Intersections
Canadian Journal of Music
Revue canadienne de musique

Analyzing Heterophony

Jonathan De Souza

Volume 41, numéro 1, 2021

Other Soundings

URI : https://id.erudit.org/iderudit/1114849ar
DOI : https://doi.org/10.7202/1114849ar

Aller au sommaire du numéro

Éditeur(s)

Canadian University Music Society / Société de musique des universités canadiennes

ISSN

1911-0146 (imprimé)
1918-512X (numérique)

Découvrir la revue

Citer cet article

De Souza, J. (2021). Analyzing Heterophony. *Intersections*, *41*(1), 9–27. https://doi.org/10.7202/1114849ar

Résumé de l'article

L’hétérophonie—où plusieurs parts développent indépendamment le même matériau—peut remettre en question les théories de la texture musicale. La recherche psychologique sur l’organisation perceptive auditive soutient un modèle théorique de la texture, créé par David Huron. Son espace de la texture a deux dimensions : synchronisation des débuts et mouvement similaire. L’hétérophonie combinerait peu de synchronisation avec beaucoup de mouvement similaire. Cependant, je soutiens que quelques méthodes analytiques ne peuvent pas détecter cette combinaison et propose une nouvelle technique moins biaisée. J’utilise ce modèle pour analyser des performances de la tradition chinoise *Jiangnan sizhu*, où les instrumentistes improvisent collectivement des ornements sur une mélodie mère. Finalement, les théories de la texture musicale nécessiteront des analyses plus approfondies de divers genres.

ANALYZING HETEROPHONY

Jonathan De Souza

Texture is a fundamental element of music, yet it is marginal to Western music theory. While theorists have extensive terminology for pitch and rhythmic organization, we conventionally place texture into four categories: *monophony*, where one or more parts form a single musical layer; *polyphony*, with multiple independent layers; *homophony*, with multiple, related layers that generally share the same rhythm; and *heterophony*, where multiple parts independently develop the same material. These labels offer a starting point but cannot capture the subtle differences within each category or ambiguous cases where the categories overlap (De Souza 2019, 160–61). By comparison, theorists also use four categories to describe triad qualities—major, minor, diminished, and augmented. But sorting a piece's triads into these categories would offer only a rudimentary analysis of its harmony, overlooking tonal functions, chord progressions, voice leading, and stylistic norms. Similarly, standard textural categories offer only an entry point into textural analysis. We need additional concepts and methods to theorize textural functions, progressions, styles, and interactions among parts.

Moreover, these four textural types are not equally represented across theory pedagogy and research. For instance, by having music students analyze and compose chorales or historically specific forms of counterpoint, our pedagogical contexts emphasize particular types of homophony or polyphony. While this curriculum helps students understand relations among melody, harmony, rhythm, meter, and texture, it limits their textural breadth. The limited scope also plays out in music-theoretical research. For example, David Huron's *Voice Leading: The Science Behind a Musical Art* draws on psychology to examine the relationship between auditory perception and traditional voice-leading rules. His goal is not to naturalize voice leading but to show how textural conventions engage perceptual capacities as they pursue historically and culturally specific aesthetic goals. *Voice Leading* is arguably the most significant recent monograph on musical texture, yet the term "heterophony" appears only once in the book within a list of the four basic categories (Huron 2016, 103). While Huron has considered heterophony elsewhere (as I will discuss), this omission nonetheless reflects the state of the field. Heterophony is studied less than other textural types. In part, this absence reflects Eurocentric biases and the historical use of "heterophony" as an imprecise, often pejorative catch-all for non-Western polyphony (e.g., Adler 1909; for critical discussion, see Napier 2006, 89–93; Coleman 2021, 277–79; and Yu Wang 2023). While texture is marginal within music theory, heterophony is marginal within research on texture.

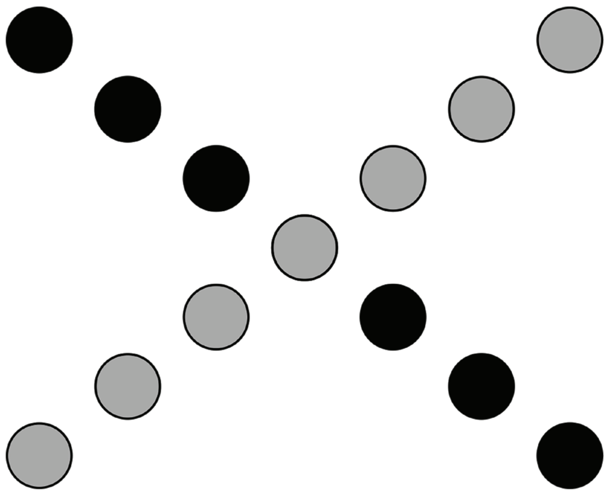
In response to this marginalization, this article begins by examining theoretical approaches to texture that combine cognitive science and formalized quantitative models, then considers how heterophony challenges these models. For example, methods developed for other textures have difficulty detecting certain types of heterophony. To mitigate these problems, I propose potential solutions to this methodological problem and illustrate them through brief analyses of heterophony in the Chinese ensemble tradition *Jiangnan sizhu*. Ultimately, this investigation points to new possibilities for theorizing and analyzing heterophony—and musical texture in general.

Textural Gestalts

Contemporary theoretical approaches to texture have their roots in mid-twentieth-century scholarship. In the 1950s, Leonard Meyer examined texture through a combination of music theory and the psychology of music. At the same time, Meyer used texture to critique these disciplines. One problem, argued Meyer (1956, 5), was that music theory and the psychology of music have traditionally been atomistic: they have treated music as a succession of independent, discrete sounds. But this approach cannot account for relational features such as texture. For Meyer, "texture has to do with the ways in which the mind groups concurrent musical stimuli into simultaneous figures, a figure and accompaniment (ground), and so forth" (185). Texture, then, is not objective. The experience of texture, in which listeners' minds interpret the sounds and their relations, has motivated an interdisciplinary approach to texture that continues today.

