Chapter 1

Introduction

This thesis uses Optimality Theory to explore the formal parallels between language, specifically prosody, and music. There are three main goals of this study. The first goal is to show that music and prosody have similar structures, providing evidence for parallels between the cognition of language and music. The second goal is to show that just as many formal universals have been found across languages, there are also music universals across different music genres. In this thesis I specifically compare Western classical music and Chinese folk songs. The third goal is to argue in favor of Gilbers and Schreuder (2002) and Van der Werf and Hendriks (2004), who have shown that Optimality Theory can be insightfully applied to music.

1.1 Theoretical Background

Though music is regarded as a type of art, it can be studied in a scientific way. Some researchers have used statistical methods to analyze music (Nicholls, 1993; Landy, 1991), focusing on the relationship between the frequency of the occurrence of various musical phenomena and the musical representation. Others have used linguistic methodology to study music (Jackendoff and Lerdahl, 1982; Jackendoff, 1989; Lerdahl and Jackendoff, 1983; Gilbers and Schreuder, 2002; Sundberg and Lindblom, 1991; Van der Werf and Hendriks, 2004), discussing the use of generative grammars and Optimality Theory in music. Our interest here is in the latter approach.

Lerdahl and Jackendoff (1983) and Sundberg and Lindblom (1991) have developed generative theories for musical grammars. Their works were primarily inspired by Chomsky's generative theory of language, in which grammars can be formulated and the formulations can be used to generate sentences. The general goal of their works is to make use of finite sets of rules to generate infinite sets of musical structures. Distinctively, Lerdahl and Jackendoff's (1983) theory is based on Western tonal music and Sundberg and Lindblom's (1991) theory is based on Swedish nursery tunes. Later, Optimality Theory was also brought to music in Gilbers and Schreuder (2002) and Van der Werf and Hendriks (2004). These two studies are grounded on Lerdahl and Jackendoff's (1983) Generative Theory of Tonal Music (henceforth GTTM), but reworked under the model of Optimality Theory. Gilbers and Schreuder (2002) show that a system like OT can describe the processes of musical cognition well. Van der Werf and Hendriks (2004), who gave an OT analysis of musical grouping and further ran a judgment experiment to support their analysis, is the specific inspiration for the present study.

This study also builds the music grammar of Chinese folk songs on Lerdahl and Jackendoff's (1983) GTTM. A particularly interesting issue is their claim that there are parallels between music and language and that the two human cognitive abilities, linguistic and musical capacities, overlap in some way. That's what we are most concerned about in this study. If Chinese folk songs do follow the GTTM, then it will be further evidence both for music universals and for the resemblance between language and music. On the other hand, since Gilbers and Schreuder (2002) and Van der Werf and Hendriks (2004) demonstrate similar processes between musical and linguistic cognition in their OT analyses, we also use OT in an attempt to find more specific resemblances between language and music, as well as music universals. However, since we conduct this study from a linguistic perspective, we don't introduce much of Gilbers and Schreuder's (2002) and Van der Werf and Hendriks' (2004) OT analyses, which focus on the special properties of music. Thus, unlike Gilbers and Schreuder's (2002) and Van der Werf and Hendriks' (2004), the OT constraints we propose for music focus on those that seem most similar to constraints proposed for linguistic prosody. Nevertheless, the judgment experiment methodology of Van der Werf and Hendriks (2004) is adopted in our study, owing to the goal of our OT analysis to predict musical judgments.

1.2 Music Universals and the Innateness of Language and Music

Much research has shown that a study of language or music can be seen as an investigation of human cognitive capacities (Jackendoff and Lerdahl, 1982; Lerdahl and Jackendoff, 1983; Jackendoff, 1989; Pinker, 1997; Raffman, 1993; Sundberg and Lindblom, 1991; Temperly, 2001; Thomas, 1995). In order to understand linguistic and musical cognition, grammars which comprise formal systems of principles or rules are constructed. A linguistic grammar describes our linguistic competence, that is, our knowledge of a language, while a musical grammar describes our knowledge of a musical genre. For a long time, there has been a debate in linguistics about whether there is an innate basis to linguistic grammar. The same question comes up with musical grammars. Jackendoff (1989) and Pinker's (1997) answers to this question support the view that a musical grammar, like a linguistic grammar, builds on innate knowledge. Jackendoff's (1989) claims that musical expertise is developed from the musical capacities we all share. Pinker (1997) believes music derives from a variety of independently necessary innate systems. Lerdahl and Jackendoff (1983) adopt two arguments from linguistics to answer the question mentioned above: the "poverty of stimulus" argument and the argument from "universals." The first argument claims that someone is able to write a song without learning a music grammar in advance. The knowledge he needs has been acquired through the music he experienced. Namely, the music grammar is constructed by some kind of inherent cognitive capacity in our mind. By musical universals, Lerdahl and Jackendoff (1983) mean the universals of musical grammars, that is, those universal rules proposed in

GTTM. However, Lerdahl and Jackendoff (1983) have noted that a universal rule in music need not be manifested in all musical genres. In other words, the rules themselves are universal but not all of them are equally used in every music genre. If it is not so, then we would not expect different genres of music found in the world.

The examination of Chinese folk songs in this paper is an attempt to test Lerdahl and Jackendoff's (1983) claim about musical universals. If their theory works for Chinese folk songs, that is to say, the universal rules they proposed are found in other music genres too, then our work can further support the argument of universals and the claim that the music capacity is innate. Since OT is a model in which constraints are universal and yet may be violated or unexpressed in some systems, this study hopes to be more precise about the nature of music universals through our OT analysis.

1.3 Parallels Between Language and Music

In most respects, language is quite different from music. However, we can still find parallels between them. Jackendoff (1989: 29) suggested that language and music are similar to each other in the same way as our fingers and toes: "They are morphologically nearly identical devices with a common evolutionary basis, but specialized to different purposes." Pinker (1997: 529) claimed that "music may borrow some of the mental software for language." Sundberg and Lindblom (1991) pointed out that both musical and linguistic structures can be decomposed into parts or constituents and constructed in a hierarchical fashion.

To us the most interesting claim about parallels between language and music is Jackendoff's (1989: 27). He claims that there are two parallels between music and phonology: 1) they both use metrical grids to "mark off temporal regularities" and 2) "the trees for time-span reduction in music are a notational variant of the prosodic trees of the Liberman-Prince (1977) theory of stress". Those are unexpected findings that emerged from the development of Lerdahl and Jackendoff's (1983) GTTM. Lerdahl and Jackendoff (1983) believe that the overlap of the two human cognitive abilities, linguistic and musical capacities, should not be a coincidence. There is a more general cognitive organization that has manifestations in both musical and linguistic structures. Borchgrevink (1982: 151) has made a similar claim that prosodic and musical rhythm are processed by the speech hemisphere and concluded that "There is a close, but complex, connection between speech function and musical function. Both apparently consist of a multitude of sub-functions, many of which rely upon the same psychological mechanisms."

Regarding the first parallel addressed by Jackendoff (1989), he finds that in both music and language, metrical weights strongly rely on length. More about this will be discussed in Chapter two. Below we give illustrations of metrical grids. Example (1) is adopted from Hayes (1995: 38) and example (2) is adopted from Jackendoff (1989: 21)[1].

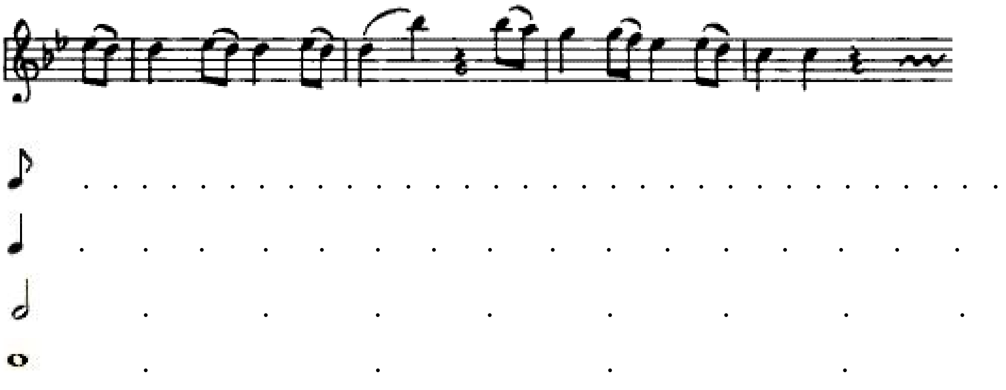
(1)

```
                x
  x             x
  x     x       x
  x  x  x  x    x
Mississippi mud
```

---

[1] The metrical grids are associated with the opening of Mozart's G Minor G Minor, K. 550 Symphony.

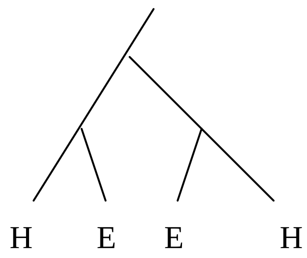
(2)



The Xs in the bottom row of (1) and the dots in the top row of (2) serve as place markers. In phonology, a place marker indicates a syllable; in music, a place marker indicates a beat of the lowest level in metrical structure. The other Xs in (1) are used to mark stressed syllables. Syllables that have more Xs than others are more stressed. Analogously, in (2), the other dots are used to mark stressed beats of a given level, and the beat with the most dots is most stressed. Another point that should be mentioned is that the metrical grids in both (1) and (2) show regular stress patterns, that is, the Xs and dots are equally spaced.

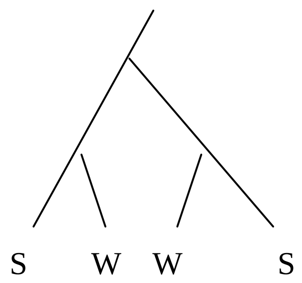
As for the time-span reduction of the second proposed parallel between music and phonology, it concerns the head-elaboration relationship among pitches in a musical piece and describes implicit knowledge about surface structures. A time-span is the interval between beats. Notice that a beat is a point in time, so it has no duration. It is the time-span which has duration. Musical heads refer to those relatively more important pitches, while elaborations refer to less important pitches. Elaborations are heard as ornamental pitches to the heads; hence they can be deleted in some way and without affecting the overall structure of a piece. The deletion of the elaborations reduces the time-spans, so this process is termed as *time-span reduction*. The reduction takes places from the smallest level of the time-spans to larger levels. It is represented with tree diagrams in GTTM. Below is a simplified version of a time-span reduction tree, where H stands for head and E stands for elaboration.

(3)



This is similar to the prosodic trees we have in phonology.

(4)



S and W refer to strong syllable and weak syllables respectively. We may take a musical head as being like a strong syllable in a prosodic tree and an elaboration as being like a weak syllable.

Along with Lerdahl and Jackendoff's (1983) findings of the two parallels between music and phonology, a theoretical issue in phonology arises. Phonologists such as Prince (1983) have argued that phonological representation does not need both metrical grids and tree structures. They claimed that the metrical grid alone can capture all of the generalizations about stress rules. Since the issue above is controversial, we will not focus on it in our analysis.

1.4 Data

Chinese folk songs are the main data in our study. Here Chinese folk songs include Taiwanese folk songs, traditional folk songs from Mainland China and so on. Unlike classical pieces, most of the composers of folk songs are anonymous. Even if a folk song is famous, usually only its original area is known, but not its composer. Besides, since folk songs were orally transmitted for centuries or decades, our readers might easily find different versions of our data.

Like folk songs from other cultures, many of the Chinese folk songs reflect people's everyday lives. In a sense, the musical grammars of folk songs should be closer to the nature of our unconscious knowledge of music because while not everyone can play Mozart's masterpieces, everyone sings some folk songs. For this reason, GTTM should be able to account for folk songs well. In order to study the similarities and differences among musical grammars, we compare Chinese folk songs with other music genres, such as Western classical music and waltz music.

Our analysis of music focuses on Grouping structure and Metrical structure, which will be introduced in the next chapter, because they are more related to our goal of studying resemblances between language and music.

1.5 Comparison of Different Music Genres

Since one of our goals is to find music universals across different music genres, we compare different music genres in our OT analysis and in the judgment experiment presented in chapter four. Chinese folk songs are compared with Western classical music and waltz music in our OT analysis, while Chinese classical musicians are compared with Western classical musicians in our judgment experiment.

It would be interesting to see a comparison of Chinese classical music and Western classical music because in the literature these two music genres are claimed to be very different from each other. 楊 (1986) mentions that unlike Western classical music, metrical shifts (i.e. switch of meters) appear a lot in Chinese classical music. 沈 (1982) and 楊 (1986) even argue that regular stresses do not occur in Chinese classical music. They claim that stresses can be imposed on any note anywhere in a piece by performers. Moreover, 楊 (1986) also claims that the segmentation of musical flow (i.e. grouping of notes) in Chinese classical music is determined by 句 (a sentence) and 逗 (a pause), and it is more flexible than in Western classical music. Furthermore, in terms of pitches, the scale has seven major pitches (Do, Re, Mi, Fa, So, La, Si) in Western classical music, while in Chinese classical music, it has only five major pitches (Do, Re, Mi, So, La).

The overall organization of the rest of this thesis is as follows. Chapter two demonstrates Lerdahl and Jackendoff's (1983) GTTM, on which our study is based. Chinese folk songs are the data we use to illustrate the theory. Weaknesses of the GTTM are outlined in the conclusion of that chapter. In chapter three, we give an OT analysis of Chinese folk songs. We show that constraints for music can be formally quite similar to constraints for language, therefore revealing more specific resemblances between language and music. In order to study the differences and similarities of constraints and ranking, we also compare Chinese folk songs with other music genres. Then, we point out how our OT analysis gives a better analysis than GTTM does. The fourth chapter describes a judgment experiment of musical grouping. By means of this experiment, we explain how our knowledge of music is used to make musical judgments and how the musical constraints should be ranked. Finally, chapter five concludes this thesis.

Chapter 2

A Generative Theory of Tonal Music

Since the present study builds on Lerdahl and Jackendoff's (1983) Generative Theory of Tonal Music (i.e. GTTM), this chapter provides an introduction to the theory. The first part of this chapter gives a brief introduction to GTTM, primarily focusing on its framework. The second part demonstrates how GTTM works on Chinese folk songs by showing some simple examples. The last section concludes this chapter by noting the weakness of GTTM, which can be avoided by a constraint-based approach, i.e. Optimality Theory, as shown in the following chapter.

