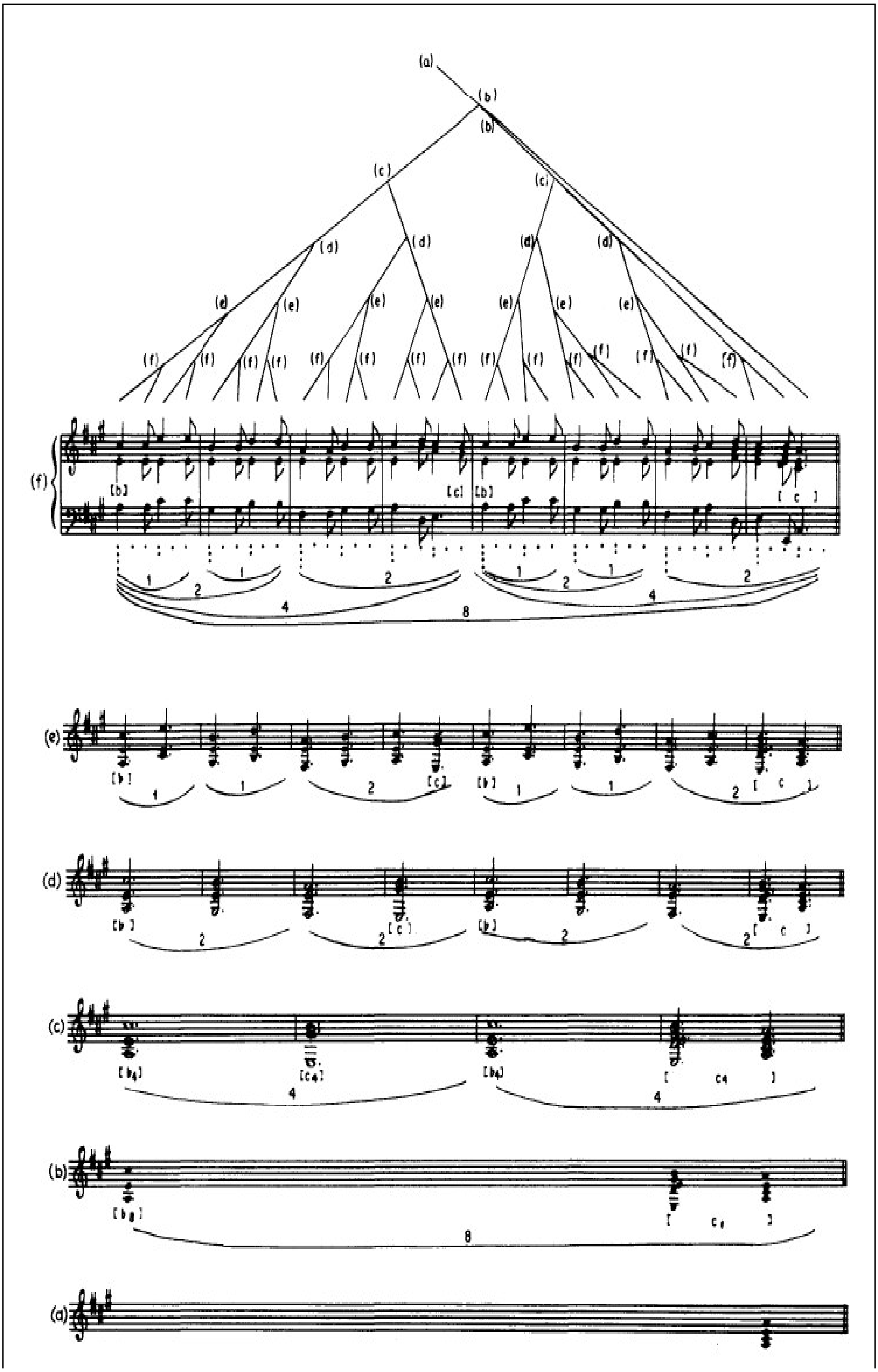
structure, rhythm, and pitch, which are the main tunnels of interactions in Jiang's verses.

2.1.2 Language and Musical Structure

The similarities of language and music mainly lie in their hierarchical structure. Lerdahl and Jackendoff (1983) develop the musical grammar and give rules for assigning analyses, i.e. structural descriptions. The features of creation and recursion are found to be shared by language and music. In their earlier investigation on time-span reduction in music and language, they found that not only language but also music is composed of recursive tree structure, as can be seen in (1).

(1) Music Structure Tree

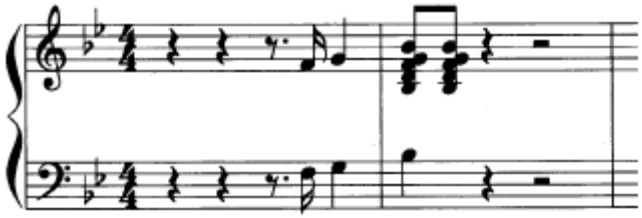
(Lerdahl and Jackendoff 1977, cited in Schreuder 2006: 22)