Meyer was partly inspired by gestalt theory, an anti-atomistic approach in psychology (Koffka 1922). The image in figure 1 can help illustrate basic gestalt principles. Here I see two dotted lines, one black and one grey. They cross in the middle, forming an X. Objectively, there are no lines, just thirteen circles! But my mind connects the dots. It associates circles in terms of proximity, continuation, and shade. While thirteen discrete objects are difficult to track, the perceived structure—two crossing lines—is simpler. As it simplifies, the mind not only links the visual shapes; it may also add other imagined shapes. If the lighter line seems to cross over the darker one, I might sense that an additional black circle is hidden beneath the grey circle at the centre. My mind would continue the black dotted line to simplify the figure. Moreover, the figure only stands out against some ground—in this case, a blank white background, although the circles could also appear over a coloured background, the surface of a wooden table, or a wallpaper-like pattern. It seems difficult to reverse the figure and ground here, though in other visual examples, such as the vase/faces illusion, the figure and ground are interchangeable, and the transition from one perceived organization to the other is experienced as a sharp switch. This visual example illustrates the relationality of perception, the mind’s associative powers.

Figure 1. A visual gestalt, composed of thirteen circles that suggest two overlapping dotted lines



The “scale illusion” discovered by the psychologist Diana Deutsch offers a musical analogue to figure 1 (Deutsch 2019, 33–38). This illusion involves stereo headphones or speakers, with one melody on the left and another on the right. The sound pattern, notated in figure 2, presents two simultaneous versions of the same major scale, ascending and descending. However, the left and right parts alternate between notes from the ascending and descending lines. Objectively, each part constantly jumps, but listeners do not typically perceive these disjunct melodies. Instead, they hear stepwise lines in contrary motion. In one ear, the melody seems to go down then up; in the other ear, up then down. As with the visual gestalt, the mind connects adjacent points and continues the resulting lines, generating a simpler perceptual organization. Gestalt principles, then, are also relevant to the perceived organization of parts in musical texture.

Figure 2. Sound patterns for the “scale illusion” (Deutsch 2019)

Because of these gestalt principles, multiple parts playing the same melody will generally fuse in our perception. Monophony is likely to emerge with unisons, octaves, and parallel perfect fifths or fourths (as in European traditions of plainchant and organum). But, while acoustic consonance enhances associations among parts (somewhat like similar colours in a visual figure), the intervals do not always have to be consistent or consonant. If the parts are similar enough, the mind might "improve" the figure, overlooking subtle variations. Monophony, like any texture, relies on perception.

While many parts can be heard as one, it is also possible for one part to be heard as many. As Meyer (1956, 186–87) explains, "if the overall articulation is simpler when a piece for a single instrument is understood as implying several 'lines' or voices, then this mode of organization is the one that will probably appear." This often happens with compound melodies, as in Johann Sebastian Bach's music for unaccompanied string instruments. Much as the mind hears stepwise motion instead of leaping parts in the scale illusion, a large leap in a compound melody is generally perceived as a switch to a different, interlocking part. The polyphonic interpretation of a compound melody connects notes that are close in pitch space, even though this separates notes that are adjacent in time. Some excerpts can support both monophonic and polyphonic interpretations. At the beginning of Steve Reich's *Piano Phase*, I usually hear the looping melody as a single monophonic line; as the piece progresses, however, I tend to hear the same pattern as a compound melody, as two parts with uneven, interlocking rhythms (see De Souza 2019, 163). The resulting melodies have no leaps, just steps and one skip. The polyphonic interpretation involves more parts than the monophonic one, but in a sense, it is perceptually simpler. According to the psychologist Albert Bregman (1990, 465), this kind of "virtual polyphony" is "not used simply as a way of getting two instruments for the price of one, but as a way of creating an interesting experience for the listener by providing two alternative organizations."

Bregman discusses psychological principles of auditory perceptual organization in great detail in his 1990 *Auditory Scene Analysis*, arguing for a continuity between ecological and musical listening. In other words, according to Bregman, the perceptual capacities that help us to make sense of complex environmental sounds are recruited for music too. A few of Bregman's core principles are particularly relevant to our current investigation. First, listeners are more likely to connect sounds that start (and end) at the same time. This grouping of simultaneous sounds allows us to hear harmonic intervals and chords. Second, we are more likely to connect sounds that move in the same pitch direction. This grouping of sonic sequences lets us hear extended melodies and rhythms. Both simultaneous and sequential grouping affect the perceived musical streams—the perceived texture (Bregman 1990, 459).

How does heterophony fit into this cognitive perspective on texture? Theoretically, we should perceive heterophony when that organization is simpler than other interpretations, when the music is too diverse for a monophonic interpretation yet too unified for a polyphonic one, or when it supports multiple interpretations. Such music would involve the emergence of a "virtual line" (Campbell 2013, 20–24). It might be described as "monophonic polyphony" or "polyphonic monophony." With heterophony, then, the mind again connects the dots, recognizing that parts are similar but not separate, different but not discrete.