2.1 Brief Introduction to GTTM

Lerdahl and Jackendoff (1983: 1) take their goal for the GTTM to be "a formal description of the musical intuitions of a listener who is experienced in a musical genre". By musical intuitions, they refer to a listener's unconscious knowledge of music that enables a listener to perceive the organization of a musical piece. Though a composer or a performer may have more musical talent than we do, they are listeners too. Our musical grammars are affected by the music genres we have experienced. In other words, Lerdahl and Jackendoff (1983) are claiming that all of us basically share the same cognitive capacities for music, but musicians get their talent from richer experience in music.

There are five levels of mental representation for music in the GTTM: 1) Musical Surface, 2) Grouping Structure, 3) Metrical Structure, 4) Time-span Reduction and 5) Prolongational Reduction (Jackendoff, 1989). The first level has a flat sequential organization, while the other four levels are all hierarchical. In addition, the four hierarchical levels are all derived from the Musical Surface. Detailed illustrations of the five levels will be given in Section 2.2, accompanied with examples of Chinese folk songs.

For the hierarchical levels of representation, the GTTM comprises a system of rules that assigns analyses to musical pieces. Note that it is not about composers. The musical pieces are fixed; GTTM derives interpretations of them. The two main sets of rules in the GTTM are *Well-Formedness rules*, which “establish the formal structures and their relationship to the string of pitch-events that form a piece” and *Preference rules*, which “establish which of the formally possible structures that can be assigned to a piece correspond to the listeners’ actual intuitions” (Lerdahl and Jackendoff 1983: 36-37).[2] The GTTM claims that Preference Rules are genre-specific. In addition, all the rules in the GTTM can be applied to structures in a single piece cyclically. However, the universality of those rules might be challenged. For instance, Chinese classical music may be argued to go against GTTM’s rules for metrical structure because it is regarded as a music genre in which regular stresses do not occur (楊, 1986; 沈, 1982).[3] Yet, this is taken into account in GTTM. Lerdahl and Jackendoff (1983) claim the stylistic norms of some musical genres don’t give the rules opportunities to apply. They give an analogous example from the visual arts. Some artistic genres may only make use of black and white, but we would not claim that viewers couldn’t make use of principles of color perception. A similar situation can be found in linguistic phenomena too. We would not claim that the ability to handle consonant clusters isn’t universal because we couldn’t find any clusters in some language.

As we have mentioned earlier, GTTM defines music universals to be universals of musical grammars—“the principles available to all experienced listeners for

---

[2] There is a minority of rules which deal with special phenomena, such as grouping overlap and grouping elision.
[3] Though regular stresses are found in certain pieces, specifically drum beats, they are rare.

organizing the musical surfaces they hear, no matter what genre they are experienced in" (Lerdahl and Jackendoff 1983: 278). In this sense, simple musical forms are constructed on the same knowledge base as complex music forms. Therefore, GTTM should be able to account for folk songs or nursery tunes as well as Beethoven's and Mozart's symphonies. In the next section, we will see how GTTM can account for simpler musical forms, specifically Chinese folk songs.

2.2 Examples of Chinese Folk songs

In this section, we'll go through GTTM's five levels of mental representation for music, and apply its rules to Chinese folk songs to see how they work. Since Grouping Structure and Metrical Structure are directly related to our study, we will particularly emphasize them.

2.2.1 Musical Surface

Musical surface is what a listener physically hears in a piece. Pitches are produced successively and are heard in a sequential order. In order to capture the musical surface, musical notations were invented. In GTTM, musical surface has less musical significance. What GTTM is concerned about are the other four hierarchical levels which are derived from the musical surface.

Below is a display of the musical surface, which has been notated into a score. This is a famous Taiwanese folk song named "四季紅." Note that only the first few measures[4] of the song are presented (see Appendix 1 for the complete score)[5].

---

[4] A measure is a musical metric unit between two bars on a staff.

[5] The score is taken from the website of 宜蘭縣鄉土教材. http://media.ilc.edu.tw/music/MS/ms_music-pic/ms_music-pic13.htm

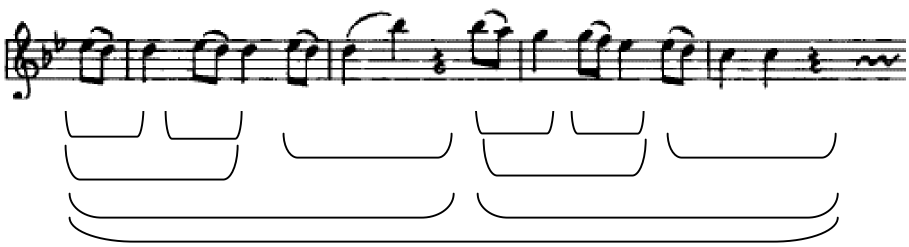
(5)

2.2.2 Grouping Structure

2.2.2.1 Introduction to Grouping Structure

Grouping here refers to the segmentation of the musical surface, decomposing it into groups (*motives*, *sections*, *phrases* and so on).[6] A smaller group can be embraced by a larger group, so the grouping structure is in a hierarchy. In GTTM, grouping structures are notated by the use of slurs under the notes of a musical score. A grouping example on the opening of Mozart's Symphony in G Minor, K. 550 is shown below (taken from Lerdahl and Jackendoff 1983: 37, figure 3.1).

(6)



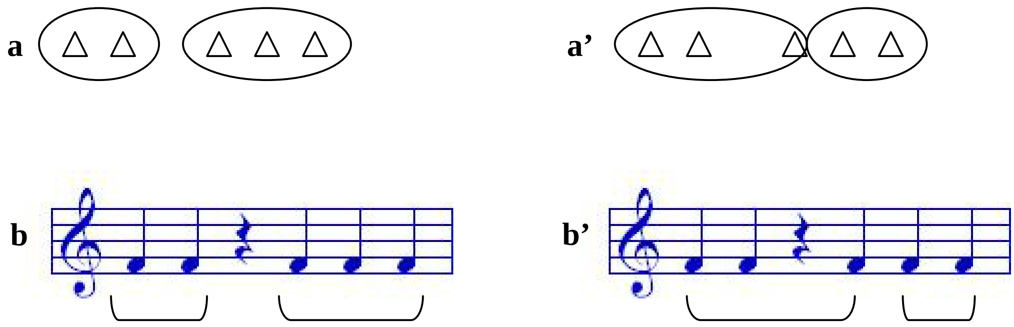
The grouping structure is not unique to musical cognition. Lerdahl and

---

6 According to Randel and Apel (1986), a *motive* is the smallest musical composition which has at least two pitches (notes) and one of them should be a strong beat. It is usually one measure long. A *section* is larger than a motive and is usually two measures long. The composition which is larger than a section is a *phrase*. Usually, it is four measures long.

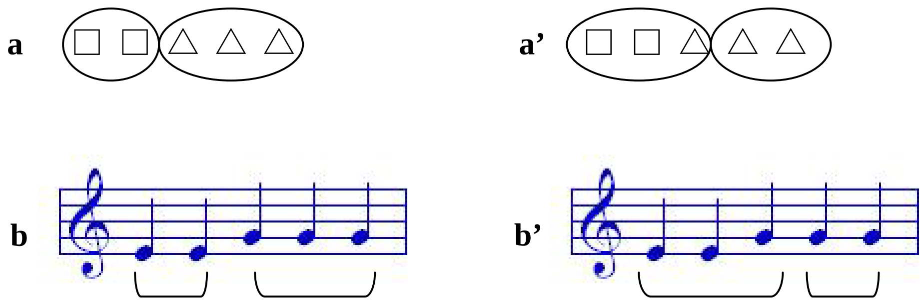
Jackendoff (1983) mentioned that the grouping structure is a common property found in many areas of human cognition. It's quite natural for us to segment a series of elements into chunks or groups in our daily life. For example, we segment a friend's telephone numbers into 3- or 4-digital chunks. In a like manner, when people listen to music, a series of pitches will be segmented into groups unconsciously. Moreover, the process of grouping is not arbitrary. There are principles behind it. Most of the factors encoded by the principles are the relative distance (*proximity*) and *similarity* among the elements. These two main principles tell us a lot about optimal groupings. Let's look at the following examples, based on discussion in Lerdahl and Jackendoff (1983).

(7)



In (7)a, the space between the second triangle and the third triangle implies a boundary between the triangles. Triangles that are relatively closer tend to be grouped together, therefore the two triangles on the left will form one group and the three triangles on the right will form another group. (7)a' shows a grouping that violates our intuitive grouping process. Analogously, the rest in (7)b functions like the space in (7)a, so intuitively we will make the two left-hand-side notes one group and the three right-hand-side notes the other group. (7)b' is an illegal grouping in musical grouping because it violates the proximity principle. Next, we will see examples of grouping that relates to the other principle—the similarity principle.

(8)



In (8)a, the five elements are equally spaced and the distinctive difference among them is their shapes. Spontaneously, we tend to group the two circles together and make the three triangles the other group. Our process of grouping here is affected by the similarity of the elements. Because of the similarity principle, grouping like (8)a’ will not normally occur. Likewise, in musical grouping, (8)b shows grouping according to the similarity principle, while (8)b’ shows an illegal grouping in music.

The overall restrictions on the grouping structure are stated into rules in the GTTM. They are termed Grouping Well-Formedness Rules (GWFRs) and Grouping Preference Rules (GPRs). There are five GWFRs and seven GPRs in total. In this present study, we will only introduce the rules that are relevant to our analysis of Chinese folk songs.

2.2.2.2 Analysis of the Grouping Structure in Chinese Folk songs

Below is the application of Grouping Well-Formedness Rules (GWFRs) and Grouping Preference Rules (GPRs) in the first four measures of “四季紅”, the Taiwanese folk song mentioned earlier. We will first introduce the GWFRs that are relevant to the grouping, then go to the GPRs, which most of the logic of our groupings relies on. Notice that the arcs which mark the groupings are numbered for our convenience in the discussion; the groupings are intuitions of a listener, which readers who know music may reconfirm the groups with their own intuitions. These groupings are intended to reflect psychologically real structure. The purpose of the rules in GTTM is to describe this structure.

(9)

All the slurs in (9) obey the first *Grouping Well-Formedness Rule* (GWFRs) because the pitch-events[7] we grouped together are all successive.

GWFR 1 Any contiguous sequence of pitch-events, drum beats, or the like can constitute a group, and only contiguous sequences can constitute a group. (Lerdahl and Jackendoff 1983: 37)

---

[7] The term *pitch-events* are the pitches played in a piece. They are represented as notes in musical notation so that a pitch is also recognized as a note. When a pitch is played alone, a pitch-event refers to the single pitch. Yet, when some pitches are played simultaneously, i.e. harmonies, a pitch-event may refer to several pitches at the same time.

As we can see, we make some bigger groups embrace smaller groups in (9). For instance, slur 5 embraces slur 1 and slur 2; slur 7 embraces slur 5 and slur 6. This embracing relationship obeys GWFR 3 (Lerdahl and Jackendoff 1983: 38).

GWFR 3 A group may contain smaller groups.

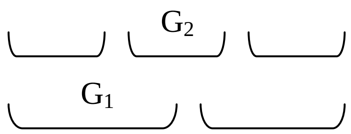
Yet, there are restrictions on the embracing relationship. These are encoded in GWFR 4 and GWFR 5 (Lerdahl and Jackendoff 1983: 38).

GWFR 4 If a group G_1 contains part of a group G_2, it must contain all of G_2.

GWFR 5 If a group G_1 contains a smaller group G_2, then G_1 must be exhaustively partitioned into smaller groups.

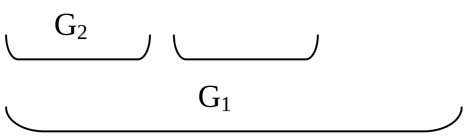
Hence, an ungrammatical violation of GWFR 4 will be like (10).

(10)



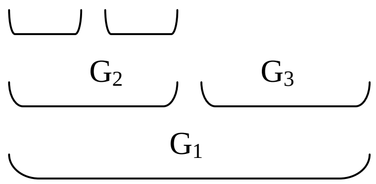
A grouping that GWFR 5 prohibits is like (11).

(11)



Note that a grouping such as (12) doesn't violate GWFR 5.

(12)



It is allowed to have a larger group like G_3 that is not subdivided into smaller groups.

Next we will consider the *Grouping Preference Rules* (GPRs).

GPR 1 Strongly avoid groups containing a single event[8] (Lerdahl and Jackendoff 1983: 43).

According to Lerdahl and Jackendoff (1983), small groups are not preferred in musical grouping because they are not easy to be perceived in musical flow. Therefore, we don't mark those pitch-events with arcs[9] beneath them into smaller groups. Another reason for the grouping is based on GPR 3 (taken from Lerdahl and Jackendoff 1983: 46).

GPR 3 (Change) Consider a sequence of four notes $n_1n_2n_3n_4$.
All else being equal, the transition n_2-n_3 may be heard as a group boundary if

a. (Register) the transition n_2-n_3 involves a greater intervallic distance than both n_1–n_2 and n_3–n_4, or if

b. (Dynamics) the transition n_2-n_3 involves a change in dynamics and n_1–n_2 and n_3–n_4 do not, or if

c. (Articulation) the transition n_2-n_3 involves a change in articulation and n_1–n_2 and n_3–n_4 do not, or if

d. (Length) n_2 and n_3 are of different lengths and both pairs n_1, n_2 and n_3, n_4 do not differ in length.

---

8 There is a more general alternative form of GPR 1: "Avoid analyses with very small groups—the smaller, the less preferable" (Lerdahl and Jackendoff 1983: 43).

9 These arcs are also termed 'slurs' in musical notation, but they do not equal the slurs we use in grouping. It is a mark to remind performers to perform the notes smoothly. Though sometimes these two kinds of slurs overlap, they are basically not the same.

Let us first give explanations for the musical terms mentioned in the rule. A *register* is a division of the range of an instrument or singing voice. It is usually defined by a change in the pitch of the sound between a lower range and a higher range. *Dynamics* refers to loudness or softness. Some most often seen markings of dynamics are ***p*** (soft), ***f*** (loud), ***pp*** (very soft), ***mf*** (somewhat loud) and so on. *Articulations* give information about the characteristics of notes to performers. They are represented by symbols above the notes in a musical score. For example, > indicates that the duration of a given note should be performed shorter than it is normally, · indicates that a note should be performed detached from its neighboring notes, and – or ⌒ (slur) indicates a note or a group of notes should be performed in a smooth, graceful and connected style. *Length* simply refers to the duration of a note (e.g. quarter-note ♩, half-note 𝅗𝅥, etc.).

To get back to the point earlier, we take group one in figure (9) as an example. According to GPR 3, the transition of the second pitch-event and the third pitch-event does not involve a change in register, so there is no boundary between them. In a like manner, we don't divide the pitch-events in measure two into smaller groups because they don't involve a change in register either.