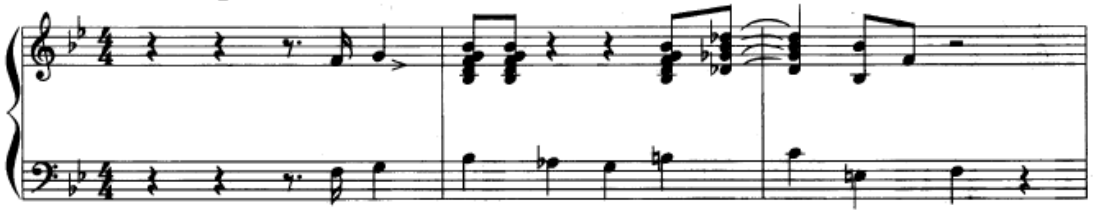
Sundberg and Lindblom (1991) indicate that both language and music are decomposable and constructed in a hierarchical structure. In linguistic and music theories, the structural events of language and music can be parsed into smaller domains. In terms of language, a sentence can be decomposed to several phrases which are divided into different words. Words are composed of syllables which are formed of one or more segments. Similarly, each musical domain can be divided into smaller domains which then decomposed into smaller domains. The smallest domain in music is motif. According to Ammer (2004), a motif is the shortest possible musical idea and can be as short as two notes. Several motifs form musical phrases and musical phrases constitute themes, which are the most recognizable parts of songs to listeners. Several themes or phrases build up musical sections, which are the largest components of a music work. Schreuder (2002) illustrates the decomposability of music by the example of “Tuxedo Junction.” As shown in (2a), the first two measures are the motif. Along with the third measure, a theme or a phrase is formed, as presented in (2b). The larger domain based on the theme is referred to as a musical section of the song, as (2c) indicates.

(2)

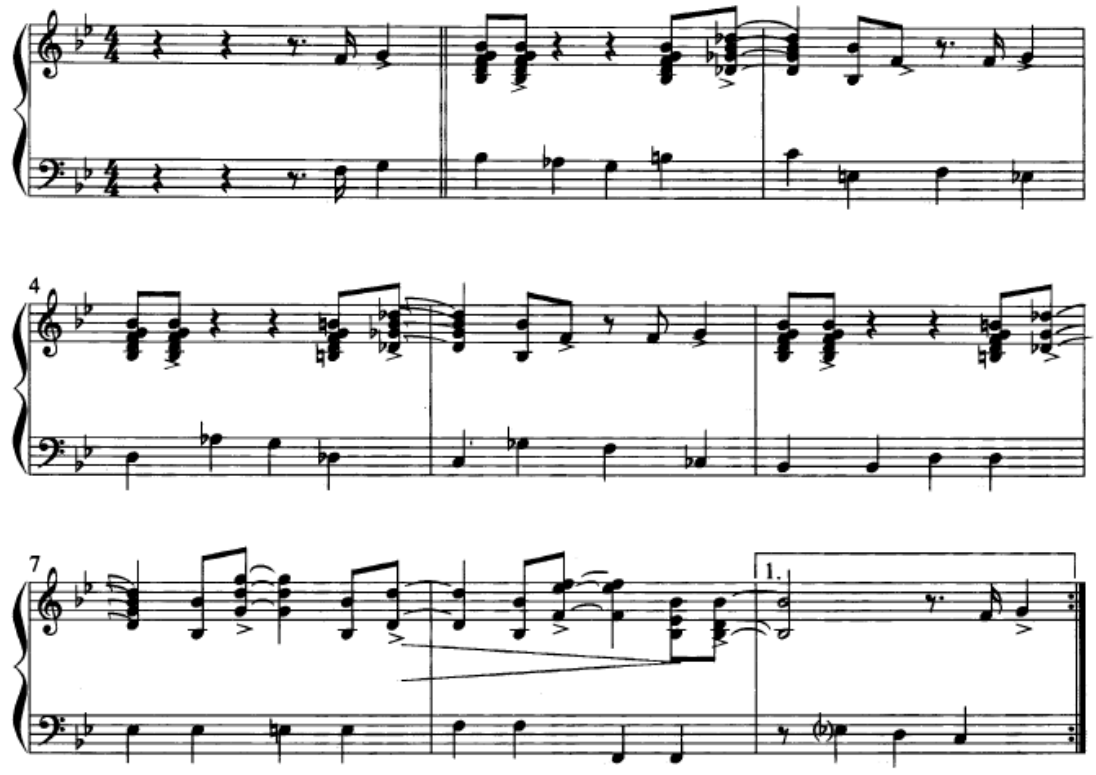
a. motif



b. theme or phrase



c. section

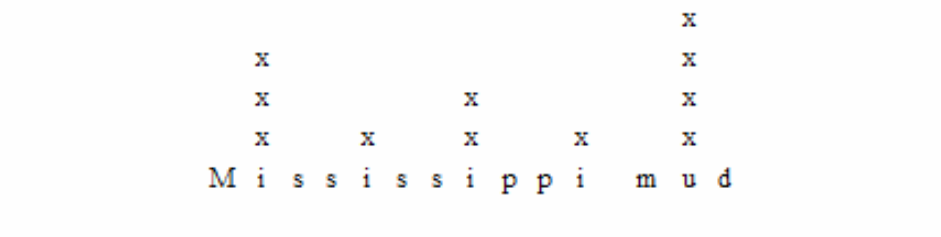


The previous studies on the similarities of music and language structures offer a solid basis for the investigation on the language-music interactions in songs. In this thesis, I will compare the musical and metrical trees of Jiang's works and exemplify how closely they are related.