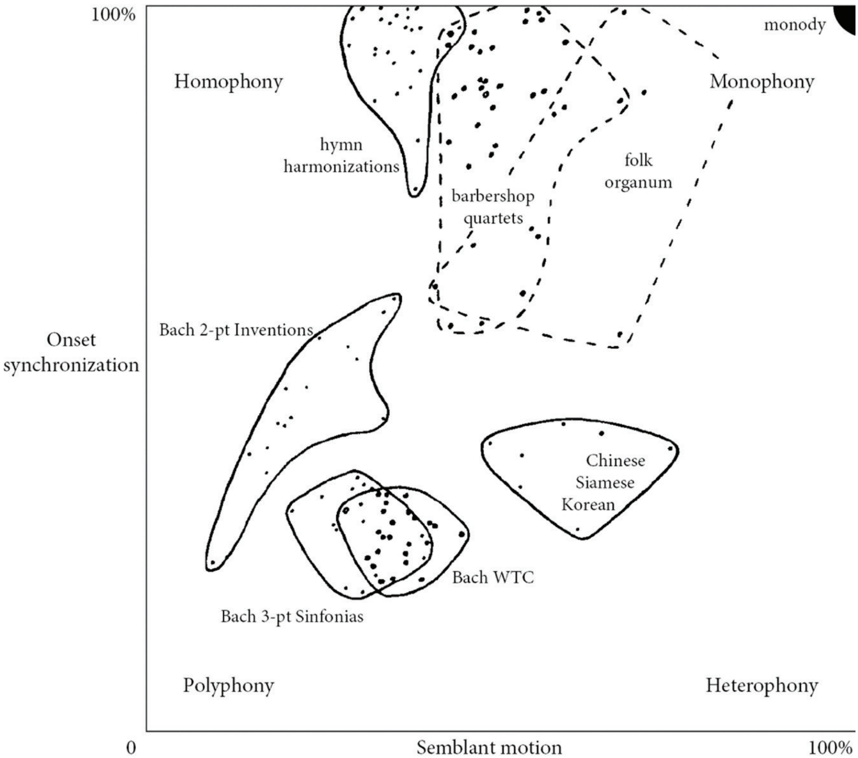
Texture Space

In a 1989 study, David Huron applied these perceptual principles to music theory. He assembled a corpus with more than 450 scores, categorized in terms of the four basic textural categories, including "37 solo folk ballads, 25 four-part hymns, 78 Bach keyboard fugues, and a selection of Chinese, South Indian and Southeast Asian heterophonic works" (Huron 1989, 132). For each score, Huron measured several musical variables, such as the number of parts, pitch, rhythm, counterpoint, and timbre. Then, he used discriminant function analysis to see which combination of musical variables predicted the textural categories. Only two were needed to accurately categorize the scores' textures: "onset synchrony" and "semblant motion." With onset synchrony, Huron calculated "the proportion of synchronous onset pairs—as compared with the total number of onset moments in a work"; with semblant motion, "the proportion of [parallel and similar] motions compared to the total number of contrapuntal motions" (133). Onset synchrony involves simultaneous grouping, while semblant motion mixes simultaneous and sequential factors. This understanding of musical texture elegantly aligns with relevant psychological research.

Huron then set up a two-dimensional space for texture, reproduced in figure 3. The vertical axis represents onset synchrony. Toward the top of the square, the parts are more synchronized. The horizontal axis represents semblant motion. Toward the left of the square, the parts' pitch motion is more independent. These axes, Huron notes, reflect "the inter-stream correlations in the time domain and the inter-stream correlations in the pitch or frequency domain" (133). The space's corners correspond to the basic textural categories. With monophony, both onset synchrony and semblant motion are high; with polyphony, both are low. Homophony combines high onset synchrony with independent motion, while heterophony combines low onset synchrony with similar motion. Each point in the square represents a piece in Huron's corpus. Shapes gather similar repertoire, clarifying that the examples from China, Thailand, and Korea appear in the heterophonic quadrant. Still, because the space is continuous, it not only assigns excerpts to a single category; it also breaks down categorical boundaries. For example, hymns and barbershop quartets are both generally homophonic. But the barbershop quartets have more semblant motion, so they tend to be experienced as more monophonic than the hymns. They represent a homophony-monophony hybrid, combining features of both categories. In my view, this theoretical approach is highly productive, as it both builds on and goes beyond the established concepts for texture.

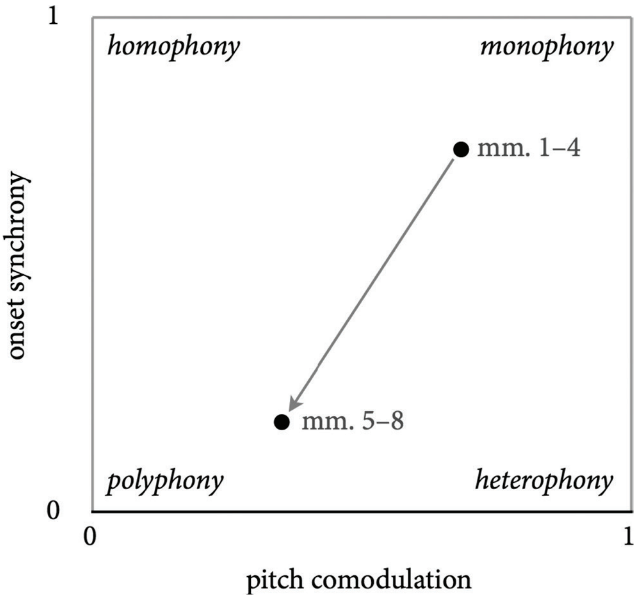
Huron aimed "to provide analytic tools which are less genre-specific and more universal," with applications to "early music, electroacoustic music, and non-western music," among others (131). And he suggested that the space could support analysis of textural variation: "To the extent that dynamic changes of texture are evident in individual musical works, these changes can be traced as *trajectories* through the texture space" (133). Yet Huron's texture space has rarely been used in analysis. That said, I have charted trajectories through texture space in a brief discussion of Bach's Sinfonia 5 in E-flat Major, BWV 791 (De Souza 2019, 166–67). When the alto part enters in m. 2, it loosely doubles the soprano. The interval between these two parts varies, but they share the same rhythm and move almost exclusively in semblant motion. So, these right-hand parts combine in a quasi-monophonic texture against the independent bass line. But in m. 5, the relationship between the soprano and alto changes. Their rhythms and pitch movement become more independent. Figure 4 represents this shift as a move in texture space. This example, however simple, indicates the model's untapped analytic potential.

Figure 3. Texture space from Huron (1989, 134)



To plot a piece in the texture space, we must measure both onset synchrony and semblant motion. Both values will be proportions (i.e., real numbers between 0 and 1, which are equivalent to percentages between 0 percent and 100 percent). I usually obtain these calculations with some Python code that draws on Music21, an open-source toolkit for computational musicology (Cuthbert and Ariza 2010). But understanding how to do the calculations by hand can help to clarify how measurement methods can inadvertently exclude heterophony.

Figure 4. Texture-space analysis of Johann Sebastian Bach, Sinfonia 5 in E-flat Major, BWV 791, mm. 1–8 (De Souza 2019, 166)



Onset Synchrony

To measure the proportion of onset synchrony between a pair of parts, take the number of shared onsets and divide it by the total number of distinct onsets. It is easiest to demonstrate this with an example that has two parts. In the opening measures of Cécile Chaminade's *Idylle*, op. 126, no. 1 (figure 5), there are five shared onsets and twenty distinct onsets overall. So, onset synchrony for the excerpt is 25 percent (5/20). The two hands are fairly independent, and this music would appear in the polyphonic/heterophonic half of texture space (the bottom half).

