QUANTIFYING AND QUALIFYING HETEROPHONY IN JINGJU MUSIC

Julian Lenz
Universitat Pompeu Fabra
julian.lenz01@estudiant.upf.edu

Matteo Fabbri
Universitat Pompeu Fabra
matteo.fabbri01@estudiant.upf.edu

ABSTRACT

The field of computational musicology marries the desire to more deeply understand the rich world of musical cultures that surround us with the power of modern day computing resources. To this end, we propose a novel analysis of jingju music, a Chinese theatrical tradition that weaves a narrative fabric from interdependent threads of lyrics, visuals, acting, and music to create vibrant stories that cover a vast range of human experiences. Specifically, we aim to precisely model how much heterophony is present between a given melody and its accompaniment part, and whether there is a consistent pattern between these degrees of similarity and the banshi of the melody in question. This paper therefor has two primary contributions. First, we provide a novel method to quantitatively analyse pitch and rhythmic similarity that can be applied to a variety of musical traditions. Second, we demonstrate a strong correlation between a number of banshis and their degree of heterophony, and argue that this plays a key role in perceptually communicating the desired emotional tonality.

1. INTRODUCTION

The music of jingju (otherwise known as Beijing or Peking Opera) has a rich history and has recently become an increasing topic of interest for musicological studies, such as with [1], [2] and [3]. With the creation of a large and well-structured dataset [4], containing both musical and lyrical features, it has become possible to quantify a variety of features and frame them in a larger cultural context. In this paper, we aim to more deeply understand the relationship that exists between the various *banshi* and their degree of heterophony. First, we will provide a concise set of definitions for *banshi* and heterophony (in addition to other relevant topics, such as *shengqiang* and polyphony), as well as their cultural significance. Then, in our methodology section, we will describe the dataset, computational methods, and how the information is musicologically taxonomized. Finally, we present our results as well as a discussion on their broader implications.

2. JINGJU MUSIC

Although it draws upon a rich variety of histories and traditions, Jingju music is considered to have formalized when it was brought to Beijing in 1790 to celebrate the eightieth birthday of the Qianlong Emperor [5]. Unlike Western classical repertoire of the same era, the music of jingju was not created by dedicated composers, but instead arranged by the performers themselves, who would combine traditional local songs with lyrics in a poetic style. The diverse character-types were eventually simplified to four major roles; the Sheng (main male), Dan (any female), Jing (painted face male), and Chou (male clown). The music, lyrics, costumes and physical acting are deeply intertwined, creating an experience that is theatrical, energetic, and highly stylized (for example, a long journey may be represented by walking several times in a small circle).

An *aria* (single song) will center on one or more vocalists, who are often accompanied by instruments defined as *wenchang* (wind and strings), *wuchang* (drums), and the conductor who is playing a wooden clapper. Every aria can have several sections, each of which is categorized as belonging to a *shengqiang*, which broadly defines the emotional content. Specifically, there are two main *shengqiang*, which are *xipi* (bright, energetic), and *erhuang* (dark, heavy). Furthermore, each section belongs to one of twelve *banshi*, a rhythmic concept that is related to a specific set of emotions, meter and tempo range. Due to the interwoven nature of the performances, the *shengqiang* and *banshi* of a given section will have implications on the lyrics, emotions and tonality of the performance.

		tempo	metre	function
main *banshi*	*manban* (*sanyan*, *mansanyan*)	slow	4/4	introspection, reflection
	zhongsanyan		4/4	
	kuaisanyan		4/4	
	yuanban	**moderate**	**2/4**	**neutral, narrative**
	erliu		2/4	
	liushui		1/4	
	kuaiban	fast	1/4	agitation, nervousness
	sanban	N/A	free	lyricism, emotion
	yaoban	(fast acc.)	free	
auxiliary *ban.*	*daoban*	N/A	free	aria's first line
	huilong	fast	1/4 (2/4)	line(s) after *daoban*
	duoban	fast	1/4 (2/4)	in-line (piled characters)
	kutou	N/A	free	grief (extra characters)

Figure 1. An overview of the most common *banshi*, as well as their associated meter, tempo and emotional function, as per [1].