2.1.3 Language and Musical Beats

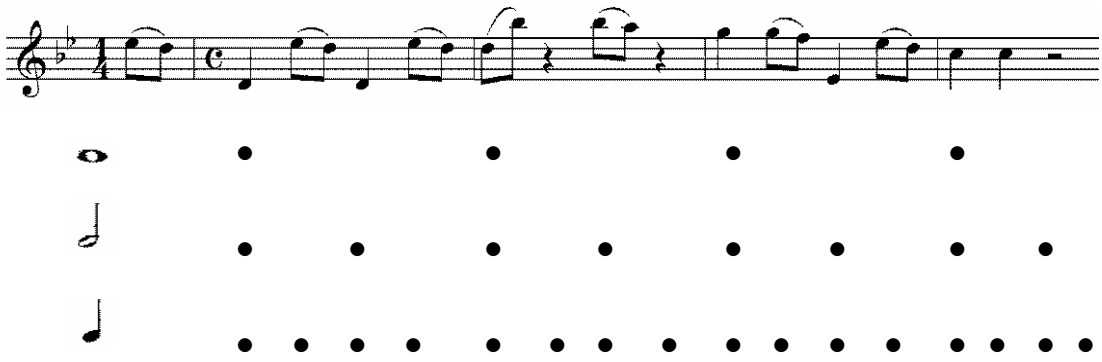
Lerdahl & Jackendoff (1983), Selkirk (1984), Chen (1984), Jackendoff (1989), Oehrle (1989), Prince (1989) and Hsiao (1991b) point out that the metrical beat and musical beat play the same role in language and musical rhythm respectively. Jackendoff (1989) claims that language and music are almost identical devices sharing the same evolutionary basis. He stresses that there are two main parallels between language and music. First, they both employ metrical grids as an indication of regular time value. Second, the musical structures for time-span reduction are similar to the prosodic trees of the theory of stress proposed by Liberman & Prince (1977). For example, for a phrase like "Mississippi mud," different metrical weights are placed on different syllables in speech form.

(3)



Similarly, in music, a musical phrase is composed of metrical grids, as illustrated in (4) (Jackendoff, 1989)

(4)



It is obvious that both language and music use place markers to show the levels of rhythm and regular stress pattern. In phonology, a place marker serves as an indication of a syllable, as represented by the Xs in the bottom row of (3). A syllable is more stressed if more Xs are placed. In music, a place marker functions as a beat of the lowest level, as shown by the dots in the bottom row of (4). In (4), 𝅝, 𝅗𝅥, and ♩ are viewed as stresses added to different beats. 𝅝 gives stress to the first beat of every measure. 𝅗𝅥 gives stress to the first and the third beats of every measure. ♩ gives stress to every beat of every measure. Therefore, the first beat of every measure

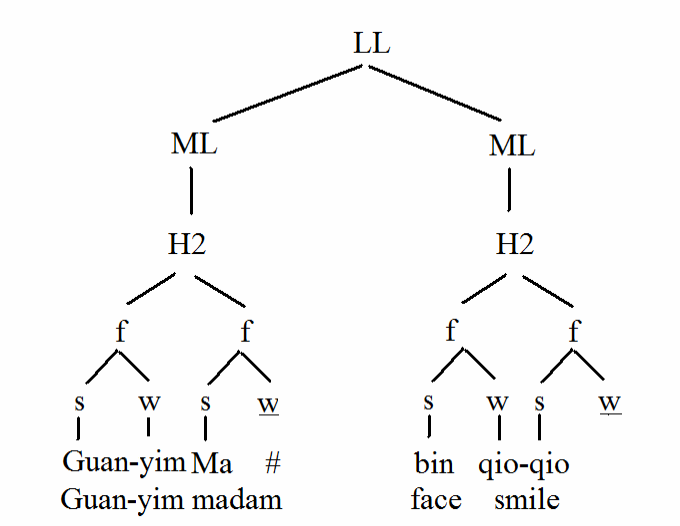
is more stressed than the other beats, as the number of dots in (4) suggest.

Hisao (1991) also states that when syllables are assigned metrical beats, a syllable-timing mechanism is presented by the quantitative value for duration. In addition to the common feature of beat assignment, Hsiao indicates that the notion of pause or lengthening is reflected both in the language and music forms, as illustrated in (5a) and (5b). The two forms are then compared to musical structures, in which the pause in language is realized by rest marks and the lengthening in language is realized by notes of longer duration, as illustrated in (6a) and (6b) respectively.

(5)

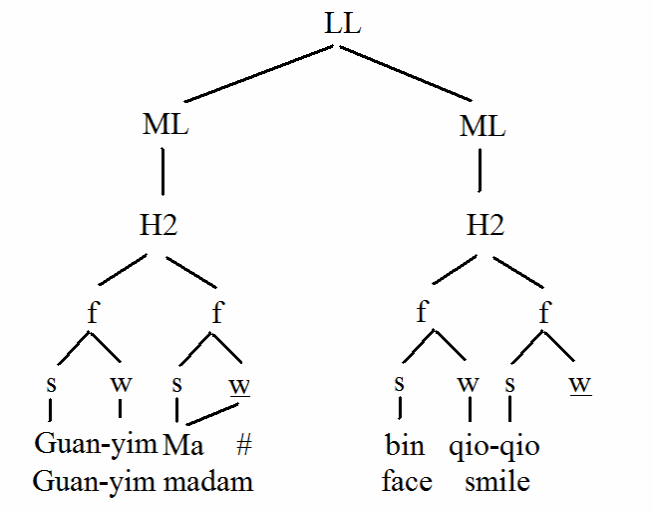
a. ‘Bodhisattva Guan-Yim smiles’ (Taiwan Minyao, 1987: 190)

Pausing



b. ‘Bodhisattva Guan-Yim smiles.’ (Taiwanese)

Syllable-Lengthening



(6)

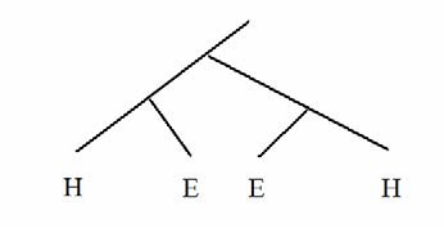
a. musical notation of (5a)

b. musical notation of (5b)

Besides, Yen (2004) proposes that the head-elaboration relationship is found both in music and in language. In music, musical heads are associated to relatively more important pitches, while less important pitches serve as ornamental part, known as elaboration. The deletion of elaboration causes the reduction of time span. However,

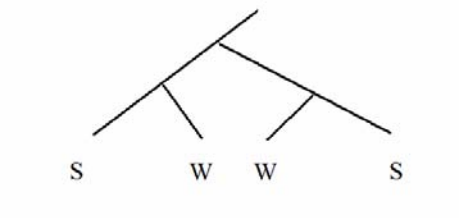
the time-span reduction does not influence the musical structure. Yen illustrates the time-span reduction tree as (7), in which H means head and E means elaboration.

