STRUCTURAL SIMILARITIES AND DIFFERENCES IN CHINESE FOLK SONGS

Ravinithesh Annapureddy, Maxime Jan, Zijun Cui
EPFL

ABSTRACT

This study explores the structural differences in a corpus of Chinese Folk Songs and compares their similarity across various geographical regions. A corpus of 2154 musical pieces was selected and categorised into 8 distinct regions. The analysis focuses on the organisation of the compositions around their central pitch. In particular, the concepts of scale degrees, pitch transitions, and their relationship with the metrical structure are studied. The analysis of scale degrees distributions and the use of intervals reveal deep cross-regional similarities. However, the pitch transition asymmetries with respect to the metrical weights show that some regions clearly distinguish from the others. Overall, these (latter) findings highlight the directedness of music across all the examined regions, while also serving as a basis to distinguish region-specific traits.

1. INTRODUCTION

A musical genre can be statistically characterized by using many musical concepts, such as its tempo, beat, or key. These characteristics can also be strong indicators of the cultural affiliation of a particular piece. Amongst those, the concept of tonality is fundamental to the study of Western music [3, 7]. This subject has attracted the attention of many scholars over the past century, but while some theorists restrict its scope to Western music characteristics, others pledge for a more embracing definition that also applies to non-Western cultures [3]. In this work, the concept of tonality was understood in its simplest meaning, as a system for interpreting pitches or chords through their relationship to a reference pitch [3] and was used to characterize folk songs from the Chinese tradition. A similar analysis was performed by David Huron in *Sweet Anticipation* [3]. He analyzed the modal organisation of a Germanic Folk Songs corpus, by computing the scale degrees distributions, transitions, and entropy. He concluded the analysis by stating that his discussion is limited to Western music and that only a few studies have approached this concept in other cultures. Therefore, this work delves into the topic of modal organisation by statistically analyzing a dataset of Chinese Folk Song (hereafter CFS). Similar computational analysis as the one performed by Huron was applied to this non-Western corpus to extract statistical information on the modal organisation. In particular, this study focuses on the similarities and differences in the modal organisation of folk songs between various regions of China.

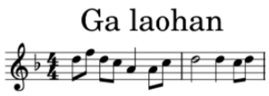
CFS stems from an ancient tradition that dates back to the Classic of Poetry (Shijing) written between the 11th and 7th centuries BC [9]. Through thousands of years of oral transmission and development, these songs have become a mirror of Chinese society, reflecting the lifestyle of the working class. Folk Songs in the Northwest Plateau, for instance, were sung in the fields and rhythmed by manual labour, whereas the prosperity of the Jiangsu and Zhejiang Plains gave birth to the Xiao Diao, songs with a soft and gentle style [1]. Based on this observation and on the theory of Carol Krumhansl [6] stating that tonality may be viewed as a set of statistically learned schema arising from sustained exposure to the music of culture, this study hypothesized that significant differences would be found as deep as in the modal organisation of the CFS of each analyzed region.

To perform this analysis, a dataset of CFS was first collected and partitioned into eight regional categories. Each song's tonic was computed using a custom algorithm, where it is defined as the pitch that is the most recurrent with a strong metrical weight. As the scale degrees are a strong indicator of the modal organisation [3], their distribution is computed in each region. This allows building pitch profiles which are then compared between the regions. This first computation does not take the temporality of the pieces into account. Therefore, the second analysis approaches this concept by considering the intervals. Similarly, their distributions are computed region-wise and then compared to highlight similarities or differences. Finally, the last analysis combines scale degrees and temporality by studying the pitch transitions with respect to their shifts from strong to weak beat strength and conversely. This allows to underline the asymmetries in the compositions of the CFS and to perform a cross-regional comparison.

2. DATA

2.1 Data description

The Essen Folksong Database [2] contains more than 20,000 scores of folk songs, whereof approximately 2250 originating from China are used for this study [5]. This




```
!!!OTL: Ga laohan
!!!ARE: Asia, China,Qinghai, Osten
*M4/4
*k[b-]
*F:
{8dd 8ff 8dd 8cc 4a 8a 8cc
=1
2dd 4dd 8cc 8dd
=2 }
```

Figure 1. First two measures of a CFS with its ***kern* encoding

corpus is in ***kern* [4] format, where each note is represented by a letter and its duration by a number. A ***kern* file also contains the song's metadata such as its title and origin. As an example, the first two measures of a CFS titled *Ga laohan* and their corresponding ***kern* encoding are given in Figure 1. Such encoding allows to access the songs' pitches and their duration, as well as easily computing the beat strength of each note.