On the other hand, a group boundary will be heard between the first measure and the second measure, caused by GPR 2 (Lerdahl and Jackendoff 1983: 45).

GPR 2 (Proximity) Consider a sequence of four notes $n_1n_2n_3n_4$. All else being equal, the transition n_2-n_3 may be heard as a group boundary if

a. (Slur/Rest) the interval of time from the end of n_2 to the beginning of n_3 is greater than that from the end of n_1 to the beginning of n_2 and that from the end of n_3 to the beginning of n_4, or if

b. (Attack-point) the interval of time between the attack point of n_2 and n_3 is greater than that between the attack points of n_1 and n_2 and that between the attack points of n_3 and n_4.

As has been noted earlier, the term "slur" in GPR 2 refers to the arcs below or above pitch-events in a musical score. These musical slurs don't correspond to our grouping slurs[10]. An attack-point is a point in time when a pitch-event begins to be played. Based on GPR 2a, since there is an unslurred transition between slurred transitions, a boundary must be placed between the first measure and the second measure.

Notice that the grouping structures in the first measure is the same as the one in the third measure. This is not purely a coincidence, but also a reflection of their parallel constructions, which is described in GPR 6 (Lerdahl and Jackendoff 1983: 51).

> **GPR 6** (Parallelism) Where two or more segments of the music can be construed as parallel, they preferably form parallel parts of groups.

As a final point, these four measures can further be treated as a bigger group in a larger level, as shown by group 7. The principle allows this is stated in GPR 5 (Lerdahl and Jackendoff 1983: 49).

> **GPR 5** (Symmetry) Prefer grouping analyses that most closely approach the ideal subdivision of groups into two parts of equal length.

The parallelism of groupings in the first four measures causes symmetry, and therefore those smaller groups form a larger group, i.e. group 7.

[10] As explained in footnote 9, a musical slur is a mark to remind performers to perform the notes smoothly. Though sometimes these two kinds of slurs overlap, they are basically not the same.

2.2.3 Metrical Structure

2.2.3.1 Introduction to Metrical Structure

Metrical structure is a hierarchical organization concerned with the patterns of strong and weak beats in a piece. Recall what we have introduced in chapter one: in both language and music, metrical weights rely on length. A heavy beat is regarded as a strong beat. That is to say, a strong beat is relatively more stressed, and the time-span of a pitch-event on a strong beat has relatively longer duration. Lerdahl and Jackendoff (1983) clarify musical stress as follows.[11] They point out there are three kinds of stress in music: phenomenal, structural and metrical. *Phenomenal stress* is free and can occur anywhere at the musical surface. A performer may use this kind of stress to emphasize any pitch-events in a piece he or she wants. *Structural stress,* as implied by its name, is caused by the musical structure. For example, the *cadence* of a phrase is likely to be stressed[12]. This kind of stress doesn't occur regularly in a piece. As for the *metrical stress*, it is a regular mental construct that listeners impose on music. This is the kind of stress that GTTM is concerned with.

In addition, the GTTM claims that metrical structure is a relatively local phenomenon. Lerdahl and Jackendoff (1983: 21) stated, "Even though the dots in a metrical analysis could theoretically be built up to the level of a whole piece, such an exercise becomes perceptually irrelevant except for short pieces".

In order to capture the metrical structure, a metrical grid is used to mark off the metrical regularity. Since the organization of metrical structure is hierarchical, there are two or more levels of beats in the metrical grid of a piece. In the metrical grid, a dot represents a beat at a given level and each dot should be equally spaced at any given level. Let's take two measures of a 3/4-meter grid for example (adopted from

[11] Stress is also called "accent" by other authors (e.g. Cooper, 1973).
[12] *Cadence* is a musical term which means "ending."

Lerdahl and Jackendoff, 1983: 20, figure 2.8a).

(13)

	(1		2		3		1		2		3	)
♪	.	.	.	.	.	.	.	.	.	.	.	.
♩	.		.		.		.		.		.	
𝅗𝅥.	.						.					

There are three levels in this particular grid. The smallest level is the eighth-note (♪) level. It is the note with the shortest time-span in a piece that occurs in this level. The intermediate level is the quarter-note (♩) level and the largest level in this grid is the dotted-half-note (𝅗𝅥.). The dots in the largest level indicate a most stressed beat. Therefore, this metrical grid indicates that the first beat of a measure in 3/4-meter is the strongest beat. Lerdahl and Jackendoff (1983: 20) also pointed out that in Western music, “the time-spans between beats at any given level must be two or three times longer than the time-spans between beats at the next smaller level.”

Like the grouping structure, the overall restrictions on the metrical structure are stated into rules in the GTTM. There are four Metrical Well-Formedness Rules (MWFRs) and ten Metrical Preference rules (MPRs) in total. Rules that are relevant to our analysis will be described in the following discussion.

2.2.3.2 Analysis of the Metrical Structure in Chinese Folk songs

Again, we will use the Taiwanese folk song “四季紅” and analyze the metrical structure of its first four measures. Each metrical level will be given the time-span note value to its left for our convenience in the discussion.

(14)

There are three levels in our analysis. The note which has the shortest time-span in this song is the eighth-note, so we make it the smallest metrical level. Since this is a 4/4-meter piece, the time-spans of the beats at each level are two times longer than the time-spans of the beats at the next smaller level. Following MWFR 1, every note in our analysis is associated with a beat at the smallest level.

> **MWFR 1** Every attack point must be associated with a beat at the smallest metrical level present at that point in the piece (Lerdahl and Jackendoff 1983: 69).

Given that an *attack point* refers to a point in time when a pitch-event begins to be played, here we can just treat it as a note or a pitch-event. Furthermore, in the metrical grid, a beat at a larger level is also a beat at a smaller level. This is a restriction stated by MWFR 2 (Lerdahl and Jackendoff 1983: 69).

> **MWFR 2** Every beat at a given level must also be a beat at all smaller levels present at that point in the piece.

As we have said earlier, a dot at a given level, except for the smallest level, indicates a strong beat at that level. So, the dots at the quarter-note level and half-note level indicate stronger beats in the metrical grid. Note that the strong beats at each level are spaced two beats apart. This is consistent with the statement of MWFR 3 and MPR 10 (Lerdahl and Jackendoff 1983: 69, 76).

> **MWFR 3** At each metrical level, strong beats are spaced either two or three beats apart.
>
> **MPR 10** (Binary Regularity) Prefer metrical structures in which at each level every other beat is strong.

As in Grouping Structure, the two parallel constructions in the first four measures will receive the same metrical analysis. MPR 1 formalizes this phenomenon.

> **MPR 1** (Parallelism) Where two or more groups or parts of groups can be constructed as parallel, they preferably receive parallel metrical structure. (Lerdahl and Jackendoff 1983: 75).

To link up our metrical analysis with grouping structure, we find that at the quarter-note level and half-note level, the strongest beat in a group is always the first beat in that group, though this is not strongly preferred in GTTM.

MPR 2 (Strong Beat Early) Weakly prefer a metrical structure in which the strongest beat in a group appears relatively early in the group. (Lerdahl and Jackendoff 1983: 76)

A strong beat in music tends to be stressed. It is shown clearly in the metrical grid that the strongest beats have the most dots, i.e. most stressed. MPR 4 describes this preference.

MPR 4 (Stress) Prefer a metrical structure in which beats of level L_i that are stressed are strong beats of L_i. (Lerdahl and Jackendoff 1983: 79).

As in language, strong beats are dependent on weights. Here are the ways to decide a strong beat listed in MPR 5 (Lerdahl and Jackendoff 1983: 84).

MPR 5 (Length) Prefer a metrical structure in which a relatively strong beat occurs at the inception of either
a. a relatively long pitch-event
b. a relatively long duration of a dynamic,
c. a relatively long slur,
d. a relatively long pattern of articulation,
e. a relatively long duration of a pitch in the relevant levels of the time-span reduction, or
f. a relatively long duration of a harmony in the relevant levels of the time-span reduction (harmonic rhythm).

The principles that MPR4 and MPR5 state are similar to those for phonology in language. Heavy syllables, also known as strong syllables, carry stress in many languages. This has been contributed to the *Weight-to-Stress Principle* in Prince (1983) and Prince and Smolensky (1993): if a syllable is heavy, then it is stressed. The weight of a syllable depends on the number of *moras* it has. A light syllable has only one mora and a heavy syllable has two moras. In terms of length, the vowels in heavy syllables are relatively longer than those in light syllables. However, in phonology, length is supposedly quantized (heavy vs. light), but in music, it’s gradient.

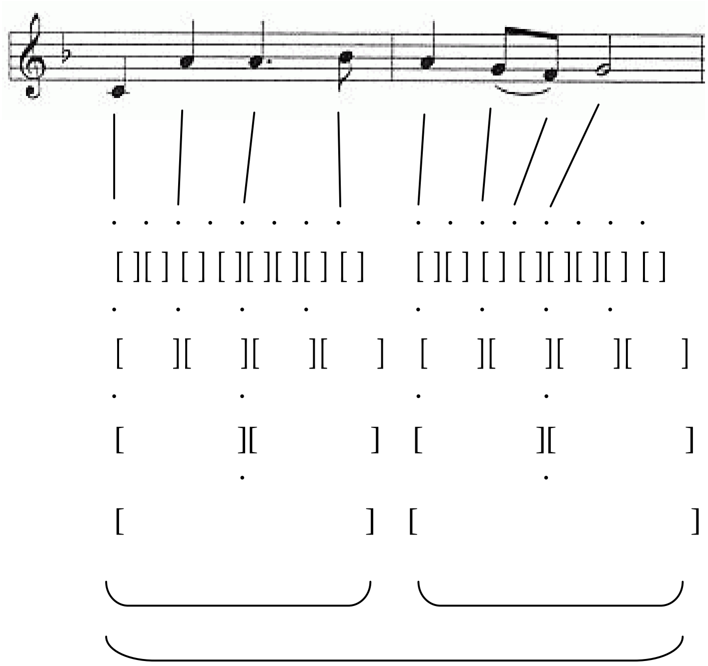
2.2.4 Time-span Reduction

2.2.4.1 Introduction to Time-span Reduction

Time-span reduction deals with head-elaboration relationship and harmonies (pitches performed simultaneously). It is a hierarchical organization and can be represented in a tree structure. The basic notion of time-span reduction is, as we have introduced in Chapter one, that some pitches are heard as relatively more important than other pitches in a piece, and the relatively less important pitches can be reduced without destroying the overall structure of the piece. The deleting of pitches can also be seen as reducing the number of the time-spans, so this process is termed time-span reduction.

In earlier sections, we have mentioned that time-spans are the intervals between beats. Hence, a time-span can be as short as the interval between two beats next to each other at the smallest metrical level or as long as the interval between the first and the last beat of a large group. The segmentation of time-spans is determined by the interaction of grouping and metrical structure and it is the input of time-span reduction. Below is an illustration of the time-span segmentation of the third and fourth measures in “四季紅”. Notice that the lines are used to associate each pitch-event with its place marker at the lowest metrical level and time-spans are indicated with brackets.

(15)



The relationship between grouping structure and metrical structure in (15) is “out of phase” because the strongest beats in the fourth tier of metrical structure don’t line up with the inception of groups. On the contrary, if all stronger beats line up with the inception of groups, then their relationship is “in phase”. When grouping structure and metrical structure are in phase, stronger beats hold structural importance. In GTTM, the time-spans of pitch-events that are heard as less important or structurally less important can be reduced.

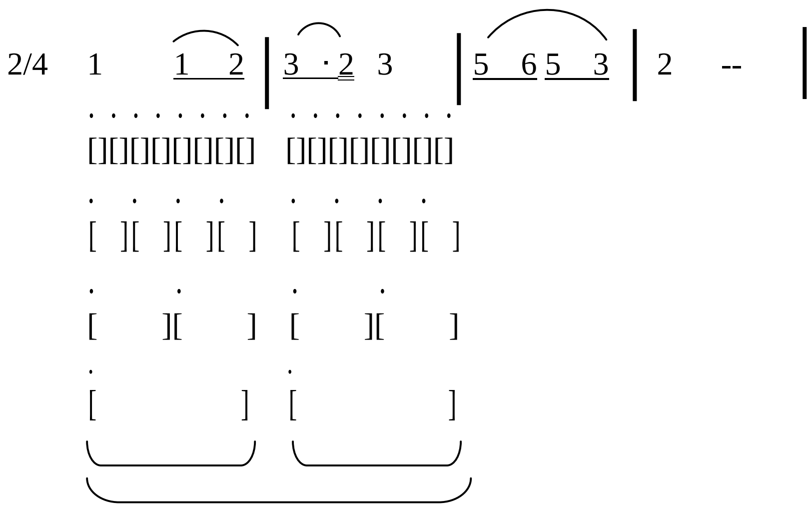
Examples of the head-elaboration relationship can be easily found in a musical genre like jazz (Jackendoff, 1989). In jazz, the organization of a piece is usually a theme and its variations. Pitches in the theme usually match those relatively more important pitches in the variations. Another type of head-elaboration relationship is harmony, which is a typical characteristic of Western music. The idea is that in harmony, some pitches are heard as more stable than others. The less stable pitches can be reduced. Since most Chinese folk songs don’t use harmonies, we illustrate time-span reduction by dealing with the relationship between a theme and its variations. In GTTM, the principles behind the time-span reduction are stated in two Segmentation Rules, four Time-Span Reduction Well-Formedness Rules (TSRWFRs), and nine Time-Span Reduction Preference Rules (TSRPRs).

Like many other folk songs in the world, Chinese folk songs were traditionally orally transmitted, leading to different versions of the same folk song. 何 (1986) has given an example of the famous folk song “孟姜女”. A comparison of the different versions can also be seen as a study of a musical theme and its variations. It reveals that listeners of Chinese folk songs also have the knowledge of time-span reduction. Therefore, two of the versions for “孟姜女”, which are adopted from 何 (1986: 154-155), are selected to compare in the next section.