Another method for measuring onset synchrony has emerged from Ben Duane's research on auditory streaming in classical string quartets. He uses these measurements in complex computational models, though I have combined his techniques with Huron's texture space (De Souza 2019). Duane divides the number of shared onsets by the number of onsets in "whichever part contain[s] more notes" (Duane 2013, 51). With the Chaminade piece, the right hand has twelve onsets, and the left hand has thirteen. The left hand, then, has the "larger" part. With this method, the value for onset synchrony is 38 percent (5/13)—still on the polyphonic/heterophonic side, though closer to the middle of texture space. The two approaches use the same numerator (shared onsets) but different denominators (distinct onsets vs. onsets in the larger part). Their results match when all onsets are shared (100 percent), when no onsets are shared (0 percent), and when the larger part includes every onset in the passage. However, because the number of onsets in one part can never be higher than the total number of distinct onsets, Duane's values will always be equal to or greater than Huron's. As a consequence, the textures might seem more monophonic, homophonic, and unified, such that an excerpt might appear in the heterophonic quadrant with Huron's method and the monophonic quadrant with Duane's method. Even though the values obtained by the two methods probably correlate overall, this subtle difference could affect the number of pieces categorized as heterophonic and polyphonic in texture space.

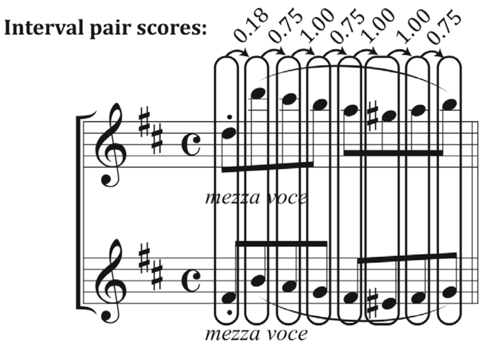
Figure 5. Cécile Chaminade, "Idylle," op. 126, no. 1, mm. 1–4

The difference between these methods might be amplified with pieces that have more than two parts, where onset synchrony is calculated for all possible pairs, and the results are averaged to produce a single value (Huron 1989, 133). Average onset synchrony can be meaningful on its own: in a forthcoming corpus study, my co-authors and I used it to argue that sonata development sections in classical string quartets are significantly more polyphonic than expositions (De Souza, Dvorsky, and Oyon forthcoming). Where Duane calculated onset synchrony in terms of the larger part in a pair, we calculated it in terms of *both* parts and then averaged the two values. Recall that there are five shared onsets in the Chaminade excerpt; the right hand has twelve notes, and the left hand has thirteen. The average of 5/12 and 5/13 is 40 percent—slightly higher than Duane's 38 percent. There is a tradeoff here: our method accounts for all the parts, but in some cases, it increases the value for onset synchrony. It seems important, then, to consider measurement techniques when analyzing heterophony.

Pitch Motion

With the pitch-related dimension, Huron's approach reflects four traditional types of "contrapuntal motion": similar, parallel, contrary, and oblique. With similar motion, two parts move in the same direction. Parallel motion is a special kind of similar motion, where parts move in the same direction as well as the same distance, maintaining the harmonic interval between the parts. With contrary motion, the melodies move in opposite directions (one up, one down), and with oblique motion, one voice moves while the other stays in place. For Huron, semblant motion combines parallel and similar motion (non-semblant motion includes contrary and oblique motion). He divides the number of semblant motions by the total number of contrapuntal motions (1989, 133). An instance of parallel octaves here would be equivalent to a similar motion where one part has an ascending step and the other leaps up an octave. Huron's value for semblant motion would highlight the melodies' common direction without registering the size of the melodic interval.

Figure 6. Method for calculating pitch comodulation (Duane 2013, 51)



Pitch Comodulation = (0.18 + 0.75 + 1 + 0.75 + 1 + 1 + 0.75) / 7 = 0.78

Psychologically, though, there is a difference between parallel and similar motion: as we know from Bregman's research, it is more likely that we will group parallel motion into a single stream (1990, 248–60). As such, Duane developed a method for measuring "pitch comodulation" (2013, 51). He created a mathematical function based on the intervals between two parts. When two successive intervals are the same (parallel motion), the value is 1. With similar motion, the value for pitch comodulation decreases as the difference between the intervals increases, and instances of contrary motion count as 0. Averaging these scores produces a proportion for the excerpt as a whole. Huron and Duane's methods will agree in examples with only parallel motion (100 percent) and contrary motion (0 percent). Because Duane's numbers decrease for similar motion, his values for pitch comodulation will be equal to or lesser than Huron's values for semblant motion, relatively located toward the homophonic/polyphonic side of texture space (the left side). In Duane's illustration (figure 6), pitch comodulation is at 78 percent; for Huron, this would be 100 percent semblant motion (and because onset synchrony is also at 100 percent, this excerpt would appear in the monophonic corner for Huron). With both dimensions, then, Duane's methods are justified in terms of psychology and music theory. However, compared with Huron's original techniques, they produce values that shy away from the heterophonic corner of texture space.

There is another theoretical issue here: Duane's method of measuring pitch comodulation requires synchronized onsets. He calculates pitch comodulation only for "instances in which one line contain[s] two successive notes whose onsets [are] shared by two successive notes in the other line" (2013, 51), meaning that pitch comodulation relies on onset synchrony for Duane. The two variables are not entirely independent. If onset synchrony is at 0 percent, pitch comodulation will also be 0 percent, as there will be no pairs with shared onsets. The method can produce low values in both dimensions, but not low onset synchrony with high pitch comodulation. As such, this method may not adequately account for heterophonic textures, leaving the bottom right corner of texture space empty. Huron is less explicit about what counts as a contrapuntal motion, but his method might have the same bias. His heterophonic examples are close to the middle of the texture space, and the bottom-right corner of figure 3 is relatively unpopulated. Evidently, analytical methods that require onset synchrony are not calibrated for heterophony.

To mitigate this problem, theorists need to segment the music without relying on the rhythms of the parts. I propose a "moving-window" approach that is often used to analyze noisy data. Throughout an analysis, the window would be set at a suitable duration (e.g., an eighth note or quarter note). This consistent window would then "move" through the score. At each step, we would take a pitch-related measurement for the music inside the window. For example, we can calculate the average pitch height for each part using a method from Matthew Poon and Michael Schutz (2015; see also De Souza, Roy, and Goldman 2020, 376): each note's pitch number (A0 = 1, C4 = 40) is multiplied by its duration (quarter note = 1), and the sum of these weighted pitch values is divided by the sum of duration values. Finally, for each window-to-window transition, we can look for semblant motion involving the parts' average pitches. To produce the relevant proportion, count the number of transitions where average pitches move in the same direction and divide by the total number of transitions. (Following Duane, it would also be possible to add a mathematical function to account for differences in distance between the two parts' average pitches). This approach accounts for similar pitch motion in music where parts are not synchronized or move at different rates. As with earlier methods, it gives 100 percent semblant motion for strict monophony; but because it does not prioritize onset synchrony, this method also engages heterophony. To test and refine these methods, then, we must put them to work in analysis.