2.2 Geographical regions

The songs were divided into nine custom regional categories which were defined based on the province information present in the geographical origins. The regions are shown in Figure 2. The division was based on the description of the geographical distribution of CFS on the Chinese Encyclopedia website [1]. For further analysis, the songs whose province information was absent along with those that do not have song data or have undefined values in the score data were excluded. Additionally, as there were only 5 songs from the Zang region and it is too little data to make any significant claims, the pieces from the Zang region are ignored in this analysis. All in all, 2154 songs, specifically 1054 from Northwest, 287 from Central, 212 from Northeast, 202 from Jiangzhe, 145 from Southwest, 143 from Southeast, 83 from Neimeng and 28 from Jiang were analyzed.

Figure 2. Geographical divisions of China as adopted in this study

2.3 Tonic inference

This study defines the tonic as the most recurrent pitch class that falls on strong beats, and a custom algorithm was created to infer it for each song. In this algorithm, the product of duration and beat strength over all the occurrences of a given pitch class is summed. Further, as the last pitch of the score often corresponds to its tonic in Chinese music [9], it receives an additional weight. Empirically, this weight was defined to be proportional to 10% of the length of the piece. In the end, the pitch class with the highest weight in a song is then defined as its tonic. Scale degrees were then be obtained by translating each score to C relative to their tonic.

3. METHODS AND RESULTS

3.1 Scale degree statistics

The goal of this experiment was to measure the regional differences in the usage of scale degrees. Using the same process as for the computation of the tonic, a weight corresponding to the product of the duration and beat strength is assigned to each scale degree. These weights were tallied for each scale degrees per region, thus obtaining a score denoting their regional importance. The scores of scale degrees per region were normalised and are visualised in Figure 3. The 95% confidence intervals were computed by bootstrapping, in which notes were randomly drawn 1000 times with replacement from the pool of notes of the region created by collecting notes from all pieces of the region.

As the scale degrees have been computed by translating the pieces to C relative to their tonic, in Figure 3, C is to be understood as the tonic, G as the fifth, F as the fourth, etc... Visually, the results of this analysis show a tonal hierarchy very similar to the one computed by Huron on Germanic Folk Songs [3]. A statistical comparison of the distributions between the regions was performed to quantify this result using the Spearman correlation.

Spearman's correlation measures the strength and direction of association by ranking the two datasets. The scale degrees distribution of each region was sorted chromatically and then converted into a vector. The missing scale degrees' of each region when compared to the remaining regions were added to the distribution and set to zero, thereby creating scale degrees vectors with the same dimensions. The similarity between the regions was computed using the Spearman rank coefficient between these vectors of each region, which shows a significant and strong positive relationship between them (Figure 4). It is not surprising to find the tonic as the most frequent scale degree since it is part of how it was defined. However, the fifth, fourth, and second were the next most frequent scale degrees in all regions. This analysis shows a great similarity between regions and thus no support for our initial hypothesis that regions would exhibit significant differences in their modal organisation was found.

3.2 Intervals

When computing the scale degrees' counts, the temporal nature of the piece was ignored i.e. the transitions between them were not considered. In the next experiment, these transitions were considered. For each piece, the diatonic