2.2.4.2 Analysis of the Time-span Reduction in Chinese Folk songs

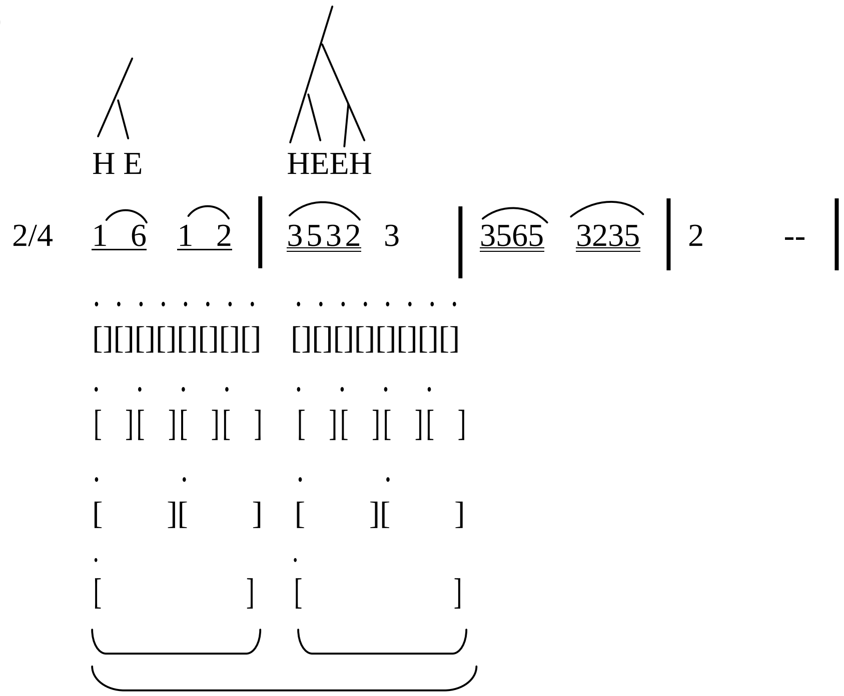
These two songs are transcribed in numbered musical notation.[13] Since there are more pitch-events in (17), we consider (16) as a theme and (17) as its variation. Our analysis will only focus on the first two measures of each version.

(16)



[13] An illustration of the numbered musical notations is given in Appendix 2.

(17)



Let’s compare the pitch-events of the first beat in the two first measures. We take the first pitch-event in (17) as a head and the second pitch-event as its elaboration according to their stresses. Since the stress pattern of this song is trochaic, the first beat of each measure is strong and the first pitch-event is most stressed. As stated in TSRPR 1, the first pitch-event of the first beat at (17) is preferred as a head pitch.

TSRPR 1 (Metrical Position) Of the possible choices for head of a time-span T, prefer a choice that is in a relatively strong metrical position (Lerdahl and Jackendoff 1983: 160).

Since the second pitch-event of the beat is heard as an ornamental pitch to the first pitch-event, it can be reduced. As predicted, the reduced form of the first measure of (17) will be like the first measure of (16). The time-span reduction can be represented by the tree diagram which is shown above the score in (17), where H stands for head and E stands for elaboration.

Now we turn to the time-span reduction in the second measure of (17). There are four pitch-events at the first beat. We choose the first and the last pitch-events to be head pitches. The reason for the first one is the same as what we have explained in the first measure. The reason for the last pitch-event to be a head is described as follows in TSRWFR 4 (Lerdahl and Jackendoff 1983: 159).

> **TSRWRF 4** If a two-element cadence is directly subordinate to the head *e* of a time-span *T*, the final is directly subordinate to *e* and the penult is directly subordinate to the final.

Since the last two pitch-events form a two-element cadence among the four pitch-events of the first beat, the final pitch-event will be chosen as a head. Comparing the second measure of (17) with the second measure of (16), we may find the one of (16) can be seen as a reduced form of this measure. The tree diagram above this measure shows its time-span reduction.

2.2.5 Prolongational Reduction

According to Jackendoff (1989: 26), prolongational reduction refers to "the sense of musical flow across phrases, the building and releasing of musical tension." In particular, prolongational reduction deals with the function of harmonies in music, a typical feature of Western music. The basic notion of it is similar to that of the time-span reduction, that is to say, the structure of a piece may be reduced to a simpler form. Since Chinese folk songs don't use harmonies, they don't employ it. Yet, we do not have to assume that this level of representation doesn't exist, but can instead suppose that it is universally available in the human mind, but Chinese folk songs just don't provide opportunities for prolongational reduction to apply.

2.3 Conclusion

This chapter provides a basic introduction to Lerdahl and Jackendoff's (1983) GTTM, which is the theory our study is based on. We have used some Chinese folk songs to test GTTM, and we found that this theory works for Chinese folk songs. This result leads us to agree with the music universals they propose in GTTM. Consequently, the implication of music universals might make us believe that some of our music capacities are innate. For this issue, we definitely need deeper study and more investigation into other music genres. In the next chapter, we will present a comparison of Chinese folk songs and other music genres to discover more about music universals. As for the parallels between music and language claimed by GTTM, they seem to be apparent in our analysis of Chinese folk songs. In order to better understand the overlap of language and music cognition, the focus on parallels between music and language will be carried on in the next chapter.

2.4 Strengths and Weaknesses of GTTM

While reviewing the GTTM, we have noticed some strengths and weaknesses of it. As we have seen above, GTTM formalizes valid insights into music cognition by proposing rules. These rules describe our cognitive capacities of music and make the study of music scientific, rather than just artistic. What's more, GTTM is concerned with the universality among musical grammars. This follows the idea of Universal Grammar (UG), the central goal of linguistic theories. Such an approach is theoretically more practical than a genre-specific grammar and even helps in arguing for the innateness of music capacity.

However, some of the rules that GTTM propose are vague. It is not clear whether a structure is *required* or *allowed* in a rule. Take GPR 6 as an example: "where two or more segments of the music can be constructed as parallel, they *preferably* form parallel parts of groups" (Lerdahl and Jackendoff, 1983: 51). This description doesn't state that parallel structures are required. Instead, it says that parallel structures are preferred. This causes the problem of identifying when the rule may be violated. The vagueness of rules would lead us not to apply the theory rigidly. Moreover, the rules in GTTM obviously should have some kind of ranking, but the theory doesn't spell it out. For instance, GPR 1 prescribes that single-pitch-event groups are strongly avoided (Lerdahl and Jackendoff, 1983: 43). According to the literal meaning of this rule, it should outrank most of the rules in GTTM and yet can be outranked by other rules of GTTM too. Yet, this ranking of rules is not made explicit in GTTM. Also, differences across genres are only described by saying that some rules don't apply, but this makes some rules seem to be genre-specific and makes the universality of the rules doubtful.

The following chapter provides an OT analysis of Chinese folk songs. Since OT has violable and ranked constraints, the above weaknesses of GTTM can be avoided.

Chapter 3

An Optimality Theory Analysis of Chinese Folk songs

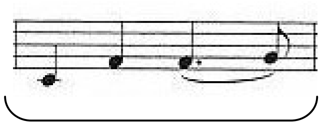
This chapter presents an Optimality Theory analysis of the Grouping Structures and Metrical Structures in Chinese folk songs to demonstrate that an approach like OT gives better formal description of music than GTTM does. Since we are more interested in exploring the resemblances between language and music, we only provide the analysis of Grouping Structure and Metrical Structure, which are highly related to linguistics. We turn those rules in GTTM into constraints and clarify how they are ranked. In addition, we compare Chinese folk songs with other musical genres to uncover the similarities and differences in their constraints and ranking. In contrast with Gilbers and Schreuder (2002) and Van der Werf and Hendriks (2004), the constraints we propose for music closely match constraints for prosody. For this reason, our analysis reveals more specific resemblances between language and music.

Our analysis starts with the grouping structure of Chinese folk songs. The second section analyzes the metrical structure of Chinese folk songs. The third section describes a comparison of Chinese folk songs and other musical genres on their constraints and ranking. The fourth section summarizes resemblances between language and music. Finally, the fifth section explains how OT gives a better description of music structure than GTTM does.

3.1 Grouping Structure in Chinese Folk songs

In order to make the difference between OT and GTTM clear, we take “四季紅” as the main example for our OT analysis of Grouping Structure. The first measure of “四季紅” is given again here.

(18)



In this measure, all four pitch-events form one group at the smallest level. This is due to an interaction of five constraints, which will be discussed one by one as follows.

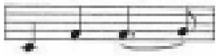
The MIN-BINARITY constraint has the greatest effect on the grouping of measure one. The principle behind this constraint is GPR 1, which strongly disfavors a group containing a single pitch-event (See p. 17).

(19) MIN-BINARITY

A group contains no fewer than two pitch-events.

Since GTTM claims that single-pitch-event groups are strongly avoided, we predict the MIN-BINARITY constraint to be ranked high. This constraint is illustrated in the tableau below. Notice that we digitize the pitch-events, 1 for the first pitch-event, 2 for the second pitch-event, and so on. Brackets are used to mark groupings.

Tableau 1: Demonstrating MIN-BINARITY in the first measure of “四季紅”

 Input: 1 2 3 4	MIN-BINARITY
a. [1] [2 3 4]	*
b. [1] [2] [3 4]	**
c. [1] [2] [3] [4]	****
d. [1 2 3] [4]	*
e. [1] [2 3] [4]	**
f. [1 2] [3 4]	
g. [1 2 3 4]	
h. [1 2] [3] [4]	**

As we can see in the tableau, the MIN-BINARITY constraint blocks all the grouping combinations with single-pitch-event groups and leaves two (f and g). One divides the four pitch-events into two groups, and the other makes the four pitch-events into one group. The constraints relating to these two groupings are given below.

(20) PITCH(X)PITCH(X)

Adjacent pitch-events that have a same pitch are grouped together.

(21) ALIGN (Group, Left/Right, Slur, Left/Right)

Align the left/right edge of a musical slur with the left/right edge of a group.

The PITCH(X)PITCH(X) constraint derives from GPR 3a, which states that a transition of two pitch-events may be heard as a group boundary if it is involved in a greater change of interval than its neighbors (see p. 18). There is one more constraint which also derives from GPR 3a – the INTERVAL(X)INTERVAL(X) constraint: “Neighboring pitch-events that have same intervals in between will put a boundary following the last pitch-event” (a list of constraints we propose for music is given in Appendix 3).

PITCH(X)PITCH(X) says the second pitch-event and the third pitch-event should not be separated into different groups. The constraint ALIGN (Group, Left/Right, Slur, Left/Right) comes from GPR 2a, in which a musical slur is said to symbolize a group, while a rest indicates a group boundary. The relationship between rests and boundaries is described in the REST/BOUNDARY constraint: “A rest must be a group boundary”.

The interaction among MIN-BINARITY, PITCH(X)PITCH(X) and ALIGN (Group, Left/Right, Slur, Left/Right), which are related to the first measure, is described by the following tableau.

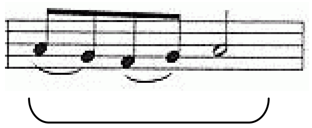
Tableau 2: Analysis of the first measure in “四季紅”

 Input: 1 2 3 4	MIN-BINARITY	PITCH(X) PITCH(X)	ALIGN (Group,L/R, Slur, L/R)
a. [1][2][3][4]	****!	*	
b. [1][2 3][4]	**!		*
c. [1][2 3 4]	*!		*
d. [1 2][3 4]		*!	
e.☞ [1 2 3 4]			*
f. [1][2] [3 4]	**!	*	
g. [1 2] [3][4]	**!	*	
h. [1 2 3] [4]	*!		*

Since PITCH(X)PITCH(X) itself has an implication of having more than one pitch-event, it’s hard for us to figure out the ranking of MIN-BINARITY and PITCH(X)PITCH(X). However, we can be sure that both of them must outrank the alignment constraint. Let us focus on their interaction with ALIGN (Group, Left/Right, Slur, Left/Right). As we can see in tableau 2, ALIGN (Group, Left/Right, Slur, Left/Right) would favor incorrect candidates like (a), (d), (f) and (g) and block the optimal candidate (e). To prevent candidate (a), (f) and (g) from winning, we should have both MIN-BINARITY and PITCH(X)PITCH(X) outrank ALIGN (Group, Left/Right, Slur, Left/Right). On the other hand, ALIGN (Group, Left/Right, Slur, Left/Right) would put boundaries around the third and fourth pitch-events, thus blocking the optimal candidate (e). To protect candidate (e), PITCH(X)PITCH(X) must outrank ALIGN (Group, Left/Right, Slur, Left/Right).

The second measure of “四季紅” shows an even clearer demonstration of MIN-BINARITY’s outranking ALIGN (Group, Left/Right, Slur, Left/Right). We represent the grouping structure of measure two again in (21).

(22)



Like measure one, measure two forms one whole group. It is ALIGN (Group, Left/Right, Slur, Left/Right) which puts the boundary between the two measures. However, ALIGN (Group, Left/Right, Slur, Left/Right) would require the two musical slurs in measure two to put boundaries around the first and fourth pitch-events, hence isolating the fifth pitch-event. To make the five pitch-events be in a group, we must have the MIN-BINARITY constraint outrank the ALIGN (Group, Left/Right, Slur, Left/Right) constraint. Tableau 3 illustrates the constraint interaction in measure two.

Tableau 3: Analysis of the second measure in “四季紅”

 Input: 1 2 3 4 5	MIN-BINARITY	ALIGN (Group, Left/Right, Slur, Left/Right)
a. [1 2][3 4][5]	*!	
b. [1 2 3 4] [5]	*!	*
c.☞ [1 2 3 4 5]		*!

The resulting ranking so far is {MIN-BINARITY, PITCH(X)PITCH(X)} >> ALIGN (Group, Left/Right, Slur, Left/Right).

In the preceding discussion, the MIN-BINARITY constraint is said to have a great effect on groupings. Yet, we can still find examples that violate it in Chinese folk songs. Below is a scrap of the Taiwanese folk song named “一隻鳥仔孝救救” with its grouping structure indicated in the first measure (see Appendix 1 for the complete score).[14]

(23)

In the musical flow, rests will result in breaks, as in the first measure above. For this reason, the grouping structure of measure one will be perceived as two small groups.

[14] This song is also known as “一隻鳥仔孝啾啾”. The score is taken from the website of 宜蘭縣鄉土教材. http://media.ilc.edu.tw/MUSIC/MS/ms_music-pic/ms_music-pic04.htm

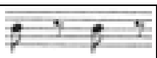
This property is stated in the REST/BOUNDARY constraint, which has been noted above.

(24) REST/BOUNDARY

A rest must be a group boundary.

Since REST/BOUNDARY makes the two pitch-events in measure one isolated, we know it should be ranked above the MIN-BINARITY constraint. What's more, these two isolated pitch-events have the same pitch, thus also violating PITCH(X)PITCH(X). Consequently, REST/BOUNDARY must outrank these two competing constraints. A simple tableau expresses this ranking:

Tableau 4: Analysis of the first measure in “一隻鳥仔孝救救”

 Input: 1 2	REST/BOUNDARY	MIN-BINARITY	PITCH(X)PITCH(X)
a. ☞ [1] [2]		*	*
b. [1 2]	*!		