Analyzing Heterophony in *Jiangnan Sizhu* Performance

Heterophony is the core texture in the Chinese ensemble music *Jiangnan sizhu*. *Jiangnan*, "south of the Yangtze river," refers to its region of origin, which includes the cities of Shanghai and Suzhou. *Sizhu*, "silk and bamboo," refers to the ensemble's combination of string instruments (whose strings were historically made of silk) and woodwinds. Common string instruments in *Jiangnan sizhu* include the *erhu* (two-stringed bowed spike fiddle), *pipa* (pear-shaped, plucked lute), and *yangqin* (hammered dulcimer); wind instruments include the *dizi* (transverse flute), *sheng* (Chinese mouth organ), and *xiao* (vertical, end-blown flute). These are usually accompanied by percussion instruments such as the *ban* (wooden clapper). Traditionally, *Jiangnan sizhu* is played by clubs of amateurs. Whether they meet in private homes or public spaces, such as teahouses, these clubs are devoted to participatory music making, not performing for external audiences. However, conservatory-trained musicians have started to play *Jiangnan sizhu* in a more presentational style (Chow-Morris 2010, 73). In both participatory and presentational settings, members of the ensemble play the same basic melody or "mother tune" (*muqu*), most often one of the Eight Great Pieces (*Ba Da Qu*) that form the core repertoire for *Jiangnan sizhu* (Thrasher 1985, 238; see also, Wong 2024). Upon this melody, each musician improvises an embellished variation, which reflects a personal style and instrumental idiom. They describe this process of variation as *jixing jiahua*, "spontaneously adding flowers" to the melody, and the resulting texture as "*zhisheng fudiao*, 'branch sound polyphony,' often explained as resembling small branches of a river that continually diverge from the main stream and then return to it" (Witzleben 1995, 89, and 106).

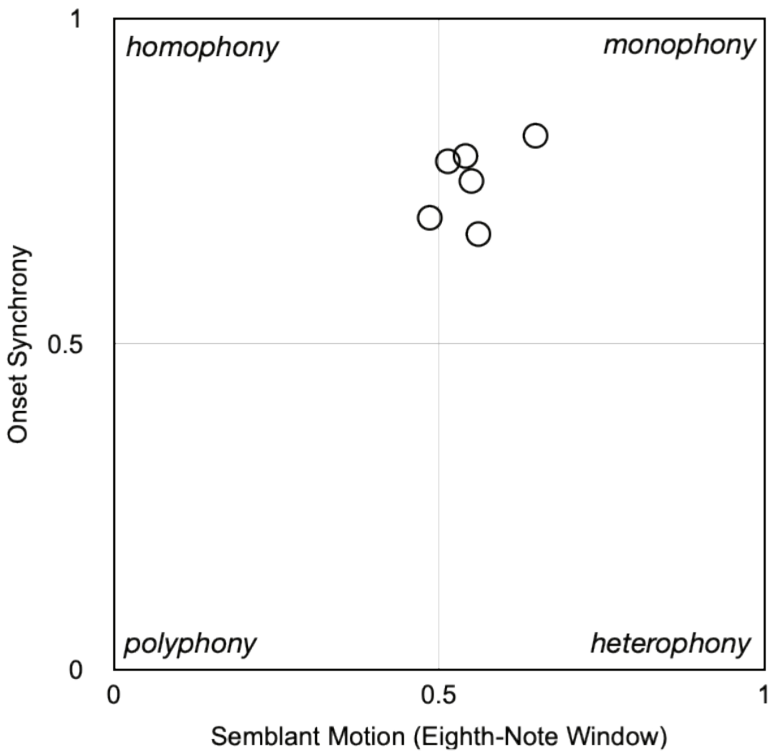
In his book on *Jiangnan sizhu*, ethnomusicologist J. Lawrence Witzleben transcribes passages from five ensemble performances of "Zhonghua Liuban," one of the Eight Great Pieces. Four are played by duos, and one by a trio (Table 1). Witzleben briefly describes each excerpt's texture and raises questions for further inquiry. How similar are the performances? Is the texture consistent throughout each performance, or does it ebb and flow? How do performers interact? To what degree do they follow the aphorism "*ni fan, wo jian; wo fan, ni jian*," meaning "when you play elaborately, I play simply; when I play elaborately, you play simply" (Chow-Morris 2010, 73)? While cultural knowledge is essential for understanding this genre, I hope that my texture-space analysis can complement insights from practitioners and scholars of *Jiangnan sizhu*.

Table 1. Performances of "Zhonghua Liuban," transcribed in Witzleben 1995

Performance 1	Sun Yude (*xiao*) and Chen Yonglu (*erhu*)
Performance 2	Chen Yonglu (*erhu*) and Zhang Zhengming (*yangqin*)
Performance 3	Zhou Hao (*erhu*) and Zhou Hui (*yangqin*)
Performance 4	Lu Chunling (*dizi*) and Zhou Hui (*yangqin*)
Performance 5	Cai Cide (*xiao*), Liu Yuehua (*erhu*), and Shi Quan (*pipa*)

Overall, the five performances' textures are highly similar. Average values for each performance, plotted in figure 7, are close in texture space. That consistency seems to reflect the established textural norms for "Zhonghua Liuban," and possibly *Jiangnan sizhu* as a genre. This "branch sound polyphony" appears at the edge of the monophonic quadrant in texture space. Because the musicians present the same mother tune, more than half of the transitions involve semblant motion (55 percent on average, measured with an eighth-note window). Because they lock on to a consistent beat, usually established by a percussionist, the players also share a substantial proportion of onsets (75 percent on average). Yet each player's improvised embellishments add variety, which keeps the texture away from the strict monophony of the space's top-right corner. This type of heterophony differs from Huron's heterophonic examples. It might be understood as "polyphonic monophony".