Etymologically, the word *ban* literally refers to the wooden clappers played by the conductor to keep tempo; and *shi* can be understood as 'pattern'. Therefor *banshi* can be considered a combination of the tempo and metrical pattern (which involves both the time signature and accented beats). As [1] explains, there is no central consensus on the total number of *banshi*, but there are a number that are widely agreed-upon, as detailed in fig. 1. Together, they cover a range of tempos and emotional contexts. In the center, you have moderate tempos with "neutral" tonalities, such as *yuanban*. On the slower end, there are *manban* and *zhongsanyan*, which are reserved for moments of introspection and a somber tonality. In the faster ranges, the meter can become 1/4, or even un-metered ("free"), such as in *sanban* and *yaoban* - these are used for moments of great excitement or intensity, such as battles. What is most critical to understand, is that a single *banshi* has broad implications, affecting not only the time signature, but also the tempo and emotional functionality.

3. MELODIC SIMILARITY

When two or more voices are playing a melody at the same time, the relationship between these melodies can be categorized into one of three 'buckets':

- **Polyphony** - the melodies are distinctly separate from one other
- **Monophony** - every melody is in perfect unison (identical)
- **Heterophony** - The melodies are similar, but not identical

Whilst initially used in Western music traditions, these terms can be applied to practically any variety of musical contexts, provided there are more than one instruments or voices performing a melodic function. It is important to note that, to the best of our knowledge, there does not exist a precise mathematical separation between these three categories. For example, if two melodies share 99 notes, but 1 differs, would this be considered heterophonic or monophonic? To this end, a principal contribution of this paper is an attempt to objectively quantify the properties of each category, whilst acknowledging the inherent ambiguity of such definitions.

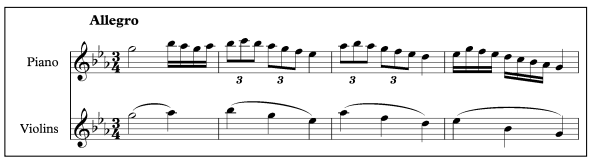

Figure 2. An example of heterophony, in which a violin and piano play distinct variations of a central melodic line. Mozart, Piano Concerto in C minor, K491, first movement, bars 211-214.

4. METHODOLOGY

4.1 Pitch & Rhythm Similarity

Similarity in a multi-voice, melodic context is best understand as a temporal phenomenon; one that is calculated for the duration of the given melody. Furthermore, it is helpful to articulate between *pitch* and *rhythmic* similarities, as an overlap in just one can have perceptual relevance. To this end, we propose a novel system that can, utilizing Music21, calculate the degree of pitch and rhythmic similarity between two melodic lines as a ratio to the temporal duration of the melody.

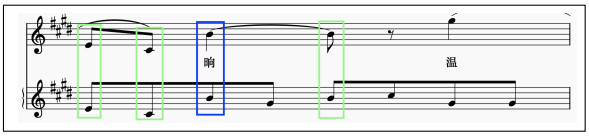

Figure 3. Example of similarity calculation. Green sections would be added to pitch and rhythm similarity sums, and blue would be added to just pitch. *Best viewed in color.*

We start by calculating the entire duration of the melody, which becomes the denominator of the final ratios. Then we proceed to analyse each note in the voice part. For a single note, the system checks for overlapping notes in the accompaniment part. If it is a single note of identical rhythmic value, we consider this rhythmically similar, and add the duration of this note to our running 'rhythmic similarity' total. Similarly, for any note with an identical pitch, we add their rhythmic values to the 'pitch similarity' total. Finally, after all notes have been analyzed and summed, we represent the degree of similarities as a ratio to the duration of the melody.

Figure 2 provides a visual aid to understand the computation process. In this example, the total duration of the voice part would be calculated as 3.5 (note the single eighth-note rest). We would then add the values of pitch similarity (2.0) and rhythm similarity (1.5). When calculated as a ratio to the duration of the melody, this would give us a final pitch similarity of 57% and rhythm similarity of 43%.