In this section, we analyze Grouping Structure in Chinese folk songs and reveal its constraint ranking. To sum up, the ranking schema for the Grouping Structure Chinese folk songs is (25).

(25) Ranking schema for Grouping Structure:

REST/BOUNDARY >> MIN-BINARITY, PITCH(X)PITCH(X) >>

ALIGN (Group, Left/Right, Slur, Left/Right)

Next, we will go into Metrical Structure in Chinese folk songs.

3.2 Metrical Structure in Chinese Folk songs

Before we start to analyze the metrical structure of Chinese folk songs, we would like to remind our readers that metrical structure is said to be a local phenomenon. That is, the analysis of metrical structure only makes sense at small scale. Lerdahl and Jackendoff (1983: 22) even claim that "the large level of metrical analysis is open to interpretation". For this reason, our analysis of metrical structure will focus on small-scale metrical structures.

In our GTTM analysis of the metrical structure in "四季紅", we presented four metrical levels to illustrate the theory. For our readers' convenience, it is given again below.

(26)

Here we will only discuss the largest two levels which directly tell us about the most stressed beat, that is, the whole-note level and the half-note level. Additionally, since the analysis of the whole-note (𝅝) level is grounded on the analysis of the half-note (𝅗𝅥) level, these two levels will be discussed separately.

With respect to the half-note level, there are three constraints involved in our analysis. They are listed as follows.

(27) BINARITY

Every other beat is stressed.

(28) WEIGHT-BY-POSITION

Beats align with the first pitch-event of a group are heavy.

(29) WEIGHT-TO-STRESS[15]

A heavy beat is stressed. The weight of beats is counted on the duration of a pitch-event's time-span, a dynamic, a musical slur, a pattern of articulation, a pitch in the relevant levels of time-span reduction and a harmony in the relevant levels of time-span reduction. The beat with most weight in a group is claimed to be heavy.

BINARITY derives from MWFR 3, which states that strong beats are spaced either two or three beats apart. TERNARITY, which says, "every other two beats is strong", is the other constraint deriving from MWFR 3. The WEIGHT-BY-POSITION constraint comes from MPR 3, which claims that pitch-events coincide with initial beats are heavy, and the WEIGHT -TO-STRESS constraint is based on MPR 5.

Below is an analysis of the first measure of "四季紅". It describes the interaction among these constraints at the half-note level. Note that this piece is in 4/4 meter,

[15] Thanks to Prof. James Myers for giving this constraint a more linguistics-like name.

therefore there are four beats in each measure. The numbers in the tableau stand for the four beats and the Xs beneath them symbolize heavy beats in the level.

Tableau 5: Analysis of the half-note level in the first measure of “四季紅”

 Input: 1 2 3 4	WEIGHT-TO-STRESS	BINARITY	WEIGHT-BY-POSITION
a. 1 2 3 4	*		*
b. 1 2 3 4 x x x x		**	
c.☞ 1 2 3 4 x x			
d. 1 2 3 4 x x	*		*
e. 1 2 3 4 x x	*	*	
f. 1 2 3 4 x x	*	*	
g. 1 2 3 4 x x		*	*
h. 1 2 3 4 x x		*	*

As has been seen in the score in (26), the third pitch-event, which coincides with the third beat, has the longest duration of time-span in the group. Thus, according to WEIGHT -TO-STRESS, it should be a heavy beat. As a result of this, candidates without a heavy-beat mark on the third beat are blocked. Equally, the BINARITY constraint has a great effect on this level, since the tune is in 4/4 meter. It stipulates that stress is imposed on every other beat, which would also favor our optimal candidate. Finally, the WEIGHT-BY-POSITION constraint blocks candidates without a heavy-beat mark on the first beat. Despite this, these three constraints don’t interact with each other within this level, so that we are not able to learn the ranking of them. Probably, it is because the composer doesn't want to violate any.

Now let's look at the whole-note level, which tell us about the heaviest beat, i.e. the most stressed beat. Since the whole-note level is built on the half-note level, we present both levels to illustrate in the tableau.

Tableau 6: Analysis of the whole-note level in the first measure of "四季紅"

 Input: 1 2 3 4	WEIGHT-TO-STRESS	WEIGHT-BY-POSITION
a. ☞ 1 2 3 4 x x x		*
b. 1 2 3 4 x x x	*!	
c. 1 2 3 4 x x x	*!	*

For the same reason mentioned earlier, the third beat is heavy, thus stressed. To prevent candidate (b) from winning, WEIGHT-TO-STRESS must outrank WEIGHT-BY-POSITION. Furthermore, to make sure candidates (a) is the optimal candidate, WEIGHT-TO-STRESS should be ranked above WEIGHT-BY-POSITION. The analysis so far reveals a rough constraint ranking for the metrical structure. The ranking schema is presented in (30).

(30) Ranking schema for Metrical Structure:

WEIGHT-TO-STRESS (, BINARITY) >> WEIGHT-BY-POSITION (,BINARITY)

Since the interaction of constraints in the second measure's metrical structure is the same as in the first measure, we will not repeat the analysis again. Furthermore, by a careful comparison, we can find that the first four measures of "四季紅" in fact have the same regular stress pattern shown in (26).

The analysis above concerned the metrical structure in a 4/4-meter Chinese folk song. Next, we will present a Taiwanese folk song named "補破網" with a 3/4-meter and analyze its metrical structure (See Appendix 5 for the complete score)[16].

(31)

In contrast to a 4/4-meter song, 3/4-meter and 6/8-meter songs have the TERNARITY constraint outranking the BINARITY constraint. This reveals the *co-phonology* issue, which will be discussed later.

---

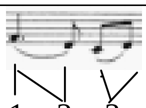
16 The score is taken from the website of 宜蘭縣鄉土教材.
http://media.ilc.edu.tw/MUSIC/MS/ms_music-pic/ms_music-pic23.htm

(32) TERNARITY

Every third beat is stressed.

The following tableau illustrates the interaction among constraints in the first measure. Note that only the dotted-half-note which tells about stress directly is listed in the tableau, due to the fact that a 3/4-meter piece, unlike a 4/4-meter one, doesn't have a secondary stress in the metrical structure. According to Lerdahl and Jackendoff (1983: 19), "if a beat is felt to be strong at a particular level, it is also a beat at the next larger level." Since the dotted-half note is the largest level in a 3/4-meter piece, it is impossible to have a secondary stress in the metrical structure. The numbers in the tableau stand for the three beats, which are associated with the four pitch-events.

Tableau 7: Analysis of the metrical structure in the first measure of "補破網"

 Input: 1 2 3	WEIGHT-TO-STRESS	TERNARITY	WEIGHT-BY-POSITION
a. 1 2 3	*!		*
b. 1 2 3 x x x		*!	
c.☞ 1 2 3 x			
d. 1 2 3 x	*!	*	*
e. 1 2 3 x	*!		*
f. 1 2 3 x x		*!	
g. 1 2 3 x x		*!	
h. 1 2 3 x x	*!	*	*

Since the first pitch-event, which also aligns with the first beat, has a longer duration than the other pitch-events in the measure, it is heavy, and therefore it is stressed by WEIGHT-TO-STRESS. WEIGHT-TO-STRESS filters out candidates without a stress on the first beat, TERNARITY filters out candidates without stresses spaced out every other two beats and WEIGHT-BY-POSITION does the same thing as WEIGHT-TO-STRESS in this score. However, the ranking of these three constraints is not shown in this score. Perhaps a study on a 6/8-meter song, in which secondary stresses exist, will help clarify the ranking. On the other hand, as we can see in the tableau, the WEIGHT-BY-POSITION constraint seems to be redundant in this score. Yet, this constraint does have a great effect on 3/4-meter scores in other music genres, such as Waltz music, as we will see in the following section.

As for the inverse ranking of BINARITY and TERNARITY, we use the co-phonology approach in linguistics to prevent 3/4-meter pieces and 4/4-meter pieces from falling into different genres. In OT, grammars differ in their ranking of constraints. However, it is impossible for us to claim that they belong to different music genres because we can find examples of switching meters in a single Chinese folk song. The following Chinese folk song, 雨夜花, switches meters from 3/4 to 4/4 and also from 4/4 to 3/4.[17]

[17] The score is taken from the website of 宜蘭縣鄉土教材. http://media.ilc.edu.tw/MUSIC/MS/ms_music-pic/ms_music-pic12.htm

(33)

Therefore, we regard 3/4 meter and 4/4 meter as subgrammatical patterns in a single musical grammar, rather than patterns in different music grammars. More on this issue will be discussed in section 3.4.

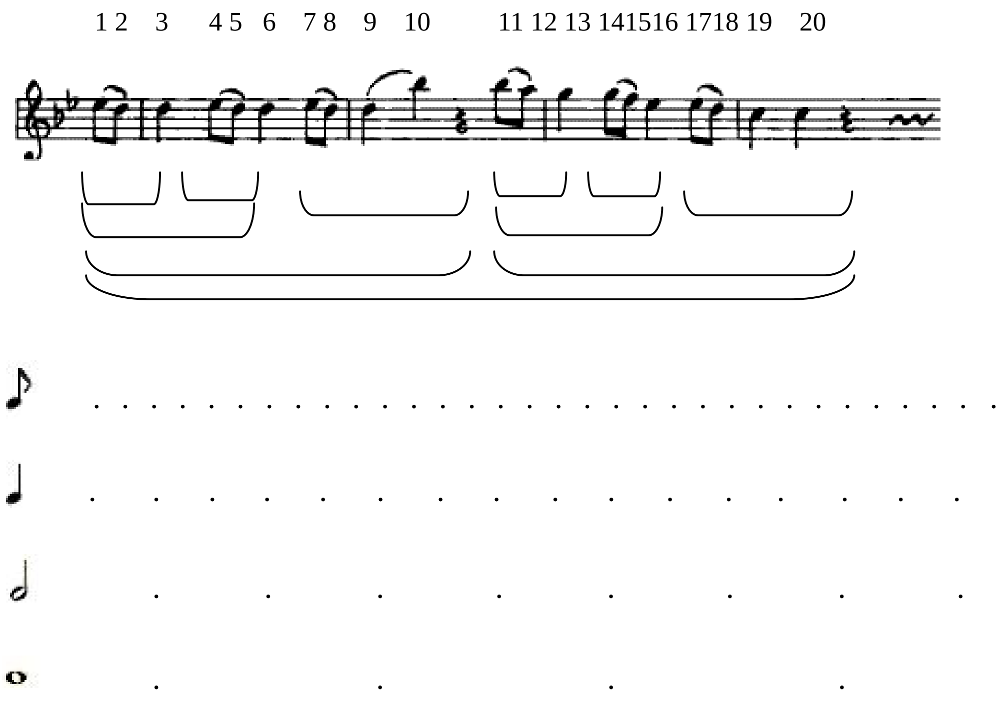
3.3 Comparison With Other Music Genres

In the previous discussion, we have seen the constraints and ranking in Chinese folk songs. Now we will give an OT analysis of a different music genre, specifically Western tonal music, to compare with Chinese folk songs. This comparison might help us identify similarities and differences cross music genres.

The opening of Mozart's Symphony in G Minor, K. 550 is analyzed below.[18] We numbered the pitch-events for our convenience in the discussion.

[18] This is also an example taken from Lerdahl and Jackendoff (1983: 37, figure 3.1 and 86, figure 4.35).

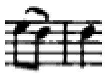
(34)



3.3.1 Grouping Structure in Western Tonal Music

The constraints involved in the grouping structure have already been motivated in our analysis of Chinese folk songs, except for ATTACK(T)ATTACK(T), which places a boundary between the third and the fourth pitch-event. ATTACK(T)ATTACK(T) states that neighboring pitch events that have the same interval of time in between attackpoints put a boundary following the last pitch-event (the constraint is derived from GPR 2b). Since the interval of time between attackpoints from the first to the third pitch-events is the same, i.e. an eighth-note long, there is a group boundary between the third and fourth pitch-events. Besides, ALIGN (Group, Left/Right, Slur, Left/Right) also puts a boundary between the third and fourth pitch-events because there is a musical slur above the fourth pitch-event. However, ATTACK(T)ATTACK(T) is violated in the third group at the smallest level, which has the seventh to the tenth pitch-events. According to ATTACK(T)ATTACK(T), there should be a boundary between the ninth and tenth pitch-event. Yet, these two pitch-events are grouped together by ALIGN (Group, Left/Right, Slur, Left/Right). Therefore we know ALIGN (Group, Left/Right, Slur, Left/Right) must outrank ATTACK(T)ATTACK(T). Below we will give tableaux to illustrate the interaction among the constraints. We start with the first musical group in the smallest level.[19]

Tableau 8: Analysis of the first musical group in the smallest grouping level of "Mozart's Symphony in G Minor, K. 550"

 Input: 1 2 3	MIN-BINARITY	PITCH(X)PITCH(X)	ALIGN (Group, Left/Right, Slur, Left/Right)	ATTACK(T)ATTACK(T)
a. [1][2][3]	***!	*		
b. [1 2][3]	*!	*		
c. [1][2 3]	*!		*	
d.☞ [1 2 3]			*	

As we can see in the tableau, ALIGN (Group, Left/Right, Slur, Left/Right) favors candidate (a) and (b), which violate MIN-BINARITY, and disfavors candidate (d), which obeys MIN-BINARITY; therefore we know MIN-BINARITY must outrank ALIGN (Group, Left/Right, Slur, Left/Right). In addition, two pitch-events that have the same pitch will be grouped together by PITCH(X)PITCH(X). In order to make sure that the second and third pitch-events are in a group, PITCH(X)PITCH(X) must outrank ALIGN (Group, Left/Right, Slur, Left/Right). The ranking we can learn so far is MIN-BINARITY and PITCH(X)PITCH(X) must outrank ALIGN (Group, Left/Right, Slur, Left/Right). But in this group, the constraint ATTACK(T)ATTACK(T) seems can be ranked anywhere.

The analysis of the third group at the smallest level can tell us about where

[19] Note that the musical group here crosses the measures in the score.

ATTACK(T)ATTACK(T) should be ranked. Below is the tableau describing the constraint interactions within this group.