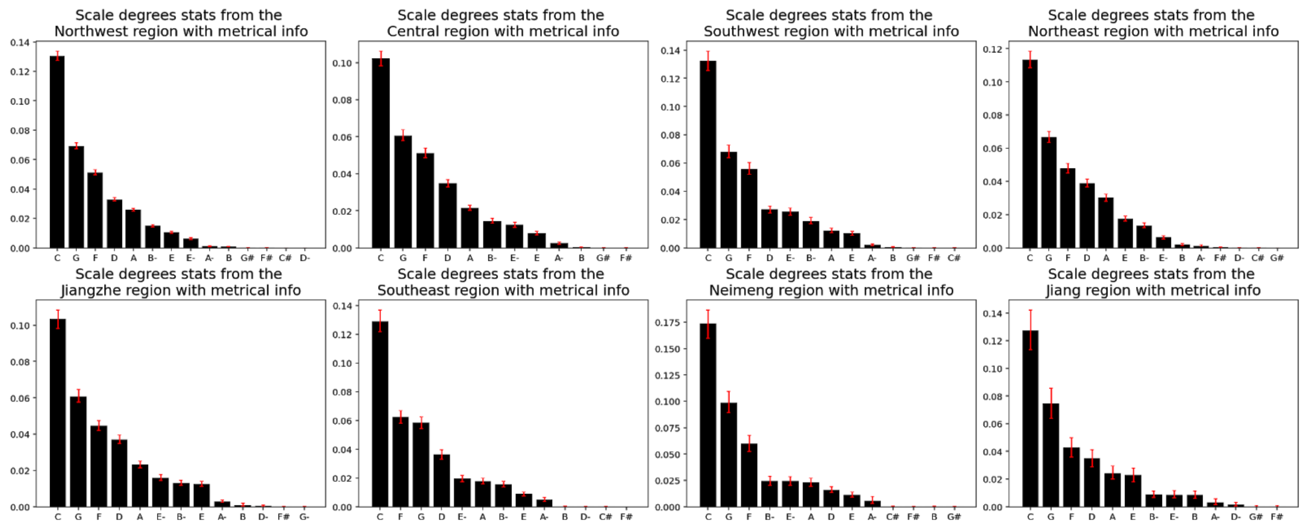


Figure 3. Scale degree statistics of the regions

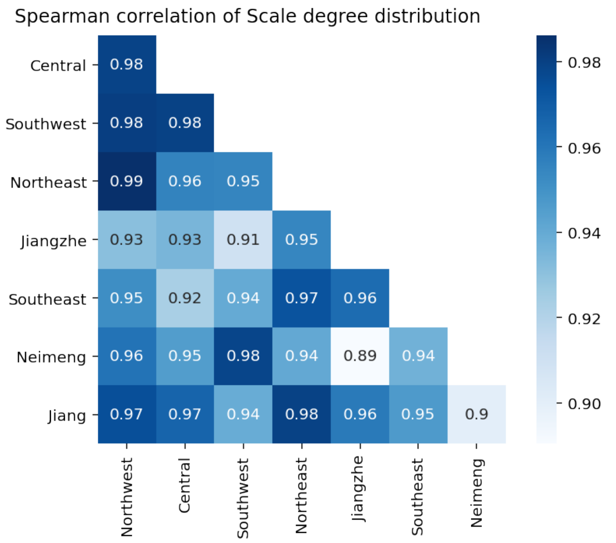


Figure 4. Spearman rank correlation coefficients for the scale degrees distribution (all $p < .0001$)

intervals of consecutive notes are computed and their count is tallied across all pieces of a region. Then, the interval counts for each region were normalized such that the sum of all fractions is equal to 1 and represented as a bar plot in Figure 5. Confidence intervals were obtained similarly as for the scale degrees statistics.

In line with the previous experiment, the regions show a similarity in interval usage as well. The major second is the most used interval, followed by the minor third and perfect unison, except for the Jiang region. This difference in the Jiang region can be probably explained by the fact that this region has a low number of pieces in the dataset. Likewise, to the previous experiment on scale degrees, the similarity in interval usage across regions is quantified using Spearman correlation. The distribution of intervals of each region was sorted alphabetically on the name of the interval and then converted into a vector. The missing intervals of each region when compared to the remaining regions were added to the distribution and set to zero, thereby creating interval vectors with the same dimensions. Figure 6 shows the correlation between intervals usage across different regions. Such quantification shows that there is a significant positive relationship between the regions. This examination also does not support our initial hypothesis. Nevertheless, the correlation values are lower than in the previous experiment, hinting that there could be subtle differences that were not identified in this analysis.

3.3 Scale-degree transitions using beat strengths

The previous experiment considers the temporal dimension of the music but does not fully consider the metrical information. In the third experiment, the metrical information was incorporated in the form of beat strengths. The asymmetry in transitions between strong and weak positions is examined, thereby looking at the interaction of melodic contour with metrical structure. The note with lower beat strength when compared with the other note was considered to be in a weak metrical position and the note with a higher beat strength was considered to be in a strong metrical position. Transitions, where both notes had equal beat strengths, were ignored. This gives rise to two possible transitions 1) transition from weak to a strong metrical position and 2) from strong to weak metrical position. To be able to see the important information while discarding the less informative entries, only the transitions between C, D, E, F, G, A, E-, A- and B- were considered as the modes of the pentatonic scale are composed by these scale degrees when translated to C [9].