(7)



A similar structure is found in the metrical tree. For example, in (8), a prosodic structure is composed of strong and weak syllables, indicated by S and W respectively. The Ss are the metrical heads.

(8)



Rhythm is not only a device to present pleasing effects but also an inner timing mechanism of structural events like language and music. The interaction of language and music therefore involves how the rhythmic structures of the two systems become accordant. The investigation on Jiang's verses also includes the relationship between language and music rhythmic patterns. It will be shown that the harmony of Jiang's

verses rely largely on the mapping of language and music rhythm.

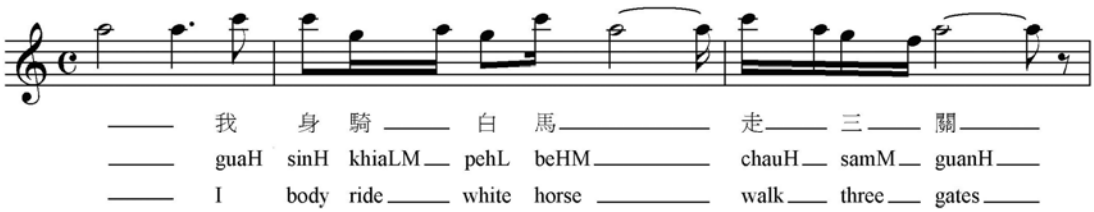
2.1.4 Language and Musical Tones

Pitch is another common trait of music and language. In terms of music, different music tones compose melodies and different intervals between notes create a variety of effects. Sound pitches also play an important role in language, especially in tonal language. Zhao (1930) uses the knowledge of music to identify the quality of language tones and marks each tone by relative music scales ranging from 1 to 5. For example, the rising tone in Chinese Mandarin is marked as 35, such as *ma35* (麻) 'sesame.' Some researchers have also investigated the correspondence of language and music tones in songs. To confirm the similarities of language and music tones, Deutsch, Henthorn, et al (2006) make an experiment on Chinese and English speakers' ability to pitches. The result shows that Chinese-speaking students have better performance in perfect pitch, which suggests the intimate relationship between language and music tones. Regardless of the factors for the performance of perfect pitch, the mapping of language and music tones has been the tradition of Chinese music. Yang (1996) points out that the traditional Chinese music is closely related to language tones, especially in *nanqu* 'southern songs.' He concludes that yin ping tone is often realized by notations with higher pitches and longer duration. Yang ping tone

is usually associated to notations with lower pitches than those associated to ying ping tone. In music, the corresponding melodic line of Ying ping tone is rising. Ying qu tone is often mapped with an arch melodic line with a peak in the middle, whereas *shang* tone usually corresponds to an arch melodic line with a foot in the middle. Agreeing with the findings of Yang, Tsao (1996) offers a statistic analysis on the language and music tones in *nanqu* and confirms that the relationship between language and music tones in songs is positive. Sun (1986) reveals that the correspondence of melodies and the four tones in Chinese language is also found in *jingju* 'Peking opera' and *hobei bangzi* '*bongzi* opera.' Wang (2002) posits several rhyming principles in Taiwanese operas. He states that the language sounds and melodies must be harmonious so that the singing can be understood and pleasing to the ear. According to Wang, the selection of rhymes with specific tonal categories is under the influence of music. The accordance of linguistic and melodious tones can be supported by the alternation of tonal categories in the line-final positions in Taiwanese opera, which is considered delicate arrangement instead of coincidence. Aside from rhymes, the relative difference of tone pitches is performed in non-rhyming position. As the example in (9), the syllables carrying high tone, such as *gua*H, *sin*H, *chau*H, and *guan*H, are associated to notes with relatively higher pitches than non-high tones. The syllable with rising tone, i.e. *khia*LM, is performed by a rising melodic line.

Similarly, a syllable with falling tone, such as *be*HM, is realized by a falling melodic line.

(9) ‘I passed the three gates by riding a white horse’



In this thesis, I will discuss the correspondence of language and music tones in Jiang’s verses in music. I will show that the parallels of tonal shapes of language and music exist in the works of the Song Dynasty.