4.2 Dataset

The dataset of Jingju Music Scores Collection (JMSC), part of the Jingju Music Corpus (http://compmusic.upf.edu/corpora) and the CompMusic project, has been used. The dataset was created with the purpose of melodic research of Jingju singing lines, and contains 108 MusicXML scores, with a total of 1084 melodic lines, grouped by 3 role types (*laosheng, dan, and laodan*), 4 *shengqiang* (*erhuang, xipi, nanbangzi and sipingdiao*), and 9 *banshi* (*manban, sanyan, zhongsanyan, kuaisanyan, yuanban, erliu, kuai'erliu, liushui and kuaiban*). In particular, the file 'linesdata.csv', containing annotations for each line in the JMSC dataset, has been used to identify the portion of the scores belonging to each line (as start and end offsets, measured in quarter notes duration). We note that the distribution is slightly unbalanced in terms of grouping of *banshi*

categories; 35.1% of the melodies have *manban banshi*, three melody groups of 12.1%, 15.9% and 19.2% have *kuaisanyan, liushui* and *kuaiban banshi* respectively, and the remaining melodies (17.7 %) are distributed over the remaining *banshi*. In other words, 82.3 % of the melodies belong to 4 *banshi* categories (*manban, kuaiban, liushui and kuaisanyan*), and 17.7 % of the melodies belong to the rest of the *banshi* categories.

4.3 Part Extraction

Most of the scores contain a single pair of voice and accompaniment parts, with only a handful containing two vocal parts and an accompaniment part (duets). Hence, for each score, a single voice part and accompaniment part is analyzed. The procedure has been carried out using Music21; counting the parts starting from the one with lower index number (in the score, the first part from the top), the first unvoiced part found (a part containing no lyrics) is labelled as the accompaniment part, and the first voiced part (a part containing lyrics) is labelled as the voice part. This pair of voice and accompaniment parts is utilised to calculate both pitch and rhythm similarity metrics. Due to the low number of occurrences of such duet cases, and to the increased complexity in computing the higher-level statistics and mappings between multiple parameters, we decided not to include this feature in the final work - thus leaving an open door for future research.

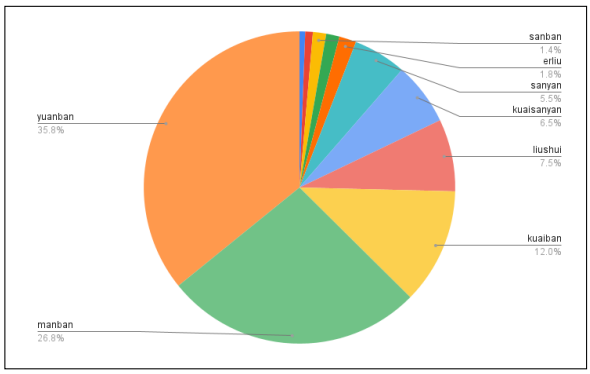

Figure 4. Distribution of melodies per *banshi* categories after processing.

Several additional actions were taken in order to clean the data after processing. Specifically, melodies which did not have an associated banshi in the metadata file were subsequently removed (approx. 2.85%), as these could not provide a meaningful contribution to our research question. Furthermore, in situations were there are no accompaniment lines played throughout the entire duration of the specified melody, it is both possible that the accompaniment is truly silent, or intended to copy the melody note-for-note. Due to the quantitative uncertainty of this, we elected to remove these sections from our final analysis, so as to avoid misrepresenting the degree of heterophony. This accounted for 8.93% of calculated melodies. Finally, a single melody of the 'kuai'erliu' banshi was manually re-taxonified to the 'erliu' banshi, as it can be considered equivalent (and thus contribute to a more balanced data distribution).

5. RESULTS

To quantify the similarity measures in regards to heterophony, we propose a 'threshold' of 20-80%. That is, if a given melody has under 20% similarity in both categories, it is perceptually unrelated from the accompaniament line. Similarly, if it is over 80%, then they are likely playing in near-total unison - thus we could consider it monophony.