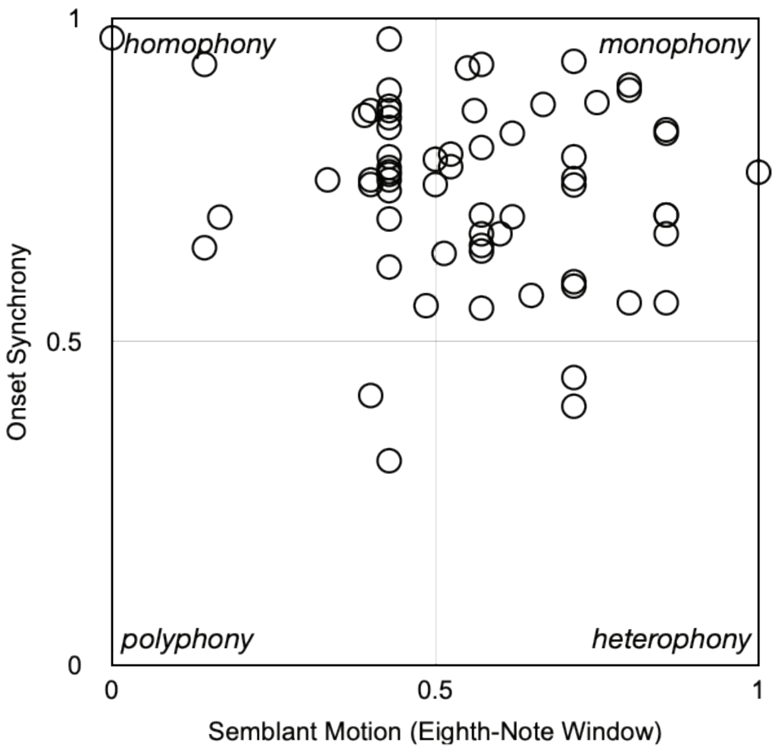
Figure 7. Texture-space analysis of "Zhonghua Liuban", mm. 9–21, where each point represents an entire performance



While average values for each performance indicate the overall texture, they obscure moment-to-moment variation. Figure 8 is based on the same data as figure 7, but each point represents one measure instead of an entire performance. Most points cluster around the same region as the averages. For example, Sun Yude and Chen Yonglu's texture is close to the average in mm. 16–17 (figure 9). On almost all downbeats, they play the same pitch, and they share three-note melodic fragments on the last beat of m. 16 (B, A, F#) and m. 17 (E, D, B). They are playing the same melody in different ways, and the result is characteristic of these performances across the board.

Nonetheless, at different times, this group of performances passes through all four quadrants of texture space. A performance of the same section by Chen and Zhang Zhengming starts in a similar zone for m. 16, then moves into the homophonic quadrant of texture space. As figure 10 shows, they share many onsets in m. 17 (59 percent). They play the same pitch class on beats 1–3, but their melodies move in different directions, with only 14 percent semblant motion. Zhang's idiomatic playing on the *yangqin* contributes to both dimensions: the instrument's precise, percussive attacks might enhance the sense of rhythmic integration, while its prominent octave leaps add melodic variation. The lines here have relatively coordinated rhythms and independent pitches.

Figure 8. Texture-space analysis of "Zhonghua Liuban", mm. 9–21, where each point represents a single measure



A performance by Lu Chunling and Zhou Hui involves larger textural shifts (figure 11). In m. 9, they share most pitches, and their average pitches create the same virtual contour (semblant motion is at 100 percent). Yet because of Lu's grace notes on the *dizi* and Zhou's *yangqin* tremolo, less than half of their onsets are synchronized (44 percent). The measure appears in the bottom-right quadrant of texture space, which Huron labelled as heterophony. After the strong pitch coordination of m. 9, however, Lu and Zhou improvise a new texture in m. 10. They share pitches on beats 1–3, but their melodies have different shapes: Lu ascends to an A5, which is the highest pitch in the excerpt, while Zhou plays a more active series of zig-zagging sixteenth notes. The parts become more independent: average semblant motion drops to 17 percent and onset synchrony to 32 percent. As a result, the texture moves into the polyphonic quadrant of texture space. This momentary textural change is striking, though on average, the entire performance is close to the other performances. Texture in *Jiangnan sizhu*, then, exhibits global consistency but also substantial local variation—variation that likely reflects broader aesthetic values. According to Alan Thrasher (1985, 246), "Both in landscape painting and traditional music, lines and melodies were thought best if they flowed freely and naturally—as if to reflect a sense of growth from within." This sense of flow and growth might be desirable not only for individual melodies but also for the textures they produce.

Figure 9. "Zhonghua Liuban," performed by Sun Yude (xiao) and Chen Yonglu (erhu), mm. 16–17 (for full transcription, see Witzleben 1995)

Figure 10. "Zhonghua Liuban," performed by Chen Yonglu (erhu) and Zhang Zhengming (yangqin), mm. 15–18 (for full transcription, see Witzleben 1995)

Other aspects of the data can help us evaluate the performers' interactions. The number of onsets in each part is already calculated when determining onset synchrony, and this value can be used to compare their complexity because a higher number of onsets should correlate with more elaborate ornamentation. Consider the performance by Chen and Zhang (presented earlier in figure 10). As Witzleben (1995, 111) notes, "The *yangqin* part is the denser of the two throughout this excerpt, but there is considerable give and take between the parts." From mm. 15–18, the players engage in subtle trading (figure 12): in odd-numbered measures, Chen plays more onsets on the *erhu*; in even-numbered measures, Zhang plays more onsets on the *yangqin*. The players alternately increase and decrease their number of onsets per measure, while gradually increasing the rhythmic density. In this performance, the performers seem to follow the principle of "*ni fan, wo jian; wo fan, ni jian*."

Figure 11. "Zhonghua Liuban," performed by Lu Chunling (dizi) and Zhou Hui (yangqin), mm. 9–10 (for full transcription, see Witzleben 1995)

Yet this trading does not occur in all the performances. For example, in a performance by Zhou Hui and Zhou Hao, Witzleben observes that "the two parts tend to become dense and sparse simultaneously rather than in alternation," so that "the ensemble aesthetic . . . is somewhat different from the ideals of contrast and exchange so common in writings on *Jiangnan sizhu*" (1995, 111–12). The chart of rhythmic density in figure 13 is consistent with Witzleben's interpretation, yet it also suggests that the *yangqin* initiates increases in density. Twice, the number of onsets in the *yangqin* increases first, the *erhu* follows, then both relax and leave more space. This duet between brothers who have played together for decades suggests that *Jiangnan sizhu* involves varied modes of interaction.