Tableau 9: Analysis of the third musical group in the smallest grouping level of "Mozart's Symphony in G Minor, K. 550"

 Input: 12 3 4	MIN-BINARITY	PITCH(X) PITCH(X)	ALIGN (Group, Left/Right, Slur, Left/Right)	ATTACK(T) ATTACK(T)
a. [1][2][3][4]	***!	*		
b. [1 2][3 4]		*!		*
c. [1 2 3][4]	*!		*	
d.☞ [1 2 3 4]			*	*
e. [1][2 3 4]	*!		*	*

Let's look at the interaction of ALIGN (Group, Left/Right, Slur, Left/Right) and ATTACK(T)ATTACK(T) first. As mentioned earlier, ATTACK(T)ATTACK(T) would favor candidate (c), thus violating ALIGN (Group, Left/Right, Slur, Left/Right) and MIN-BINARITY. To prevent candidate (c) from winning, ATTACK(T)ATTACK(T) must be ranked lower than ALIGN (Group, Left/Right, Slur, Left/Right) and MIN-BINARITY. On the other hand, ALIGN (Group, Left/Right, Slur, Left/Right) favors candidates like (a) and (b). To prevent them from winning, we must have PITCH(X)PITCH(X) be ranked higher than it.

Through our examination of Western tonal music above, we have found that there are the same constraints and ranking for grouping structure between Chinese folk songs and Western tonal music.

3.3.2 Metrical Structure in Western Tonal Music

Next observe the metrical structure of the first group. Again, we will discuss the half-note level and whole-note level respectively. Note that this score begins with an up-beat[20], so we use 4 to indicate the up-beat, which bears two pitch-events, and 1 to represent the first beat in a measure (the score in discussion is in 4/4 meter).

Tableau 10: Analysis of the half-note level in the first group of
"Mozart's Symphony in G Minor, K. 550"

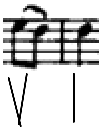
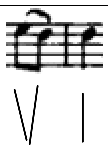
 Input: 4 1	WEIGHT-TO-STRESS	BINARITY	WEIGHT-BY-POSITION
a. 4 1 x	*!		
b. ☞ 4 1 x			
c. 4 1 x x		*!	

Tableau 11: Analysis of the whole-note level in the first group of
"Mozart's Symphony in G Minor, K. 550"

 Input: 4 1	WEIGHT-TO-STRESS	WEIGHT-BY-POSITION
a. 4 1 x x	*!	
b. ☞ 4 1 x x		
c. 4 1 x x x	*!	

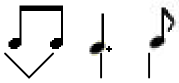
[20] An up-beat is the last beat of any measure, usually a weak beat. It is called "up-beat" because the conductor always directs it with an upward swing of the baton or hand (Randel and Apel, 1986).

Let us look at the constraint interaction shown in tableau 10. Since the pitch-event on the second beat is heavy, it should be stressed. Thus we need the WEIGHT-TO-STRESS constraint for this analysis. Also, in order to block candidates with adjacent stresses, we should also have the BINARITY constraint. In addition, though the WEIGHT-BY-POSITION constraint doesn't have any effect in the analysis, we list it in the tableau to remind our readers this is an example in which the first pitch-event doesn't coincide with the first beat.

To summarize, our analysis of Western tonal music has shown a number of similarities between Chinese folk songs and Western tonal music. For the grouping structure, we have found the same constraints and ranking in these two music genres. Moreover, the same constraints for metrical structure are also found in both of them.

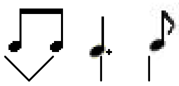
Now we will return to the topic we left in the end of section 3.2 concerning the role of the WEIGHT-BY-POSITION constraint in waltz music. Waltz is a special musical genre that has only 3/4-meter tunes. It is also well-known that the most stressed beat in Waltz is always the first one. In other words, a heavy beat in waltz is assigned by its structure, rather than its weight. Let's consider a four-pitch-event score which consists of two eighth-notes followed by one dotted-quarter-note and one eighth-note. In a musical genre like Chinese folk songs, the stress of the score will fall on the third pitch-event because its musical grammar imposes stresses on beats which have the most weight. Below is a tableau illustrating this. Note that the numbers in the tableau indicate beats, rather than pitch-events.

Tableau 12: Demonstrating the ranking of WEIGHT-BY-POSITION in Chinese folk songs

Input:  1 2 3	WEIGHT-TO-STRESS	TERNARITY	WEIGHT-BY-POSITION
a. 1 2 3 x	*!		
b. 1 2 3 x x x	**!	**	
c. ☞ 1 2 3 x		*	*
d. 1 2 3 x	*!		*

On the contrary, such a score in waltz will always have a stress on the first beat. The tableau below describes this property.

Tableau 13: Demonstrating the ranking of WEIGHT-BY-POSITION in waltz

Input:  1 2 3	WEIGHT-BY-POSITION	TERNARITY	WEIGHT-TO-STRESS
a. ☞ 1 2 3 x			*
b. 1 2 3 x x x		**!	*
c. 1 2 3 x	*!		*
d. 1 2 3 x	*!	*	

Interestingly, we find waltz music has a converse ranking of WEIGHT-TO-STRESS and WEIGHT-BY-POSITION to Chinese folk songs. This finding tells us how music genres differ in their grammars.

3.3.3 Conclusion

In this section, we compared Chinese folk songs with other musical genres and the comparison reveals evidence for music universals. We have found that different musical genres use the same constraints in their grouping and metrical structures, especially in grouping structure. Likewise, in GTTM, the rules for grouping structure are claimed to be more universal than those for metrical structure. Furthermore, some constraints even have the same ranking in different music genres, such as the constraints of grouping structure in Chinese folk songs and Western tonal music. This supports GTTM's claim about music universals. On the other hand, the comparison of Chinese folk songs to waltz music tells us that musical grammars can differ in their constraint rankings. We thus conclude that like in linguistic grammars, constraints for music are universal and grammars of music genres only differ in rankings.

3.4 Resemblances and Differences Between Language and Music

In chapter one, we reviewed the resemblances between language and music found in other studies. Now we will discuss about the resemblances we've found in our OT analysis. Since the constraints we propose for music adhere to the conventions developed for linguistic constraints, our analysis provides more specific resemblances between language and music. On the other hand, since the resemblances cannot be exactly the same, differences between language and music will also be pointed out.

We have found a number of corresponding constraints in language and music. First of all, language and music both use alignment constraints to mark edges. A musical slur in music marks a group, so the ALIGN (Group, Left/Right, Slur, Left/Right) constraint aligns the left and right edge of a group with the left and right edge of a musical slur. With a similar process, a linguistic constraint like ALIGN-RED-L (Kager, 1999: 226) aligns the left edge of a reduplicant with the left edge of a prosody word.

Secondly, the constraint BINARITY is found in the metrical structures of both language and music. In language, stress feet are typically binary. Like heavy syllables in phonology, a strong beat is stressed. In a 4/4-meter piece of music, strong beats are spaced two beats apart at each level in music, that is, the stress pattern is binary. The TERNARITY constraint works similarly. This constraint is disfavored by some linguists, such as Liberman and Prince (1977) and Selkirk (1984) who argue that a monosyllabic word is sometimes claimed to be followed by an empty beat, and therefore its footing is still binary. However, in music, it is difficult for us to assume that there is an empty beat in 3/4 meter, because there are no pauses in musical flows.

The switch of BINARITY and TERNARITY in different measures of the same piece reveals another parallel with language: *co-phonology*. Some linguists, such as Antilla (2002), argue that OT constraints are general but there are distinctive rankings for subgrammatical patterns. It seems that this linguistic phenomenon appears in music too. In the phonology of a language, words linked to different syntactic categories may involve inverse rankings of certain constraints. Analogously, in a musical genre, pitch-events linked to different meters can reveal inverse rankings of constraints, specifically BINARITY and TERNARITY. In a 3/4-meter piece, TERNARITY outranks BINARITY; while in a 4/4-meter piece, BINARITY outranks TERNARITY. With the co-phonology approach, we regard 3/4 meter and 4/4 meter as subgrammatical patterns, rather than different music genres. Note that in waltz music, which is regarded as a musical genre, the situation is not the same. In waltz grammar, TERNARITY always outranks BINARITY.

Thirdly, the WEIGHT-BY-POSITION principle proposed by Prince (1983) and Prince and Smolensky (1993) is also found in music. It shows a structural property in music that pitch-events coinciding with the first beat of a row bear stresses. In addition, our WEIGHT-TO-STRESS constraint is similar to that proposed for phonology whereby heavy syllables are determined by theirs weights. Finally, we can say "groups" in music are like "feet" in phonology, and it is "metrical structures" that spell out the "heads" of the "groups".

Speaking of differences between language and music in our analysis, we have pointed out earlier that in phonology, weights of syllables are quantized, however, in music, weights of beats are gradient. Besides, in music, a composer may intend to follow as many constraints as he can to construct his works. This is quite different from languages. From this aspect, music is much more like artificial and idealized creation.

3.5 Comparison With Other Analyses

3.5.1 Comparison With GTTM

Recall the weaknesses of GTTM that we mentioned in the conclusion of chapter two. We pointed out two problems: some of the rules proposed by GTTM are vague and there seemed to be a ranking hidden behind the rules. Our OT analysis solves both problems. The constraints we proposed are more specific than GTTM's rules. We get rid of modifiers like "strongly preferred" and "weakly preferred". Instead, we make the descriptions of our constraints as clear as possible. Take the MIN-BINARITY constraint for example. It derives from GPR 1 which claims that single-pitch-event groups are strongly avoided. MIN-BINARITY, however, states that a group contains no fewer than two pitch-events.

On the other hand, our OT analysis has shown that the constraints we proposed are actually ranked. The ranking of constraints helps a lot while explaining a musical analysis. In other words, our cognition of music is explained by the interaction of constraints.

3.5.2 Comparison With Other OT Analyses of Music

The constraints used in the OT analysis of Gilbers and Schreuder (2002) simply copy GTTM's rules, such as TSRPR 1 (Time Span Reduction Preference Rule 1). Van der Werf and Hendriks (2004) further modify GTTM's rules and rename the constraints. For example, they turn GPR 1 into the SINGLES constraint: "Groups never contain a single element" and GPR 2a the PROXIMITY SLUR/REST constraint: "No group contains a contiguous sequence of three notes, such that the interval of time from the end of the second note to the beginning of the third note is greater than that from the end of the first note to the beginning of the second". In contrast with these two OT analyses of music, the constraints we propose in this thesis closely match constraints for prosody. What's more, the constraints in Gilbers and Schreuder (2002) and Van der Werf and Hendriks (2004) are still vague. On the contrary, as mentioned above, our constraints are more specific. Take GPR 2a as an example. In our analysis, it is divided into two constraints: ALIGN (Group, Left/Right, Slur, Left/Right), which states the left/right edge of a musical slur must align with the left/right edge of a group, and REST/BOUNDARY, which says that a rest must be a group boundary. These two constraints allow more explicit descriptions of musical structures. Also, though Van der Werf and Hendriks (2004) has modified GTTM's rules, their constraints still lack the clear indication of boundaries, which our constraints possess. They simply claim that pitch-events with identical properties should be grouped together. The absence of clear boundaries would cause problems in an OT analysis.

Chapter 4

An Experiment on Musical Grouping

In previous chapters, we have assumed certain grouping structures in our analyses. But these were based on our own intuitions, and therefore we need to be sure they are reliable. In order to see how the constraints we propose actually work and provide those constraints and their ranking with an empirical basis, a judgment experiment of musical grouping, following Van der Werf and Hendriks (2004), is presented in this chapter. Since the output of a musical grammar is an abstract mental structure that must be tested via judgments, such an experiment is crucial. To find similarities and differences across musical genres, we have two groups of participants with experience in two different music genres, specifically Chinese classical music and Western classical music. As we will see, the result of the experiment supports the reliability of our constraints and their ranking, and even reveals the universality among different music genres.

4.1 Task

A grouping judgment task was used. Participants listened to twenty recordings of five-notes scores and were asked to give their judgments of grouping notes on each score.

4.2 Participants

20 music department students were asked to participate in the experiment, 10 from the Ethnomusicology Department of Nan Hua University and 10 from the Music Department of Chia Yi University. All participants had no problems with hearing. We divided the participants into two groups: one with the average of 6.5 years of playing

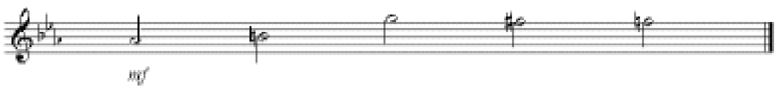
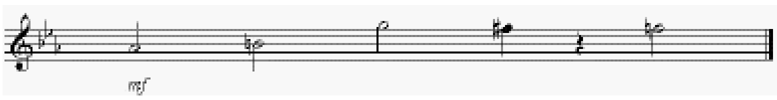
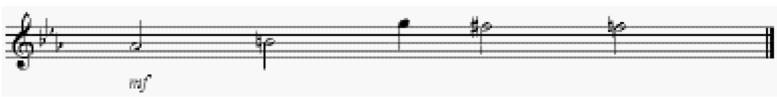
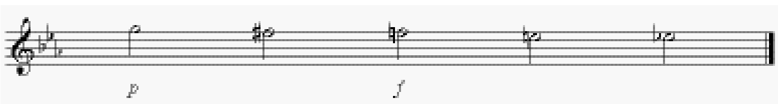
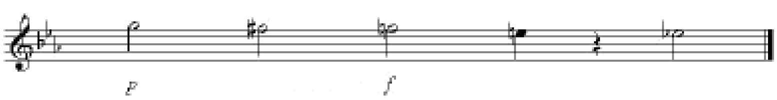
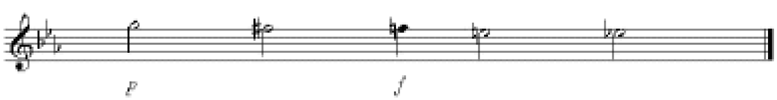
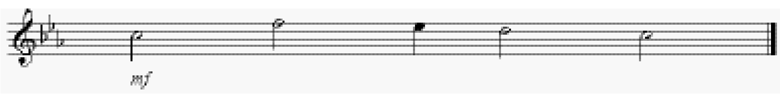
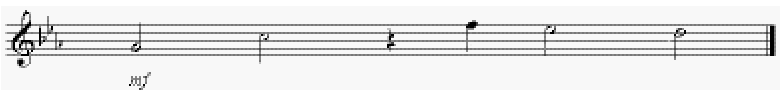
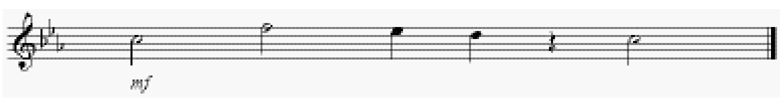
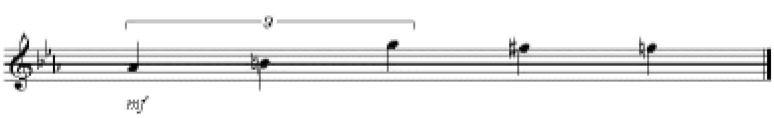
Chinese classical music (8 participants) and the other with the average of 12.3 years of playing Western classical music (12 participants). We included both groups because we might expect them to have different judgments on grouping, because Chinese classical music and Western classical music, as mentioned in chapter one, are claimed in literature to be very different from each other. Owing to differences of participants' training, we expect differences in their judgments would also be found. If so, we would be able to study how exactly these two music genres differ in the ranking of constraints.