5.1 Overall Heterophony

To understand our results from a global perspective, we first present table 1, which details the share of all melodies per bucket of both pitch and rhythm similarity. 50.98% and 44.09% of melodies had virtually no pitch or rhythmic similarity (respectively), whereas 10.04% and 2.56% are in near-perfect unison. Thus, from our total dataset, 38.98% have pitch heterophony, and 53.35% have rhythmic heterophony. [1]

Similarity Score	P-Sim Share	R-Sim Share
0% - 20%	50.98%	44.09%
20% - 40%	3.15%	18.31%
40% - 60%	15.55%	20.67%
60% - 80%	20.28%	14.37%
80% - 100%	10.04%	2.56%

Table 1. Share of melodies per pitch (P-Sim) and rhythmic (R-Sim) similarity score buckets. We propose anything between the ranges of 20 - 80% to have a considerable degree of heterophony.

5.2 Heterophony by Banshi

As a central part of this paper's research question, we aim to understand the relationships that exist between the emotional context of a given Banshi, and the subsequent degree of heterophony. To accomplish this, we have calculated the average pitch and rhythmic similarity per banshi, which is detailed in **table 5.2**.

Banshi	n. Melodies	P Sim (Mean)	P Sim (Std Dev)	R Sim (Mean)	R Sim (Std Dev)
daoban	3	28.79%	49.86	19.70%	34.12
zhongsanyan	4	0.00%	0.00	0.00%	0.00
sanban	7	1.22%	3.24	1.83%	4.85
yaoban	7	6.25%	10.28	6.64%	10.19
erliu	9	68.36%	13.27	54.67%	17.84
sanyan	28	59.93%	18.90	44.35%	16.97
kuaisanyan	33	10.83%	23.75	7.17%	15.80
liushui	38	72.75%	21.98	57.99%	22.41
kuaiban	61	40.80%	36.29	36.60%	33.51
manban	136	24.27%	30.40	24.30%	23.37
yuanban	182	29.45%	32.54	24.29%	25.70
Total	508	32.88%	34.08	28.05%	27.49

Table 2. Average Pitch (P-sim) and Rhythmic (R-sim) similarities per Banshi, as well as standard deviations.

[1] Whilst not a primary focus of our research, it is a notable finding that over half of all sampled melodies have a high degree of similarity with their accompaniment parts - a testament to the interdependent nature of jingju musical structure.

6. DISCUSSION

As discussed in **section 2**, the type of banshi will determine the rhythmic meter, tempo and emotional tonality, which is further communicated by the lyrics, acting and visual elements. We note that three banshis stand out with consistently high degrees of heterophony: *liushui*, *erliu*, and *sanyan*. The first of these two are in meters of 1/4 and 2/4 respectively, with a moderate tempo and sit within the 'neutral' tonality. They typically convey a 'conversational' tonality, which can be emphasized to great effect by a heterophonic accompaniment style, such as in **figure 5**. On the other hand,*sanyan* is the slowest banshi, which is used for elements of introspection and grief.

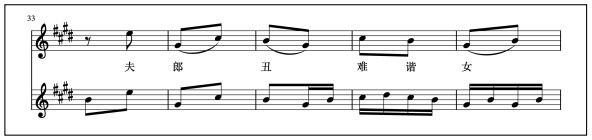

Figure 5. A typical example of a melody from the *liushui* banshi. Note the 'conversational' style with heterophony throughout.

There are also banshi that have little-to-no heterophony; in other words, the accompaniment part is playing lines almost entirely detached from the melody. The lowest scores are found in melodies of *zhongsanyan*, which had 0% similarity in pitch and rhythm - this banshi is similar to *sanyan*, and communicates slow moments of introspection. As detailed in **fig. 6**, the melodies are instead based around a "call-and-response" mechanic, where the voice and accompaniment are taking turns to make melodic statements, each giving space (both musically and emotionally) to the other.