At the same time, the texture is shaped not only by interaction but also by the players' instruments and personal styles. Witzleben even argues that "one could recombine parts transcribed from different recordings . . . with little change in the overall musical texture of either example" (1995, 115). He specifically imagines exchanging Chen's *erhu* melodies from Performances 1 and 2. Texture space can facilitate comparison of the original and recombined versions. My analytical results are consistent with Witzleben's claim: the overall texture of the recombined performances is close to the other performance-level averages (56–61 percent, semblant motion; 72–74 percent, onset synchrony). For Witzleben, "the simple/complex contrasts mentioned in aphorisms . . . seem to be descriptions of a musical result rather than that of a process, and the constantly shifting musical texture owes as much to differing individual styles as to conscious spontaneous interaction among the performers" (1995, 115). This apparent interchangeability of parts might also rely on the mother tune. For performers, the structural melody grounds each improvisation; for listeners, this melody appears as a virtual line, a simpler gestalt that emerges from complex ensemble playing. Although the mother tune is not directly played, it is essential for this type of heterophony.

Figure 12. Onset density in "Zhonghua Liuban," performed by Chen (erhu) and Zhang (yangqin), mm. 15–18

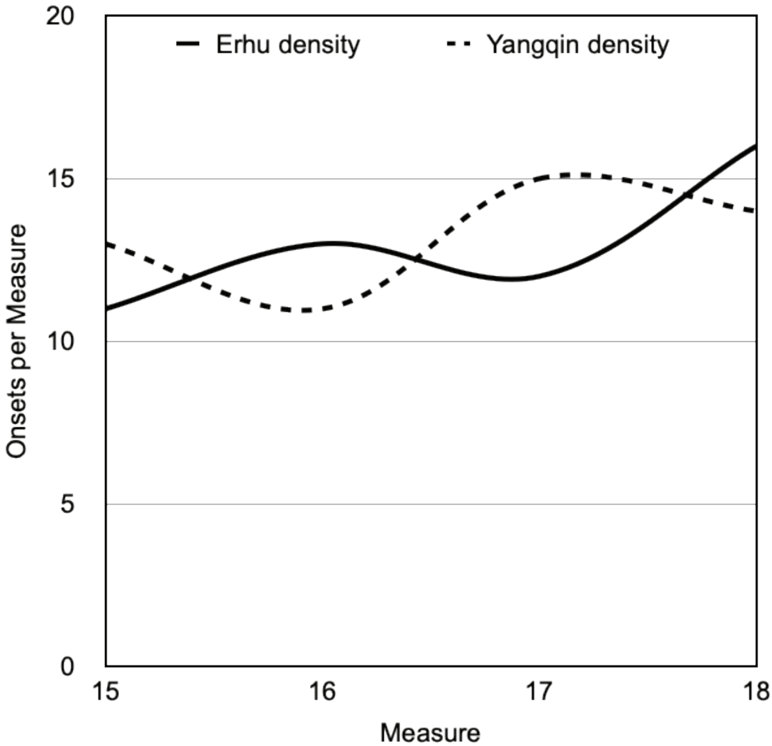
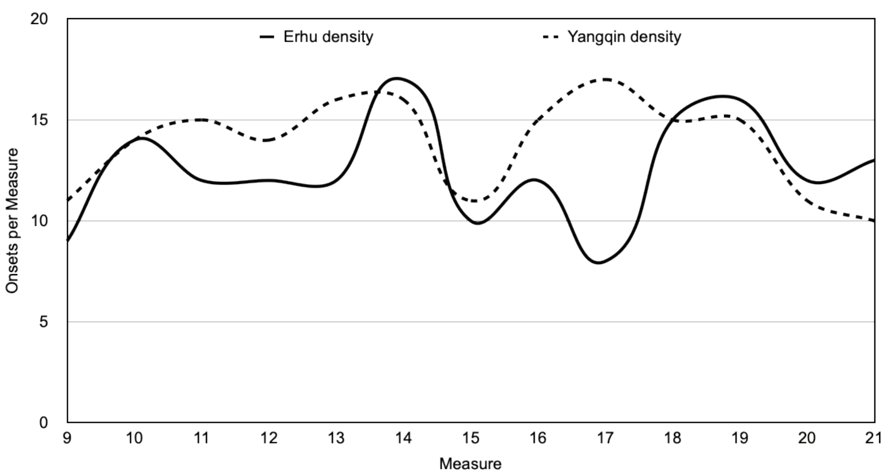


Figure 13. Onset density in "Zhonghua Liuban," performed by Zhou Hao (erhu) and Zhou Hui (yangqin), mm. 9–21



Conclusion

In this article, I have examined heterophony—and texture in general—from three interrelated perspectives. First, I have considered texture from the standpoint of music cognition, which emphasizes how texture involves musical structure and listeners' perception. In this approach, heterophony can be understood as a texture in which multiple parts create an emergent melodic gestalt, which seems both unified and varied. Second, I have touched on music theory's formal tools for representing texture, noting that Huron's two-dimensional texture space is particularly useful, though analytical methods must be adapted to account for heterophony. Finally, I have showed how music analysis, informed by ethnomusicology, supports close readings of specific heterophonic performances. Each perspective can support future research on heterophony, which might involve psychological experiments, new theoretical models, and further analysis.

Moreover, all of these perspectives could interact in productive ways: cognitive science may help to shape theoretical models, and models may guide analysis. Analysis can also challenge and refine the theory and psychology of musical texture, revealing theoretical biases or generating testable hypotheses. For example, analyzing performances of "Zhonghua Liuban" complicates the association between heterophony and the bottom-right quadrant of texture space. The "branch sound polyphony" of *Jiangnan sizhu* represents a particular type of heterophony, based on improvised embellishment of a mother tune and metrical coordination. Yet these principles are not shared across heterophonic musics: pre-existing structural melodies are not required for heterophonic *sangat* (vocal accompaniment) in North Indian music, where an instrumentalist imitates and elaborates an improvised vocal line in real-time (Napier 2006, 93); synchronization is looser in Scottish Gaelic psalm singing (Campbell 2013, 20–24); and the dense group improvisation of Ornette Coleman's *Free Jazz* is fundamentally decentralized (Coleman 2021, 279), drawing on a "heterogeneous sound ideal" expressed in various African and Afro-diasporic musics (Wilson 1992). This leads us to a critical point: different types of heterophony might inhabit different regions of texture space. Ultimately, a robust theory of musical texture requires analysis of many examples from diverse genres. And much as heterophony combines musical sameness and multiplicity, as we move forward with this research, its content calls for connected but distinct research methods, including case studies that reveal similarities and differences among the world's musics.