2.2 Interactions of Language and Music

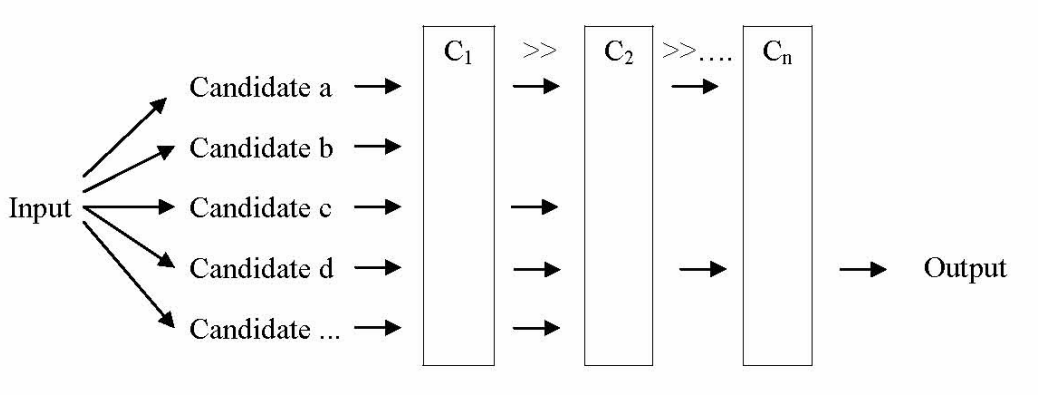
In addition to the similarities of language and music, the research on language and music has also focused on their interactions in songs. Some studies have been conducted to give description analysis on the mutual influence of language and music (Dobrian 1992, Besson and Schön. 2001). However, several researchers account for the interactions of language and music based on a set of constraints and consider the output forms the optimal choices in the competition. The approach is known as optimality-theoretical approach. Before introducing the application of the optimality theory in language-music interface studies, I will first summarize the theory and its

sub theories, which are the theoretical basis of the thesis.

2.2.1 Optimality Theory

Unlike process-oriented approaches, OT employs a set of ranked constraints to evaluate all the possible candidates produced by GENERATOR. The candidate whose violation of the constraints is the least serious is considered the optimal choice.

(10)



As the model by Kager (1998) shows, the input is generated to an infinite number of candidates evaluated by a set of constrained ranked in a hierarchy of strict domination. Therefore, violations of higher-ranked constraints are more serious than those of the violable constraints in lower ranking. In terms of the five candidates in (10), candidate (b) is ruled out because of the fatal violation of $C_{1.}$ Candidates (c) and (e) are eliminated as well because they do not conform to $C_{2.}$ Comparing candidate (a) and candidate (d), the former is inferior owing to the violation of $C_{n.}$ Although Candidate (d) may also violate some lower ranking constraints, it is still considered

the optimal choice among all candidates.

OT is based on the one-step evaluation rather than a process-oriented derivation. In the framework of OT, all generalizations or alternations are governed by constraints. The constraints are universal and are classified into two main types, namely faithfulness constraints and markedness constraints. Faithfulness constraints impose the requirements on the identity between input and output forms, while markedness constraints focus on the structural well-formedness of the output. The two groups of constraints serve as conflicting forces. On the one hand, faithfulness constraints have to prevent inputs from being realized as unmarked forms. On the other hand, markedness constraints lead to variations.

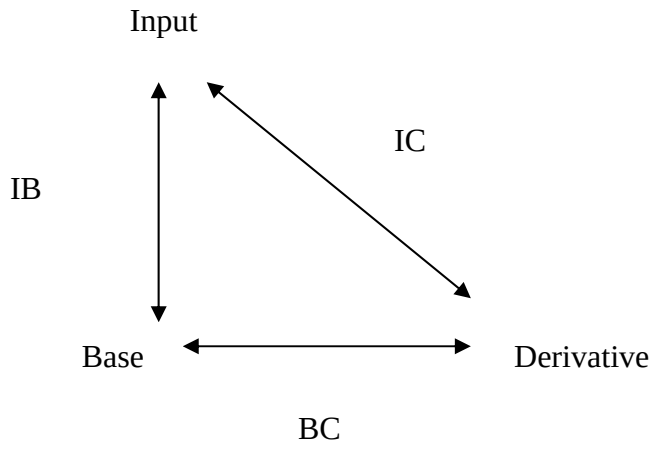
In addition to phonology, OT has been applied in many other fields, including music-related research. In this thesis, I will propose a series of constraints that determine the output forms of Jiang's verses in music.

2.2.2 Correspondence Theory

Correspondence theory is rooted in the 1970s as linguists probed into the similarity between the output and the reduplicated forms. As one of the main sub-theories of OT, Correspondence Theory was established by McCarthy and Prince (1995) and was elaborated in many studies (Ito, Kitagawa and Mester 1996,

Kenstowicz 1996, Kager 1999, Nelson 2003). As can be seen in (11), this theory explores an output-to-output correspondence, i.e BC, based on the input-to-output relationship, namely IB and IC.

(11)



The main concept of this theory is to use faithfulness constraints to investigate how outputs are affected by markedness constraints. It has achieved much success in explaining phenomena such as reduplication, vowel harmony, cyclicity, derived environment conditions, level ordering, opaque rule interactions, and others. In Correspondence Theory, the maintenance of the identity of input-to-output relationship is accomplished by six types of faithfulness constraints below:

(12) Faithfulness constraints in Correspondence Theory

A. Maximality: no deletion of segments

B. Dependence: no epenthesis of segments

C. Featural Identity: no featural changes

D. Contiguity: no medial deletion or epenthesis

E. Linearity: no metathesis

F. Anchoring: no epenthesis or deletion at edges

In this study, featureal identity is observed in Jiang's verses and will be applied to constrain the realization of rhymes in the language form.

2.2.3 Generalized Alignment

Enlightened by the edge-based in Selkirk (1986), McCarthy and Prince (1995) extends OT by accounting for the morpho-syntax and phonological issues under the general constraint format of Generalized Alignment (GA), in which constraints are subsumed by aligning two categories in either edge. The definition of GA is listed as follows:

(13)

Align(Cat1, Edge1, Cat2, Edge2) =def $\forall$ Cat1 $\exists$ Cat2 such that Edge1 of Cat1 and ge2 of Cat2 coincide.