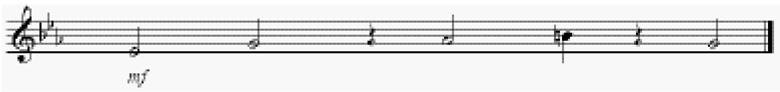
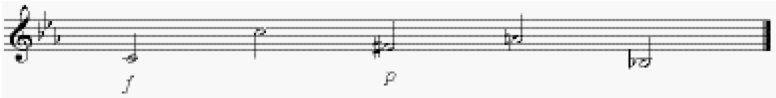
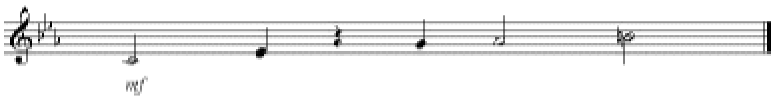
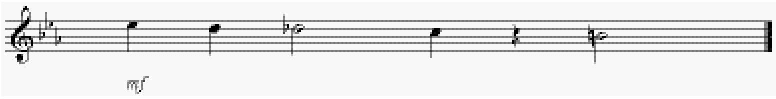
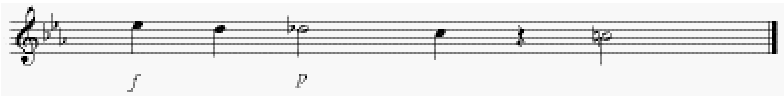
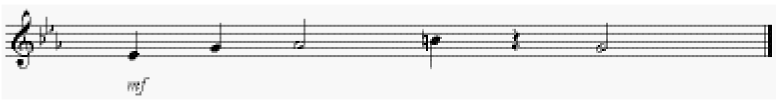
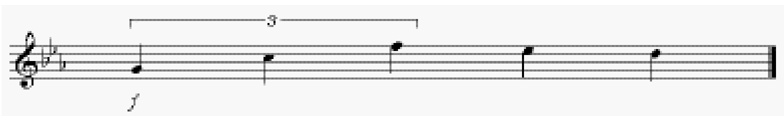
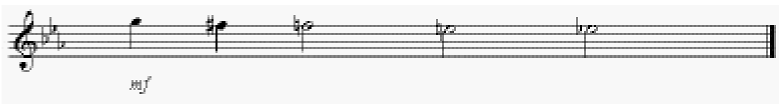
Due to the realities of musical experience in Taiwan, in our Chinese musician group, 6 of the 8 participants also had experience with Western classical music. This might raise an objection that their grammar would be affected by Western classical music. However, it is almost impossible for us to find pure Chinese classical musicians.

4.3 Design and Materials

All of the materials used in this experiment were borrowed from Van der Werf and Hendriks (2004). They reported that the pitches used in the materials were taken from Johann Sebastian Bach's "Theme Regis from the Musical Offering". The twenty five-notes scores were converted into MIDIs by NoteWorthy Composer 1.70. Each score involved different musical grouping parameters. These parameters were encoded in GTTM's GPR 1 (see p. 18), GPR 2 (see p. 19) and GPR 3 (see p. 18). Linking it up with our OT constraints, they were MIN-BINARITY, ALIGN (Group, Left/Right, Slur, Left/Right), REST/BOUNDARY, PITCH(X)PITCH(X), INTERVAL(X)INTERVAL(X), DYNAMICS(X)DYNAMICS(X), ATTACK(T)ATTACK(T) and LENGTH(X)LENGTH(X). In the discussion we will mainly look at the interaction of the constraints applied in chapter three, so we leave the explanation of the additional constraints in Appendix 3, which is a complete list of the OT constraints that we propose for music. Note that the change of articulation (GPR 3c) is not included because Van der Werf and Hendriks claim that it is too difficult to distinguish in musical flow. Figure 1 gives the materials we used in the experiment.

Figure 1: Materials used in the experiment

	score	**competing constraints**
A		none
B		LENGTH(X)LENGTH(X) REST/BOUNDARY INTERVAL(X)INTERVAL(X)
C		LENGTH(X)LENGTH(X) INTERVAL(X)INTERVAL(X) MIN-BINARITY
D		DYNAMICS(X)DYNAMICS(X) LENGTH(X)LENGTH(X)
E		DYNAMICS(X)DYNAMICS(X) INTERVAL(X)INTERVAL(X) MIN-BINARITY REST/BOUNDARY
F		ATTACK(T)ATTACK(T) DYNAMICS(X)DYNAMICS(X) LENGTH(X)LENGTH(X) MIN-BINARITY
G		ATTACK(T)ATTACK(T) INTERVAL(X)INTERVAL(X) LENGTH(X)LENGTH(X) MIN-BINARITY
H		MIN-BINARITY REST/BOUNDARY LENGTH(X)LENGTH(X) ATTACK(T)ATTACK(T)
I		MIN-BINARITY REST/BOUNDARY LENGTH(X)LENGTH(X) ATTACK(T)ATTACK(T) INTERVAL(X)INTERVAL(X)
K		ALIGN (Group, Left/Right, Slur, Left/Right) LENGTH(X)LENGTH(X) ATTACK(T)ATTACK(T) INTERVAL(X)INTERVAL(X)

score		competing constraints
L		MIN-BINARITY REST/BOUNDARY LENGTH(X)LENGTH(X)
M		DYNAMICS(X)DYNAMICS(X) MIN-BINARITY LENGTH(X)LENGTH(X)
N		MIN-BINARITY REST/BOUNDARY LENGTH(X)LENGTH(X) ATTACK(T)ATTACK(T)
O		REST/BOUNDARY ATTACK(T)ATTACK(T) MIN-BINARITY LENGTH(X)LENGTH(X)
P		REST/BOUNDARY DYNAMICS(X)DYNAMICS(X) INTERVAL(X)INTERVAL(X) MIN-BINARITY ATTACK(T)ATTACK(T)
Q		REST/BOUNDARY MIN-BINARITY LENGTH(X)LENGTH(X)
R		MIN-BINARITY LENGTH(X)LENGTH(X) ATTACK(T)ATTACK(T)
S		DYNAMICS(X)DYNAMICS(X) MIN-BINARITY LENGTH(X)LENGTH(X) ATTACK(T)ATTACK(T)
T		ALIGN (Group, Left/Right, Slur, Left/Right) LENGTH(X)LENGTH(X) ATTACK(T)ATTACK(T)
U		ATTACK(T)ATTACK(T) LENGTH(X)LENGTH(X)

According to Van der Werf and Hendriks (2004), 17 of the scores contain conflicts between constraints and 3 of the scores (i.e. A, D and U) satisfy all rules. However, we found there are more than two constraints involved in the 17 scores and even Score D and Score U have a conflict between two constraints. Moreover, since they don't separate GPR 3a into two different constraints, our PITCH(X)PITCH(X) constraint doesn’t have the chance to be tested.

Our hypothesis is that the participants, who are experienced listeners by definition, will be aware of the parameters in each score and their judgments will be affected by the parameters. Therefore, we will be able to study how our constraints compete with each other.

The scores were printed on piece of paper in numbered musical notation. To avoid possible orthographic cues for grouping, those numbers on the printed scores are not indicated with values (e.g. length, dynamic, musical slurs and so on).

4.4 Procedure

Each participant was given a short introduction before the experiment began. Participants listened to the 20 recordings with earphones and were asked to group the notes on the printed scores by circling. The scores were played one by one to the participants. Participants could ask to listen to a recording again. Each recording was not played more than twice.

4.5 Results and Discussion

Following Van der Werf and Hendriks (2004), for each score, we calculate the chance, probability of at least as many participants choosing the most chosen grouping. These p-values were calculated with the formula in (35). Since each score has five pitch-events, there are 16 possible groupings for each score. N is the total number of participants (10, 12 or 8), while K is the total number of participants who gave the most given response (shown in the # column in Figure 2). A value of $p<0.05$ is considered statistically significant.

(35) $$P = \sum_{k}^{N} \begin{pmatrix} k \\ N \end{pmatrix} \cdot \left(\frac{1}{16}\right)^{k} \cdot \left(\frac{15}{16}\right)^{(N-k)}$$

The results are listed below in figure 2. We also list Van der Werf and Hendriks' (2004) results in order to make a comparison. Groupings are represented by numbers and plus marks (e.g. "4+1" represents the grouping [x x x x] [x]). *European musicians* were the participants of Van der Werf and Hendriks (2004), *Western musicians* were the Taiwanese participants who play Western classical music, and *Chinese musicians* were the Taiwanese participants who play Chinese classical music. The # column gives the numbers of participants who gave the response and the P column gives the p-value of each most given response.

Figure 2: most given response (MGR) for each score

Score	European musicians (10 participants)			Western musicians (12 participants)			Chinese musicians (8 participants)		
	MGR	#	P	MGR	#	P	MGR	#	P
A	2+3	6	$1.00x10^{-5}$	5	5	0.00052	5	4	0.008871
B	2+2+1/2+3	4	0.00236	4+1	8	$9.17x10^{-8}$	4+1	4	0.008871
C	2+3	6	$1.00x10^{-5}$	2+3	4	0.00503	4+1/2+2+1	2	0.085
D	2+3	7	$3.78x10^{-7}$	5	8	$9.17x10^{-8}$	5	5	$4.55x10^{-5}$
E	2+2+1	6	$1.00x10^{-5}$	2+2+1	7	$2.23x10^{-6}$	4+1	5	$4.55x10^{-5}$
F	2+3	8	$9.35x10^{-9}$	2+2+1	4	0.00503	2+2+1	4	0.008871
G	2+3	5	0.000184	2+2+1	6	$3.96x10^{-5}$	2+2+1	4	0.008871
H	2+3	7	$3.78x10^{-7}$	2+2+1/2+3	5	0.00052	1+1+2+1	3	0.0108
I	2+2+1	6	$1.00x10^{-5}$	2+2+1	5	0.00052	2+2+1	3	0.0108
K	3+2	5	0.000184	4+1	6	$3.96x10^{-5}$	4+1/5	3	0.0108
L	2+2+1	6	$1.00x10^{-5}$	2+2+1	12	$3.55x10^{-15}$	2+2+1/2+1+1+1	3	0.0108
M	2+3	6	$1.00x10^{-5}$	2+3	5	0.00052	5	3	0.0108
N	2+3	8	$9.35x10^{-9}$	2+2+1	7	$2.23x10^{-6}$	2+3	3	0.0108
O	4+1/3+2	3	0.0210	3+1+1	5	0.00052	4+1/3+1+1	2	0.085
P	4+1	5	0.000184	4+1	4	0.00503	4+1	2	0.085
Q	2+2+1	3	0.0210	4+1	6	$3.96x10^{-5}$	3+1+1	6	$1.5x10^{-6}$
R	3+2	4	0.00236	5	5	0.00052	5/3+1+1	3	0.0108
S	2+3	3	0.0210	5	5	0.00052	3+2	3	0.0108
T	3+2	6	$1.00x10^{-5}$	5	5	0.00052	4+1	3	0.0108
U	3+2	3	0.0210	3+1+1/5	4	0.00503	3+2	3	0.0108

4.5.1 Constraints and Ranking

Since the scores were designed to contain competing constraints, we expected participants' judgments to reveal the ranking of constraints. Let us first look at the results of the Western musicians group. We will only discuss the scores which have more than half of the participants giving a same response, i.e. # > 6 in this group.

In Score B, the most given response (4+1) shows REST/BOUNDARY must be ranked higher than LENGTH(X)LENGTH(X) and INTERVAL(X)INTERVAL(X). Otherwise, LENGTH(X)LENGTH(X) would put a boundary following the third pitch-event and INTERVAL(X)INTERVAL(X) would group the third to the fifth pitch-events together. For Score D, 8 of 12 participants grouped all the five pitch-events as a whole group without being affected by the loudness and softness imposed on the pitch-events. Thus, according to this score, we have LENGTH(X)LENGTH(X) outranking DYNAMICS(X)DYNAMICS(X). Next, in Score E, the five pitch-events have the same sized intervals in between. The most given grouping (2+2+1) reveals that DYNAMICS(X)DYNAMICS(X) outranks INTERVAL(X)INTERVAL(X) because the first to the fourth pitch-events were divided into two groups due to their dynamic differences. Besides, the fifth pitch-event was isolated by a rest, thus violating MIN-BINARITY. Therefore, we know REST/BOUNDARY must outrank MIN-BINARITY. Then in Score N, the most given grouping is 2+2+1. The first group was due to the rest between the second and third pitch-event, which have equal length. It shows REST/BOUNDARY outranks LENGTH(X)LENGTH(X). However, the later two pitch-events with equal length, i.e. the fourth and fifth pitch-events, were divided into two groups, thus isolating the fifth pitch-event. Both LENGTH(X)LENGTH(X) and MIN-BINARITY were violated by some constraint. It seems that the different interval time between these pitch-events' attackpoints has an effect on this grouping. As we can see, the interval time between third and fourth pitch-events' attachpoints is shorter than that of the fourth and the fifth pitch-events. Based on this response, we might have to fix our ATTACK(T)ATTACK(T) constraint a bit, adding this consideration to this condition and have ATTACK(T)ATTACKH(T) outranking LENGTH(X)LENGTH(X) and MIN-BINARITY. Yet, we compare this response with the European musicians and Chinese musicians for Score N, and find that neither LENGTH(X)LENGTH(X) nor MIN-BINARITY were violated. Finally, we will look at Score L, which all of the 12 participants gave the same response (2+2+1). The three constraints competing in this score are REST/BOUNDARY, LENGTH(X)LENGTH(X) and MIN-BINARITY. As shown in the figure, the second and third pitch-events are both half-notes; however, the rest between these two notes caused participants to separate them into different groups. This tells us that REST/BOUNDARY outranks LENGTH(X)LENGTH(X). Additionally, the second rest in Score L breaks the fourth and fifth pitch-events, thus isolating the fifth pitch-event. This result supports our claim in chapter three that REST/BOUNDARY is ranked higher than MIN-BINARITY, which Score B and E also suggested.

The overall ranking of grouping constraints we have discovered in the Western musicians is as in (36).

(36) REST/BOUNDARY>>LENGTH(X)LENGTH(X)>> DYNAMICS(X)DYNAMICS(X) >> INTERVAL(X)INTERVAL(X).