Bibliography

Adler, Guido. 1909. "Über Heterophonie." *Jahrbuch der Musikbibliothek Peters* 15: 17–27.

Bregman, Albert S. 1990. *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press.

Campbell, Edward. 2013. *Music after Deleuze*. Bloomsbury.

Chow-Morris, Kim. 2010. "Going with the Flow: Embracing the 'Tao' of China's 'Jiangnan Sizhu.'" *Asian Music* 41 (2): 59–87.

Coleman, Kwami. 2021. "*Free Jazz* and the 'New Thing': Aesthetics, Identity, and Texture, 1960–1966." *Journal of Musicology* 38 (3): 261–95.

Cuthbert, Michael Scott, and Christopher Ariza. 2010. "Music21: A Toolkit for Computer-Aided Musicology and Symbolic Music Data." In *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, edited by J. Stephen Downie and Remco C. Veltkamp. International Society for Music Information Retrieval.

De Souza, Jonathan. 2019. "Texture." In *The Oxford Handbook of Critical Concepts in Music Theory*, edited by Alexander Rehding and Steven Rings. Oxford University Press.

De Souza, Jonathan, Calvin Dvorsky, and Orko Oyon. Forthcoming. "Texture and Sonata Form in Classical String Quartets: A Corpus Study." *Empirical Musicology Review*.

De Souza, Jonathan, Adam Roy, and Andrew Goldman. 2020. "Classical Rondos and Sonatas as Stylistic Categories." *Music Perception* 37 (5): 373–91.

Deutsch, Diana. 2019. *Musical Illusions and Phantom Words: How Music and Speech Unlock Mysteries of the Brain*. Oxford University Press.

Duane, Ben. 2013. "Auditory Streaming Cues in Eighteenth- and Early Nineteenth-Century String Quartets: A Corpus-Based Study." *Music Perception* 31 (1): 46–58.

Huron, David. 1989. "Characterizing Musical Textures." In *Proceedings of the 1989 International Computer Music Conference*. San Francisco: Computer Music Association.

Huron, David. 2016. *Voice Leading: The Science Behind a Musical Art*. Cambridge, MA: MIT Press.

Koffka, Kurt. 1922. "Perception: An Introduction to the *Gestalt-Theorie*." *Psychological Bulletin* 19 (10): 531–85.

Meyer, Leonard B. 1956. *Emotion and Meaning in Music*. Chicago: University of Chicago Press.

Napier, John. 2006. "A 'Failed' Unison or Conscious Differentiation: The Notion of 'Heterophony' in North Indian Vocal Performance." *International Review of the Aesthetics and Sociology of Music* 37 (1): 85–108.

Poon, Matthew, and Michael Schutz. 2015. "Cueing Musical Emotions: An Empirical Analysis of 24-Piece Sets by Bach and Chopin Documents Parallels with Emotional Speech." *Frontiers in Psychology* 6. http://dx.doi.org/10.3389/fpsyg.2015.01419

Thrasher, Alan R. 1985. "The Melodic Structure of Jiangnan Sizhu." *Ethnomusicology* 29 (2): 237–63.

Wilson, Olly. 1992. "The Heterogeneous Sound Ideal in African-American Music." In *New Perspectives on Music: Essays in Honor of Eileen Southern*, edited by Josephine Wright and Samuel A. Floyd, Jr. Harmonie Park Press.

Witzleben, J. Lawrence. 1995. *"Silk and Bamboo" Music in Shanghai: The Jiangnan Sizhu Instrumental Ensemble Tradition*. World Musics. Kent State University Press.

Yu Wang, Anna. 2023. "What Is the Difference Between Polyphony and Heterophony?: Music Theory Classifications as Instruments of Social

Class-Making." Paper presented at the Annual Meeting of the Society for Music Theory. Denver, CO.
Wong, Chuen-Fung, ed. 2024. *Silk and Bamboo: Instrumental Chamber Music from the Chinese South*. A-R Editions.

ABSTRACT

Heterophony—where multiple parts independently develop the same material—can challenge theories of texture. Psychological research on auditory streaming supports a theoretical model of texture created by David Huron. His texture space has two dimensions: onset synchrony and semblant motion. Heterophony would theoretically combine low onset synchrony with high semblant motion. However, I argue that some analytical methods cannot detect this combination and propose a new, less biased technique. I use this model to analyze performances from the Chinese *Jiangnan sizhu* tradition, where players improvise embellishments on a mother tune. Ultimately, I call for further theory and analysis of musical texture across diverse genres.

Keywords: Texture; Heterophony; Music Analysis; Auditory Streaming; *Jiangnan Sizhu*

RÉSUMÉ

L'hétérophonie—où plusieurs parts développent indépendamment le même matériau—peut remettre en question les théories de la texture musicale. La recherche psychologique sur l'organisation perceptive auditive soutient un modèle théorique de la texture, créé par David Huron. Son espace de la texture a deux dimensions : synchronisation des débuts et mouvement similaire. L'hétérophonie combinerait peu de synchronisation avec beaucoup de mouvement similaire. Cependant, je soutiens que quelques méthodes analytiques ne peuvent pas détecter cette combinaison et propose une nouvelle technique moins biaisée. J'utilise ce modèle pour analyser des performances de la tradition chinoise *Jiangnan sizhu*, où les instrumentistes improvisent collectivement des ornements sur une mélodie mère. Finalement, les théories de la texture musicale nécessiteront des analyses plus approfondies de divers genres.

Mots-clés: Texture musicale; Hétérophonie; Analyse musicale; Organisation perceptive auditive; *Jiangnan Sizhu*

BIOGRAPHY

Jonathan De Souza is an Associate Professor in the Don Wright Faculty of Music, an Associate Member of the Western Centre for Brain and Mind, and a Core Member of the Centre for Theory and Criticism at the University of Western Ontario. He is the author of *Music at Hand: Instruments, Bodies, and Cognition*, which received an Emerging Scholar Award from the Society for Music Theory, and a co-editor of *The Oxford Handbook of the Phenomenology of Music*.