Where Cat1, Cat2 $\in$ PCat $\cup$ GCat

Edge1, Edge2 $\in$ {Right, Left}

Generalized Alignment is typically used to confine two categories in prosodic hierarchy, such as syllable, foot, PrWd, or prosodic phrase. For example, the

alignment constraint ALIGN-WD-R requires every PrWd to end in a foot:

(14)

ALIGN-WORD-R

Align (PrWd, Right, Ft, Right)

“The right edge of every PrWd must coincide with the right edge of some foot.”

In the investigation on Jiang’s verses, alignment constraints will be applied to explain the mapping of language and music rhythm.

2.2.4 OT in Language-Music Interface Studies

OT is firstly applied in music-related research by Gilbers & Schreuder (2002), who show the conflicts between positional and segmental markedness constraints by examples in language and music respectively. In terms of language, the surface segments in child language are the winners in OT competition. As for music, the hierarchical structure is constructed by a set of constraints that select the optimal outputs. For instance, the selection of chords is based on the following rules:

(15)

TSRPR 1: Choose as the head of a time-span the chord (or the note) which is in a relatively strong metrical position (positional markedness).

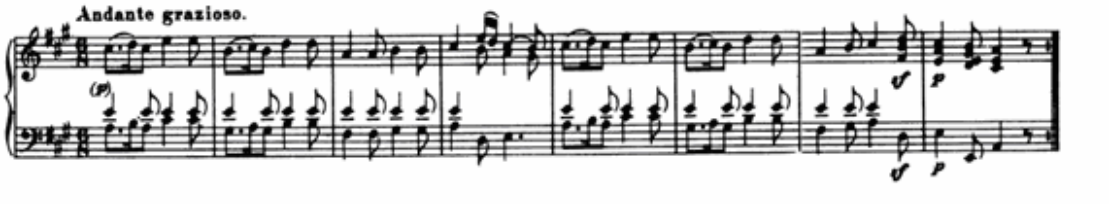
TSRPR 2: Choose as the head of a time-span the chord (or the note) which is

relatively harmonically consonant (segmental markedness).

TSRPR 7: Choose as the head of the time-span the chord (or the note) which emphasizes the end of a group as a cadence (comparable to the boundary marking effect of alignment constraints in language).

To prove the operation of OT in music, Gilbers & Schreuder adopt Mozart K. 331, I and evaluate two possible cords. With the constraints in (15), the examination shows that the winner, i.e. E chord, is the optimal choice and most pleasing to the ears.

(16) Mozart K. 331, I



(17)

constraints → A^6 – E Candidates ↓	TSRPR 7	TSRPR 2	TSRPR 1
☞ E			*
A^6	*!	*	

Gilbers & Schreuder (2002) also indicate that language and music must be parsed into smaller parts exhaustedly and the boundary phenomena can be accounted for by alignment constraints. Besides, the adjacency of identical elements is avoided both in language and in music, which is due to the Obligatory Contour Principle (OCP)

(McCarthy, 1986). Aside from Gilbers & Schreuder's study on language and music, Hendriks and Werf (2004) conduct a 10-subject experiment to investigate the similarities and differences between language and music. They present a set of constraints that best fit the musical grouping of their data by computational model. Furthermore, OT has been applied in the study of Chinese folk songs. Yen (2004) discusses the parallels of language and music. She shows that both language and music are universal, though they may exist with different forms. Besides, arguing in favor of Gilbers & Schreuder, Yen uses Western classical music and Chinese folk songs to prove that OT can also be applied in music. She proposes three relevant constraints as follows:

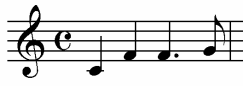
(18) MIN-BINARITY: a group contains no fewer than two pitch-events.

(19) PITCH(X)PITCH(X): adjacent pitch-events that have the same pitch are grouped together

(20)ALIGN (Group, Left/Right, Slur, Left/Right): Align the left/right edge of a musical slur with the left/right edge of a group

Yen uses the Taiwanese folk song "Four Season Red (四季紅)" to demonstrate the OT evaluation of musical arrangement. As can be seen in tableau (21), the actual output of the first measure of the song is the optimal choice among the relevant possibilities.

(21) the first measure of “Four Season Red”

 Input: 1 2 3 4	MIN-BINARITY	PITCH(X)PITCH(X)	ALIGN (Group, Left/Right, Slur, Left/Right)
a. [1][2][3][4]	*!***	*	
b. [1][23][4]	*!*		
c. [1][234]	*!		*
d. [12][34]		*!	*
☞ e. [1234]			
f. [1][2][34]	*!*	*	*
g. [12][3][4]	*!*	*	
h. [123][4]	*!		*

The previous research bridge language and music with their features in common. It is also confirmed that language and music theories are mutually compatible in certain aspects.

2.3 Metrical Phonology

The measurement of metrical tension was once focused on the linguistic stresses with the S and W positions. However, Kiparsky points out that metrical units, such as feet, play an important role in metricality. Meyer (1956) also notices a hierarchical

structure of verse and suggests that the structure of verse and song are alike. Jakobson (1970) destructs a heptasyllabic line into different constituents based on footing, as illustrated in (22).