In order to understand the asymmetry in transitions, the difference between the number of transitions from a strong to weak metrical position and weak to a strong metrical position were investigated. Mathematically, for each region, the asymmetry in transitions from A to B is defined as,

$$M_{AB} = \frac{\sum(A_S \to B_W) - \sum(A_W \to B_S)}{\sum(A \to B)} \qquad (1)$$

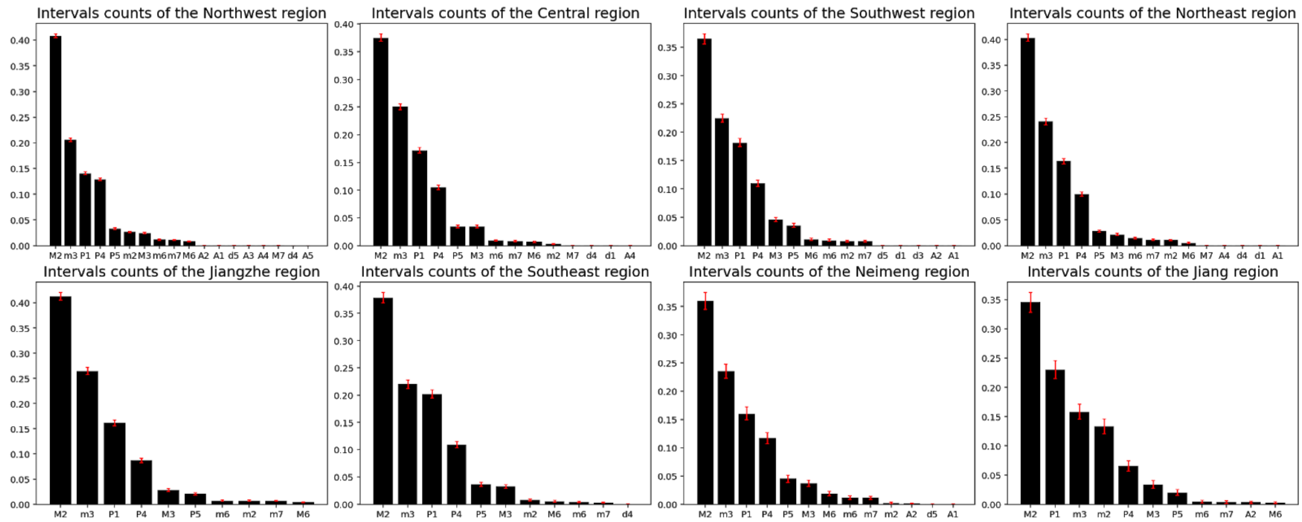


Figure 5. Interval statistics of the regions

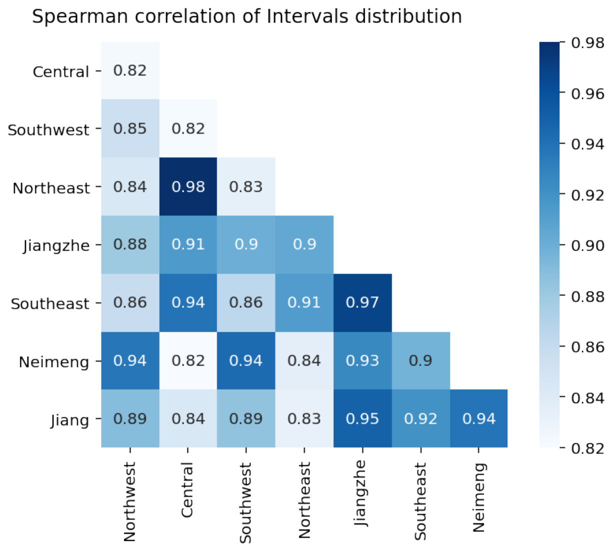


Figure 6. Spearman rank correlation coefficients for the intervals distribution (all $p < .0001$)

Where M is the asymmetry matrix with rows and columns representing the possible starting and landing scale degrees in a transition, subscript S denotes a strong metrical position and W denotes the weak metrical position. $\sum(A_S \rightarrow B_W)$ denotes the number of transitions from A in a strong metrical position to B in a weak metrical position, whereas $\sum(A_W \rightarrow B_S)$ denotes the number of transitions from A in a weak metrical position to B in a strong metrical position and $\sum(A \rightarrow B)$ denote the total number of transitions from A to B including those when A and B have the same metrical weight i.e. same beat strength. Here A and B can be any possible scale degree from the list mentioned above. Figure 7 shows the asymmetry values as a heat map. The red values indicate that the probability of transition from a weak metrical position to a stronger one is more probable than the inverse. The inverse explanation applies to the blue values.