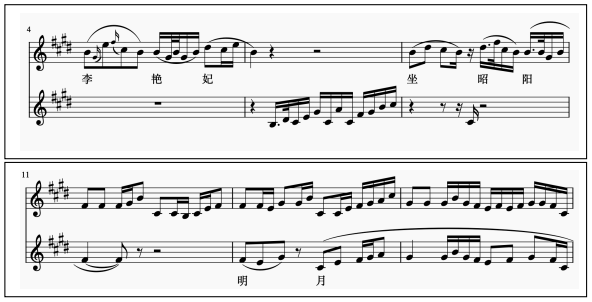

Figure 6. A comparison of melodies from the *zhongsanyan* (top) and *sanyan* (bottom) banshi. Note how, despite a similar goal to communicate a somber tonality in a slow tempo, the former relies on a 'call-and-response' mechanic whereas the latter is more heterophonic.

We also note that there is a remarkable consistency among the main 'family' of banshi. According to [1], *yuanban* is a 'central' banshi, which has a slower variation (*manban*), and a faster one (*kuaiban*), as well as an unmetered variant (*sanban*). All four of these sit in the lower range of pitch and rhythmic similarity scores, with *sanban* representing one of the lowest scores with <2% similarities.

Finally, it is critical to mention that both in *sanban* and *yaoban*, it is common for the accompaniment parts to duplicate the melody, which is commonly notated in the dataset as rests. As discussed in section 4.3, these parts were removed from our analysis as there is no way to quantifiably determine whether these rests communicated duplication or actual silence. We note that the remaining parts that were calculated by our system were clear moments of 'call-and-response'. It can thus be stated with reasonable certainty that, in the *sanban* and *yaoban* banshi, the role of the accompaniment is either to duplicate (monophonic) or to alternate (polyphonic) - but rarely is it heterophonic.

In conclusion, we can determine that there are specific and consistent relationships between a given banshi and its likely degree of heterophony. Specifically, the fastest, unmetered banshi, *sanban* and *yaoban* alternate between perfect unison and "call-and-response", which rarely results in noticeable amounts of heterophony. Similarly, any banshi related to the *yuanban* 'family' (*manban*, *kuaiban*, and aforementioned *sanban*) are likely to have lower degrees of heterophony. On the other hand, the moderately-tempoed *liushui* and *erliu* are consistently some of the most heterophonic banshi. Finally, in the more somber and slow regions, *sanyan* and *zhongsanyan* take remarkably opposite approches, with the former relying on a high degree of heterophony, and the latter consistently utilising "call-and-response" mechanics.

7. REFERENCES

[1] R. Caro Repetto, "The musical dimension of Chinese traditional theatre: An analysis from computer aided musicology," Ph.D. dissertation, Universitat Pompeu Fabra, Sep. 2018. [Online]. Available: https://doi.org/10.5281/zenodo.2030600

[2] S. Zhang, R. C. Repetto, and X. Serra, "Predicting pairwise pitch contour relations based on linguistic tone information in beijing opera singing," in *Proceedings of the 16th International Society for Music Information Retrieval Conference, ISMIR 2015, Málaga, Spain, October 26-30, 2015*, M. Müller and F. Wiering, Eds., 2015, pp. 107–113. [Online]. Available: http://ismir2015.uma.es/articles/282_Paper.pdf

[3] G. Dzhambazov, Y. Yang, R. Caro Repetto, and X. Serra, "Automatic alignment of long syllables in a cappella beijing opera," 06 2016.

[4] R. C. Repetto and X. Serra, "Creating a Corpus of Jingju (Beijing Opera) Music and Possibilities for Melodic Analysis." in *Proceedings of the 15th International Society for Music Information Retrieval Conference*. Taipei, Taiwan: ISMIR, Oct. 2014, pp. 313–318. [Online]. Available: https://doi.org/10.5281/zenodo.1416030

[5] A. L. Silverberg, "A brief introduction to beijing opera," Aug 2020. [Online]. Available: https://www.asianstudies.org/publications/eaa/archives/a-brief-introduction-to-beijing-opera/