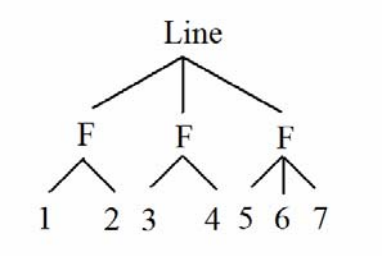
(22)

a. Heptasyllabic line: 12/34/567

b. Pentasyllabic line: 12/345

Furthermore, Chen (1979) proposes the metrical structure of Chinese regulated verses in terms of tonal arrangement, arguing that hierarchy and binarity characterize Chinese poetic forms. The principle of the metrical division is that adjacent syllables are similar in terms of tones, while distant syllables are different. For example, in terms of the heptasyllabic line in (22a), contiguous syllables 1 and 2, 3, and 4 bear the same tone. By contrast, syllables 5 and 7 are distant and thus their tonal distinctions are opposite. Based on Jakobson's tripodal division, as illustrated in (23), Chen suggests that metrical trees are hierarchical and composed of binary branches.

(23)

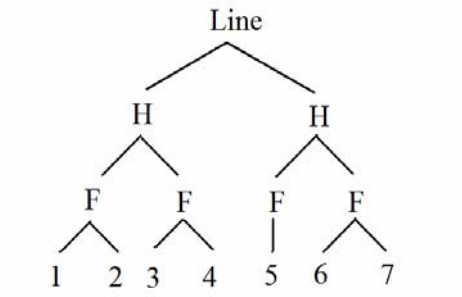


The structure of verse is supported by the pause in verse recitation. For instance,

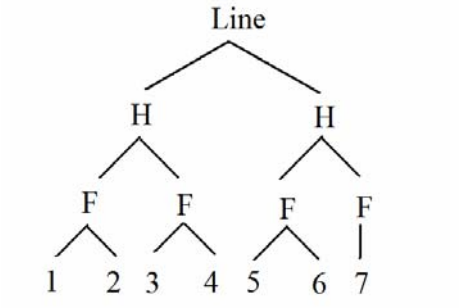
a caesura can be recognized between the second and third feet. Thus, the pause between the fourth and the fifth syllables is longer than that after the second, the fifth or the sixth positions. The implication of the optional minor pause after the fifth or sixth syllable is that the last three syllables of a heptasyllabic line is either right or left branching. That is to say, as illustrated in (24), a metrical line consists of two half lines or hemistichs, which are further divided into two metrical feet. The metrical line can be classified into two types, depending on whether the last trisyllabic hemistich is right- or left- branching.

(24)

a. Right-branching

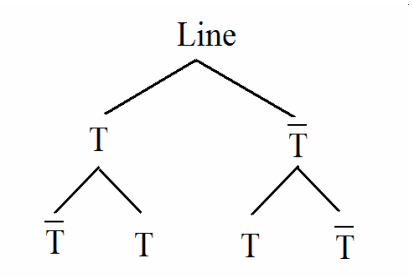


b. Left-branching



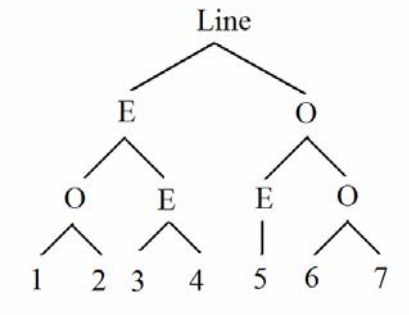
Such proposal has been supported by the tonal requirements on regulated verses. Suppose T is either E or O, the barred T should carry the opposite tonal category. The tonal tree is schemed as follows.

(25)



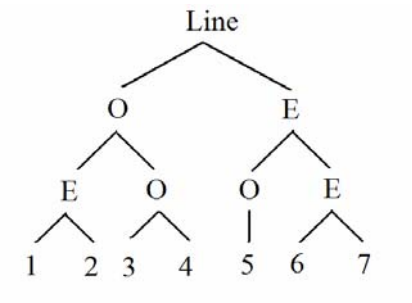
Thus, a heptasyllabic line, beginning with an E hemistich would be destructed as follows.

(26)



On the contrary, a heptasyllabic line starting with an Oblique tone should be in the form of (27)

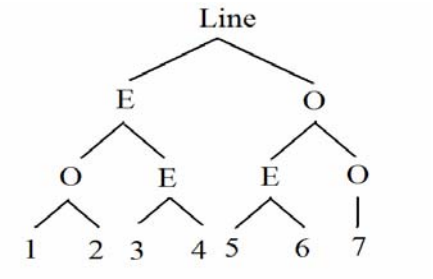
(27)



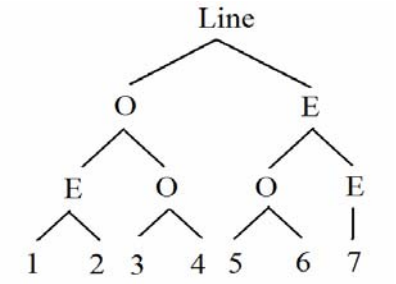
The two trees can be transferred to left-branching structures, as illustrated in (28a) and (28b) respectively.

(28)

a.



b.

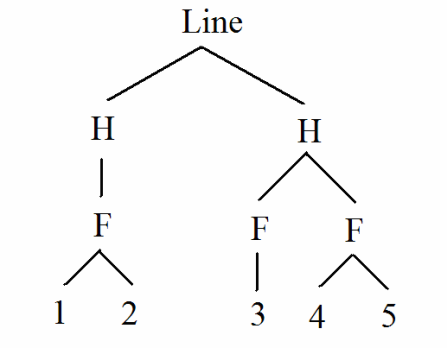


As for pentasyllabic lines, it is unnecessary to provide extra explanation because they are identical to the heptasyllabic counterparts except that the first two syllables

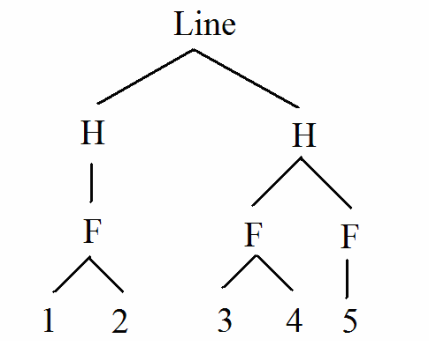
are empty. The metrical tree of a five-syllable line is thus in the form of (29a) or (29b), depending on the branching direction.