These asymmetry plots indicate a clear directedness in the structure of the songs of each region. While similarities can be observed, such as a weak second translating to a strong tonic, this analysis also displays notable differences. To get a formal understanding of these differences, these asymmetry matrices were compared statistically by flattening them into a vector and comparing them using Spearman rank correlation, similarly as in previous experiments. Figure 8 shows the Spearman rank correlation coefficient between two regions. The comparisons with the Jiang region were not statistically significant ($p > 0.05$) and hence are not discussed further.

With all correlation values below 0.45, the Southwest and Neimeng regions distinguish themselves clearly from the others. This observation also applies to the Southeast region, with a single value greater than 0.5. In contrast, the Northwest, Central, Northeast and Jiangzhe regions appear to be similar as all correlation coefficients are above 0.6. To cluster the regions based on the asymmetries, the flattened vectors of size 81 were reduced to two dimensions using Principal Component Analysis (PCA) and were represented in a two dimensional space. Figure 9 shows this clustering. While Southeast, Southwest and Neimeng, Jiang regions are separated from others and among themselves as well, the Northwest, Central, Northeast and Jiangzhe regions are clustered together. These latter four regions are also geographically closer, as seen visually in Figure 2.

4. DISCUSSION

The main objective of this work was to understand the modal organisation of Chinese Folk Songs to characterize cross-regional similarities and differences. The findings of this study on scale degree frequency reveal a tonal hierarchy similar to the one suggested by Krumhansl [6]. When considering pitch transitions with metrical weights, a significant asymmetry was found. These skewed transitions support the idea of directedness in pitch transitions. These

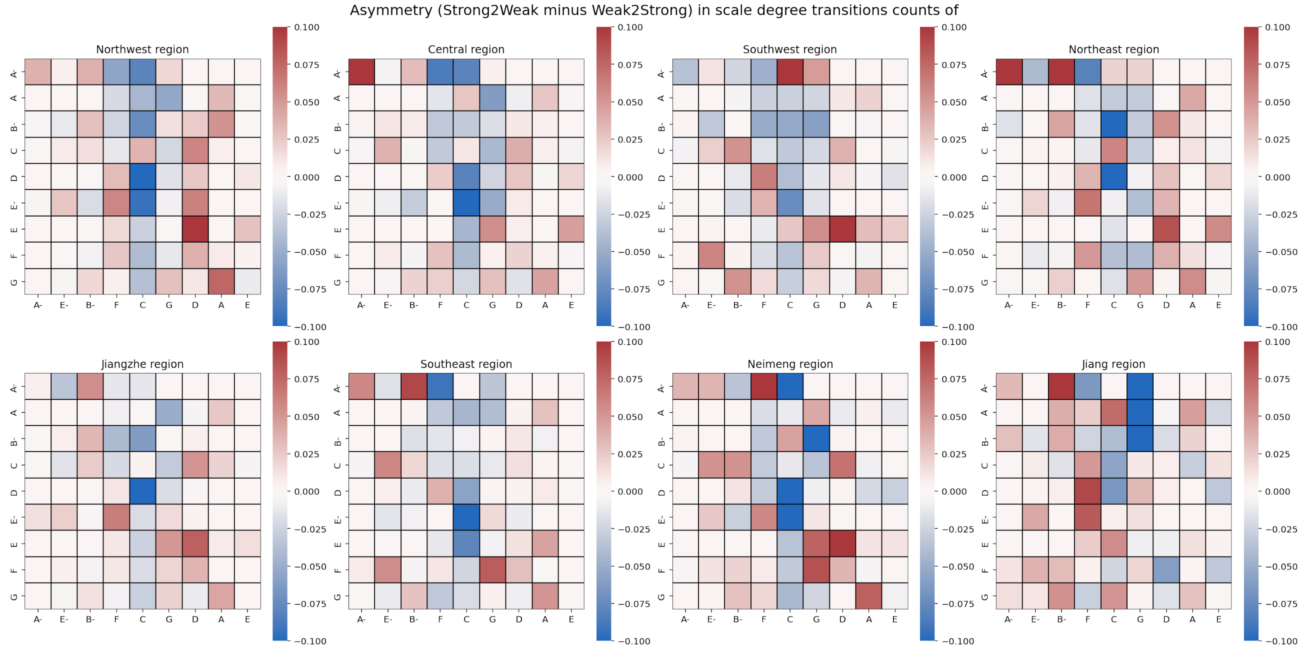


Figure 7. Asymmetry (Strong to Weak minus Weak to Strong) in Scale degree transitions counts of the regions