Also, REST/BOUNDARY >> MIN-BINARITY and a revised ATTACK(T)ATTACKH(T) constraint might outrank LENGTH(X)LENGTH(X) and MIN-BINARITY.

Now we will turn to the results of the Chinese musicians. We also look at the scores which which have more than half of the participants giving a same response, i.e. # > 4 in this group. First, for Score D, the most given response is the same as Western musicians: all of the five pitch-events were grouped together. So, we also have LENGTH(X)LENGTH(X) outranking DYNAMICS(X)DYNAMICS(X). But in Score E, the response (4+1), unlike the European and Western musicians, only tells us INTERVAL(X)INTERVAL(X) is outranked by REST/BOUNDARY. Then in the most given response of Score Q, we find ATTACK(T)ATTACK(T) outranks LENGTH(X)LENGTH(X). Since the first two pitch-events have the same length, there would be a boundary between the second and third pitch-events. But there isn't; instead, the first three pitch-events are grouped together because they have same interval time between their attackpoints. Thus, we get ATTACK(T)ATTACK(T) outranking LENGTH(X)LENGTH(X) from this grouping. Aslo, the isolated pitch-event in the end of this score once again shows REST/BOUNDARY outranks MIN-BINARITY in musical grouping. From the observation of Chinese musicians group, we have got two partial rankings of constraints as listed in (37).

(37) a. ATTACK(T)ATTACK(T)>>LENGTH(X)LENGTH(X)>>
DYNAMICS(X)DYNAMICS(X)

b. REST/BOUNDARY >> INTERVAL(X)INTERVAL(X), MIN-BINARITY.

To sum up, the results of the experiment agree with our previous claim that our constraints or GTTM's rules should be ranked in a particular order. Though the results of our experiment didn't reveal a complete ranking of the constraints, they have provided empirical support for our OT analysis. The low chance probabilities of those most given responses show that our participants' judgments are not random. There are principles behind their judgments, namely, constraints of that musical grammars. In addition, through examining the most given responses, we have found that those musical constraints are actually ranked consistently.

4.5.2 Comparison of Groups

In the beginning of this chapter, we have mentioned our intention of comparing Chinese classical music with Western classical music. Since these two music genres are claimed to be so different that we compared the judgments of the Western musicians with the Chinese musicians. We found there were 11 identical most given responses across these two groups of participants. In order to know if this similarity was due to pure chance, again we calculated its probability, using the same formula above. We assumed that the probability of a match between two groups for any score is 1/16, because each group has 16 choices of grouping for each score. The probability of 11 matched responses out of 20 scores is 5.62×10^{-9} ($p<0.05$). Since the p-value is extremely low, this similarity is unlikely to be due to chance. Our comparison of Chinese musicians and Western musicians didn't reveal significant differences; instead, it shows that there is something universal. For this reason, we further compared Western musicians with European musicians and Chinese musicians with European musicians.

Across the Western musicians and European musicians, there were 7 matched responses ($p=0.00014$), and across the Chinese musicians and European musicians, 6 matched responses were found ($p=0.00108$). As one can see, these similarities were also significant ($p<0.05$). That is, these three groups show more matched responses than we would expect by pure chance. It seems there is universality of grouping structure among these three groups of participants. In terms of constraints and ranking, the rankings of grouping constraints in the music grammars that the participants have had in their brains are similar to each other. Connecting this finding with Lerdahl and Jackendoff's (1983) claim that GTTM's rules for grouping structure are universal, we conclude there are indeed music universals across different music grammars, at least in grouping structure.

Chapter 5

Conclusions

This thesis has demonstrated how OT can be applied to music, in which many resemblances with language have been found (Borchgrevink 1982; Lerdahl and Jackendoff 1982, 1983; Jackendoff 1989; Pinker 1997; Sundberg and Lindblom 1991). Through our OT analysis of music, we have found more specific resemblances between constraints for music and prosody, as listed below. These parallels imply that somehow language and music share cognitive mechanisms.

Purpose	Music	Prosody
Making boundaries	ALIGN(Group,Left/Right,Slur,Left/Right)	ALIGN-RED-L, etc.
Feet(Groups)	BINARITY, TERNARITY	BINARITY, TERNARITY
Stress	WEIGHT-BY-POSITION WEIGHT-TO-STRESS	WEIGHT-BY-POSITION WEIGHT-TO-STRESS

Our comparison of Chinese folk songs with other musical genres has also revealed similarities in constraints and differences in constraint rankings. The former supports Lerdahl and Jackendoff's (1983) claim about music universals and the later leads us to claim that as in language, music grammars only differ in the ranking of constraints.

However, further study will need more data to provide more sufficient evidence. Since this thesis only looks at simple structures of music, examination into more complex structures might help establish improved constraints and rankings.

References

English

Anttila, Arto. 2002. Morphologically Conditioned Phonological Alternations. *Natural Language and Linguistic Theory* 20: 1-42.

Borchgrevink, Hans M. 1982. Prosody and Musical Rhythm are Controlled by the Speech Hemisphere. In *Music, Mind and Brain,* M. Clynes, ed., pp. 151-157. New York: Plenum Press.

Cooper, Paul. 1973. *Perspectives in Music Theory: A Historical-analytical Approach*. New York: Dodd, Mead & Company, Inc.

Gilbers, Dick and Maartje Schreuder. 2002. *Language and Music in Optimality Theory*. Unpublished manuscript, University of Groningen, also available as ROA # 517-0103.

Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: The University of Chicago Press.

Jackendoff, Ray. 1989. A Comparison of Rhythmic Structures in Music and Language. In *Phonetics and Phonology, vol. 1: Rhythm and Meter*, P. Kiparsky and G. Youmans, eds., pp. 15-44. New York: Academic Press.

Jackendoff, Ray and Fred Lerdahl. 1982. A Grammatical Parallel Between Music and Language. In *Music, Mind and Brain*, M. Clynes, ed., pp. 83-117. New York: Plenum Press.

Kager, Rene. 1999. *Optimality Theory*. New York: Cambridge University Press.

Landy, Leigh. 1991. *What's the Matter with Today's Experimental Music? : Organized Sound too Rarely Heard.* Philadelphia : Harwood Academic Publishers.

Lerdahl, Fred and Ray Jackendoff. 1983. *A Generative Theory of Tonal Music.* Cambridge, Mass.: MIT Press.

Liberman, Mark and Alan Prince. 1977. On Stress and Linguistic Rhythm. *Linguistic Inquiry* 8 (2), pp. 249-336.

Nicholls, David. 1990. *American Experimental Music, 1890-1940*. New York: Cambridge University Press.

Pinker, Steven. 1997. *How the Mind Works.* New York: W. W. Norton.

Prince, Alan. 1983. Relating to the Grid. *Linguistic Inquiry* 14(1), pp. 19-100.

Prince, Alan and Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Manuscript. Rutgers University, New Brunswick, N.J., and University of Colorado, Boulder. Also available as ROA # 537-0802.

Raffman, Diana. 1993. *Language, Music, and Mind.* Cambridge, Mass.: MIT Press.

Randel, Don Michael and Willi Apel. 1986. *The New Harvard dictionary of music.* Mass.: Belknap Press of Harvard University Press.

Selkirk, Elisabeth O. 1984. *Phonology and Syntax: the Relation Between Sound and Structure.* Cambridge, Mass.: MIT Press.

Sundberg, Johan, and Björn Lindblom. 1991. Generative Theories for Describing Musical Structure. In *Representing Musical Structure,* P. Howell, R. West, and I. Cross, eds., pp. 245-272. London: Academic Press.

Temperley, David. 2001. *The Cognition of Basic Musical Structures.* Cambridge, Mass.: MIT Press.

Thomas, Downing A. 1995. *Music and the Origins of Language.* UK: Cambridge University Press.

Van der Werf, Sybrand and Petra Hendriks. 2004. A Constraint-based Approach to Grouping in Language and Music. In *Proceedings of the Conference on Interdisciplinary Musicology*, R. Parncutt, A. Kessler and F. Zimmer, eds.. Graz/Austria.

Chinese

沈洽 1982 音腔論,《中央音樂學院學報》第四期。

何為 1986 論戲曲音樂的民間性，載於楊蔭瀏等箸，*語言與音樂*，頁 145-167,

台北： 丹青圖書有限公司。

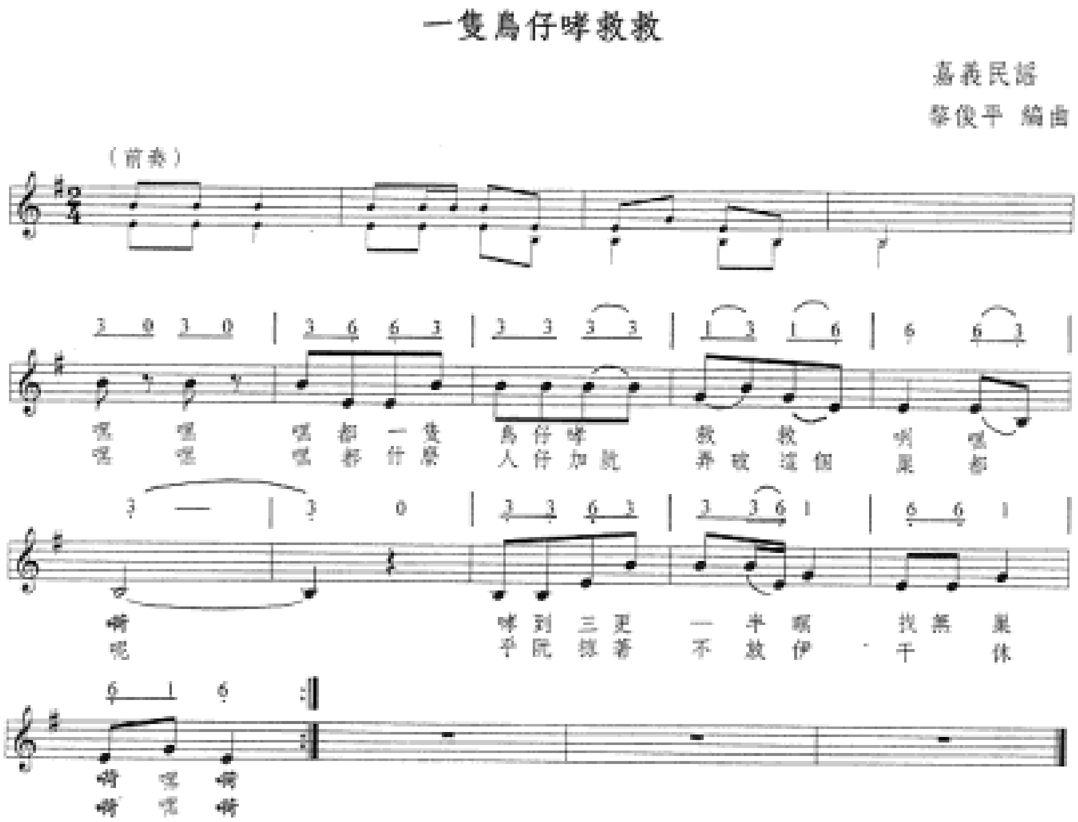
楊蔭瀏 1986 語言音樂學初探，載於楊蔭瀏等著，*語言與音樂，* 頁 1-92,

台北： 丹青圖書有限公司。

Appendix 1: Sample Musical Pieces

1.1

1.2



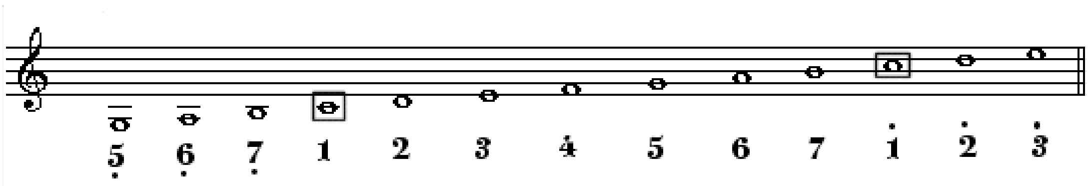
1.3

補破網

李臨秋 作詞
王雲峰 作曲
孫瀛洲 編曲

Appendix 2: Numbered Notation

Numbered musical notation provides a simply way to transcribe music, but it is not useful for complex pieces. The relationship between the numbers and pitches is illustrated as follows.



Lengths of notes and rests also have simplified representations in this notation.

Notes	*Numbered musical notation*	*Rests*	*Numbered musical notation*
whole-note	1 - - -	whole rest	0 0 0 0
half-note	1 -	half rest	0 0
quarter-note	1	quarter rest	0
eighth-note	$\underline{1}$	eighth rest	$\underline{0}$
dotted-quarter-note	1 ·	dotted-quarter-note rest	0 ·
sixteenth-note	$\underline{\underline{1}}$	sixteenth-note rest	$\underline{\underline{0}}$

Appendix 3: Constraints

Constraints for grouping structure

MIN-BINARITY: A group contains no fewer than two pitch events. (GPR 1)

ALIGN (Group, Left/Right, Slur, Left/Right): Align the left/right edge of a musical slur with the left/right edge of the group (GPR 2a).

REST/BOUNDARY: A rest must be a group boundary. (GPR 2a)

PITCH(X)PITCH(X): Adjacent pitch-events that have a same pitch are grouped together. (GPR 3a)

ATTACK(T)ATTACK(T): Neighboring pitch events that have same interval of time in between attackpoints will put a boundary following the last pitch-event. (GPR 2b)

INTERVAL(X)INTERVAL(X): Neighboring pitch events that have same intervals in between will put a boundary following the last pitch-event. (GPR 3a)

DYNAMICS(X)DYNAMICS(X): Neighboring pitch events that have the same dynamics will put a boundary following the last pitch-event. (GPR 3b)

ARTICULATION(X)ARTICULATION(X): Neighboring pitch events that have the same articulation will put a boundary following the last pitch-event. (GPR 3c)

LENGTH(X)LENGTH(X): Neighboring pitch events that have the same length will put a boundary following the last pitch-event. (GPR 3d)

Constraints for metrical structure

BINARITY: Every other beat is stressed. (MWFR 3)

WEIGHT-BY-POSITION: Beats align with the first pitch-event of a group are heavy. (MPR 3)

LENGTH-TO-STRESS: A heavy beat is stressed. The weight of beats is counted on the duration of a pitch-event's time-span, a dynamic, a musical slur, a pattern of articulation, a pitch in the relevant levels of time-span reduction and a harmony in the relevant levels of time-span reduction. The beat with most weight in a group is claimed to be heavy. (MPR 5)

TERNARITY: Every other two beats is stressed. (MWFR 3)