(29)

a. Right-branching

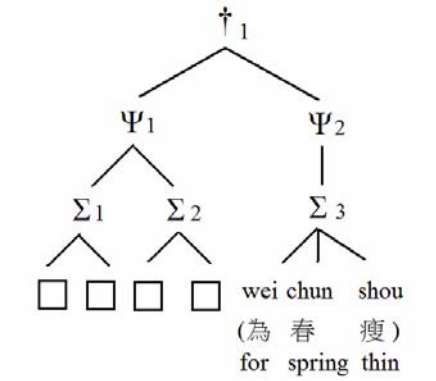


b. Left-branching



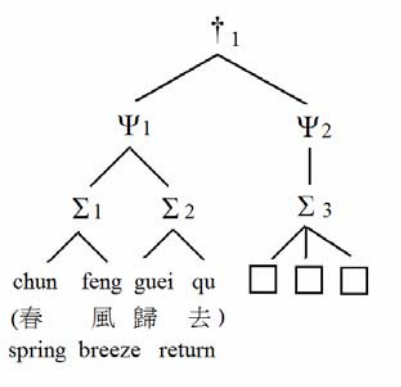
As Song verses are irregulated lines, the scheme of metrical structure should be broadened to a larger sphere aside from five- and seven- syllable lines. Hsiao (1995) proposes the notion of Partial Template to account for metrical lines composed of fewer than seven syllables. First, a trisyllabic line can be viewed as the last hemistich of a heptasyllabic line, with the first hemistich omitted.

(30) ‘becoming thin because of spring’



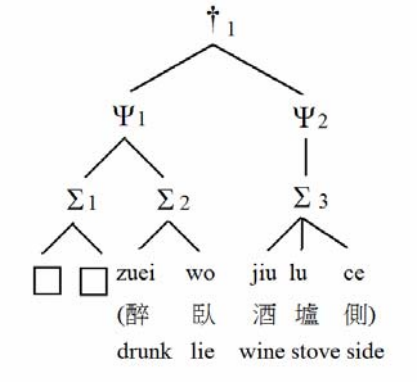
Similarly, a quadrasyllabic line is taken as the first hemistich of the heptasyllabic line.

(31) ‘the spring breeze retreats’



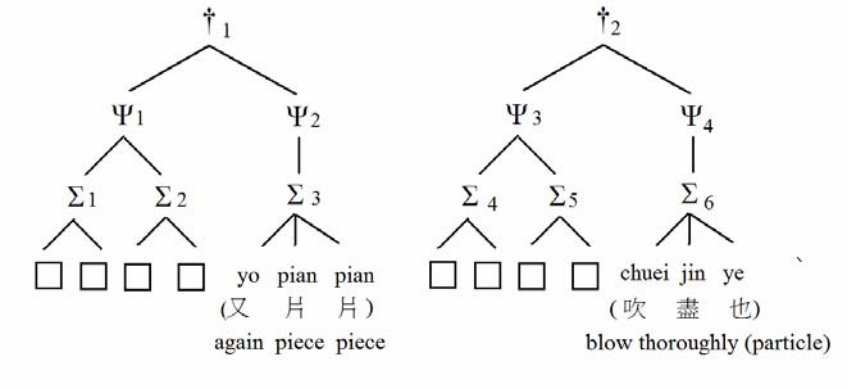
A metrical line is composed of five syllables also shares the same metrical structure except that the firs two syllables are omitted.

(32) ‘lie drunk beside the wine stove.’



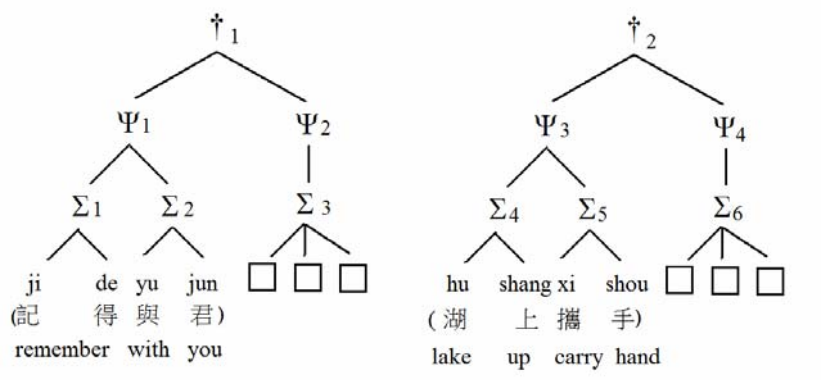
As for lines containing six syllables, the metrical template is considered a combination of two ternary feet in two separate metrical trees.

(33) ‘the leaves are blown thoroughly again.’



The integration on hemistichs in different metrical trees also justifies the metrical patterns of eight-syllable lines, which can be simply referred to as the combination of the first hemistich of the separate trees in (33).

(34) ‘I remember that we were hand in hand when we were on the lake.’



Hsiao’s explanation on metrical structure provides a solid basis for the study of irregulated verses. In Song verses, metrical lines from one to six syllables are formed by the partial template, while those composed of more than seven syllables arise from the combination of two partial templates.