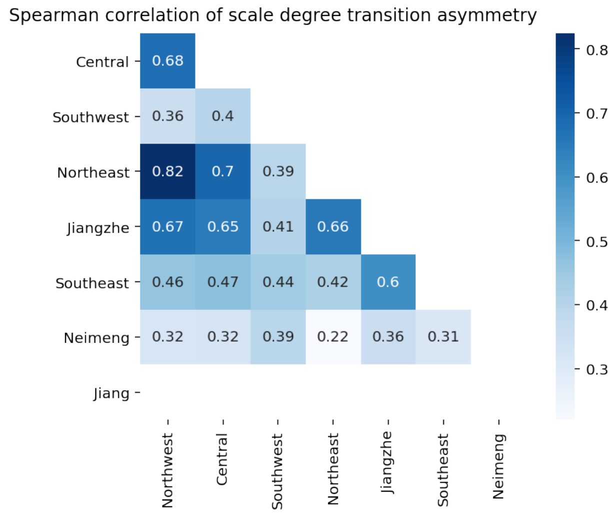


Figure 8. Spearman rank correlation coefficients for the scale degrees transition asymmetry (only significant values at the 95% confidence level are reported)

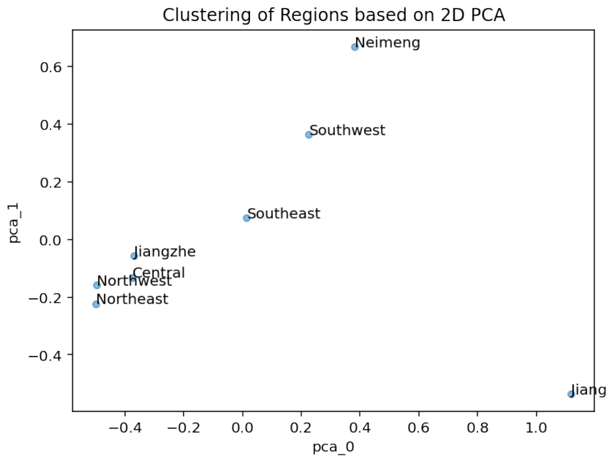


Figure 9. Scatter plot of the two components of the PCA model showing clusters of regions.

characteristics (hierarchicity and directedness) provide evidence for the tonal organisation in the CFS [3, 8]. They also indicate a way to distinguish the folk music produced in various regions, but it cannot be claimed that these statistical features will combine to form a set of rules that govern Chinese folk music. As music-making is deeply rooted in long historical traditions, it reflects subtleties of culture but cannot be merely seen as the notes in a software. Thus quantitatively establishing the asymmetry in transitions is not sufficient to make assertive claims on the regions music-making culture.

5. REFERENCES

[1] Baidu. Chinese folk songs. URL: `https://bit.ly/35dOdWM`.

[2] ESAC Data. URL: `http://www.esac-data.org/`.

[3] David Huron. *Sweet Anticipation: Music and the Psychology of Expectation*. MIT Press, 2008.

[4] David Brian Huron. *The Humdrum toolkit: reference manual*. Center for Computer Assisted Research in the Humanities, 1994.

[5] KernScores. URL: `https://kern.humdrum.org/cgi-bin/browse?l=essen%2Fasia%2Fchina`.

[6] Carol L. Krumhansl. *Cognitive foundations of musical pitch*. Oxford University Press, 2001.

[7] Fabian Moss. *Transitions of Tonality: A Model-Based Corpus Study*. PhD thesis, EPFL, 2019.

[8] Martin Rohrmeier and Ian Cross. Statistical properties of tonal harmony in bach's chorales. *Proceedings of the 10th International Conference on Music Perception and Cognition*, 01 2008.

[9] Jiazi Shi. *East Meets West: A Musical Analysis of Chinese Sights and Sounds, by Yuankai Bao*. PhD thesis, Louisiana State University, 2016.