Intersections
Canadian Journal of Music
Revue canadienne de musique

Understanding Microtuning Through Recording Analysis: The Transcription and Analysis of Microtuning in Five Recordings from a Variety of Musical Traditions

Taylor Brook

Volume 41, numéro 2, 2021

URI : https://id.erudit.org/iderudit/1117692ar
DOI : https://doi.org/10.7202/1117692ar

Aller au sommaire du numéro

Éditeur(s)

Canadian University Music Society / Société de musique des universités canadiennes

ISSN

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

Citer cet article

Brook, T. (2021). Understanding Microtuning Through Recording Analysis: The Transcription and Analysis of Microtuning in Five Recordings from a Variety of Musical Traditions. *Intersections*, *41*(2), 89–113. https://doi.org/10.7202/1117692ar

Résumé de l'article

Cet article présente cinq analyses assistées par ordinateur d'enregistrements musicaux, en mettant l'accent sur le microaccord, décrivant les détails de la hauteur au niveau inférieur au demi-ton. Le microaccord englobe l'intonation, les systèmes d'accord, l'inflexion de la hauteur, le vibrato et le glissando. En tant que tel, l'article se concentre sur l'analyse de la musique enregistrée et examine une grande variété de traditions musicales : une chanson folklorique américaine, une complainte Ainu, une dévotion chrétienne éthiopienne, un quatuor à cordes de Beethoven et une chanson de Tom Waits. Cet article s'interroge sur la manière dont la notation musicale peut rendre compte le plus efficacement possible du microaccord dans les transcriptions et sur la manière dont l'analyse des enregistrements, plutôt que des partitions, peut enrichir la pédagogie et la pratique de la théorie, de l'interprétation et de la composition musicales.

UNDERSTANDING MICROTUNING THROUGH RECORDING ANALYSIS: THE TRANSCRIPTION AND ANALYSIS OF MICROTUNING IN FIVE RECORDINGS FROM A VARIETY OF MUSICAL TRADITIONS

Taylor Brook

A sounding note is never at a precise frequency for more than a moment. Particularly when music is performed on non-tempered instruments, staff notation is either an approximation or an idealization of pitch. A bent note is a clear example, but subtle variations in pitch are omnipresent in musical sound.

Listeners experience the expressivity of pitch at a level that notation struggles to represent: bending a note, vibrato, sliding tones, the intonation of spoken word, etc. I will label this detailed level of pitch *microtuning* to complement the related concept of *microrhythm*, which describes a similarly zoomed-in focus on time. In musical performances, microtuning contributes to expressive and affective qualities and stylistic markers. For this reason, an investigation of microtuning could be fruitful to musical analysis, connecting sound with perception and informing musicians about how to be (or how they are already being) deliberate in their use of microtuning in performance.

Conservatory and university-trained living composers are as likely to cite as inspiration for their work popular music, folk music, or non-Western traditions as they are to cite music from the written Western tradition. This ever-expanding eclecticism is reflected not only in the sound of the music but in how it is produced and valued. As a composer myself, I am no different in terms of my wide-ranging musical origins, but my formal training was focused on developing an intimate knowledge of the written Western musical tradition. All other traditions and modes of musical production, while often admired and studied to various degrees, are usually not the subject of rigorous analysis and dissection in music theory and composition classes. At conservatories and universities, Western classical music is researched and analyzed in minute detail, dismantled, dissected, theorized, and criticized rigorously. Conversely, all other musical traditions are often studied from a primarily sociological standpoint, and are rarely the subject of rigorous analysis or informed criticism by music theorists and composers. It can be argued that divorcing music from its historical and social context diminishes our understanding of it, but in a music theory or composition class it becomes a privilege (or perhaps the cost of entry), and momentarily setting these contexts aside is a privilege that is usually enjoyed exclusively by music in the notated Western tradition.

In his influential article "Music Theory and the White Racial Frame" (Ewell 2020), Philip A. Ewell calls for a greater diversity in the range of music included in music theory classes. The work presented in this article aims to help answer Ewell's call in some small way while simultaneously considering the necessity for structural change in curriculum as outlined by Dylan Robinson who states that, "Without substantive change to the structures that underpin what I characterize as 'additive' inclusion, however, these changes can in fact maintain the larger system of white supremacy within which music programs operate" (Robinson 2019).

By eschewing performance scores as a prerequisite for analysis, music from any tradition may be considered—a fact that I will embrace in this article by selecting recorded excerpts from a variety of traditions. Given that no individual has the same level of expertise across musical traditions, misunderstandings in musical meanings and values are likely to occur both in this article as well as when curriculums diversify the repertoire they analyze. With that caveat in mind, this article aims to respect, understand, and empathize with the musical traditions discussed, to treat them as worthy of study and consideration with the same spirit expressed by Michael Tenzer in the introduction to *Analytical Studies in World Music*:

> But is the metaphor of analysis-as-discovery valid for all music? Is it alright to analyze music independent of its political, geographical, or cultural distance from the analyst or reader? Shall we allow ourselves to become absorbed in music's sound, conceiving of it as if in isolation from the world? Polemics and traditions of debate surround these questions which we visit in a moment, but no clean resolution exists. The response proposed here is nonetheless that yes, it is valid to do so, and anticipation of pleasure and refinement to be had is sufficient rationalization. We are all creatures of culture and ideology, but there is a moment in analysis at which we must curtail our penchants for modernist universalism, postmodern irony, or other language-based responses in order to confront music as elementally as possible. We submit that analysis is a path to musical awareness and better musicianship. Our purpose is to make the diverse systems of musical thought under consideration available for creative musicians looking for an informed basis on which to know, assimilate, model, or borrow from world musics. (Tenzer 2006, 5)

Background

Recording analysis has diversified the subject matter of music research and enriched our understanding of how we hear, perform, and make sense of music. Recording analysis is a large and growing field, and an exhaustive summary of the many directions and subdisciplines of recording analysis is beyond the scope of this article. One area of lively research in recording analysis is the concept of microrhythm/microtiming, which describes the minuscule deviations in timing from a perfectly even beat in musical performance. The ongoing research at the Timing and Sound in Musical Microrhythm (TIME) project at the University of Oslo focuses on microrhythm. Additionally, researchers such as Haugen and Danielsen (2020), Madison and Soiros (2014), Frühauf et al. (2013), Iyer (2002), and many others have published material on microrhythm, arguing that it is an essential component of musical expression. Notable recording analysis studies focusing on pitch and tuning include Chen (2013), Six et al. (2013), and Bozkurt (2014). The Centre for the History and Analysis of Recorded Music (CHARM) at the University of London is another major resource, developing software tools for recording analysis alongside musicological research.

Recording analysis and considerations of microtuning are central to ethnomusicology. Beginning from the inception of ethnomusicology as a discipline, initially termed comparative musicology, transcription of musical performances and the collection/transcription of folk and Indigenous musics formed the basis of their research. Ethnomusicologists continue to grapple with how to represent the object of their research in notation, and have often employed technology to aid their transcriptions, from the nineteenth-century Tonometer to Seeger's melograph in the 1950s to the computer software we use today (Cooper, 2006). Computer-aided analysis of musical recordings has proliferated in recent years in projects seeking to measure and analyze musical performance in order to understand synchronization (Timmers et al, 2014), tuning (Tavani, 2021; Yang et al, 2015), and expressivity (Marcini et al, 2014). This article will synthesize computer-aided analysis of pitch with manual, by-ear transcription, extending the methodology of the studies above by comparing and verifying the computer analysis through trained human perception.

Recent research in the discipline of music theory has considered music performance through recording analysis, notably in the work of Cook (2014), Duinker (2019), Aaslid (2022), Smith (2024), and others. This article builds on this background of growing research in music theory with a specific focus on microtuning. Shet et al (2022) identify the Eurocentric role of pitch-based analysis in music theory, stating that "pitch expertise has been leveraged to promote discriminatory musical attitudes and practices across history. Likewise, pitch-based musical parameters such as harmony have historically been regarded as the most legitimate music-theoretical pursuits" (Shea et al. 2022, 1.7), echoing the arguments of Ewell (2020), Yust (2024), Deveaney (2019), and others. This article contributes to the rectification of the historically discriminatory focus on pitch by valuing the expressive role of pitch on the microtuning level and in musical performance in a way that does not privilege the forms of pitch hierarchy and patterning that conventional pitch-based analysis in music theory has been developed to analyze.

The musical recordings that I will analyze in this paper are varied, each situated in a distinct musical tradition and genre. Additionally, the recordings include a variety of musical textures: unaccompanied voice, accompanied voice, and polyphonic instrumental music. This variety in the musical corpus renders a thorough review of the literature on these traditions unworkable, but some notable research will be discussed within the case studies below to provide necessary context.

Methodology

For this article, I have transcribed recordings by combining computer analysis with transcription by ear. Both by-ear transcription and the computer-generated data are prone to misreadings and subjective interpretation, and by switching between these tools and testing one against the other, my goal is to approach how a listener might normally perceive the pitch.

For the computer analysis, I used the software Sonic Visualiser (Cannam et al., 2010) with the pYIN (probabilistic YIN) algorithm to detect frequency for monophonic music in combination with Fast Fourier Transforms (FFT) to produce spectrograms. For polyphonic music, I used the software Melodyne 5 (Celemony n.d.) with its proprietary polyphonic pitch-detection algorithm, which was then verified through manual comparisons with spectrograms showing peak-frequency bins produced by Sonic Visualizer. With Sonic Visualizer being free and open-source software that is well supported by a community of developers as well as a public SDK and several projects developing open-source plugins, it is ideal for researchers and teachers who wish to verify and reproduce the methodology used in the case studies below or to realize similar analyses. As Sonic Visualizer improves, it will likely not be necessary to use Melodyne in the near future; however, at the time of writing it was useful to expedite and verify the pitches in polyphonic textures.

While eschewing staff notation as a starting point, the methodology used to analyze the recordings does rely upon staff notation for transcription. This choice accords with the goals of diversifying the music considered by music theory and composition while building a deeper understanding of microtuning. I come to this work from the perspective of a composer who was trained in North American universities and is proposing a broadening of that training. In this context, staff notation is a practical way of conveying information about the sound that can be expanded upon to include microtuning information via sonograms, shape visualizations of pitch contour, and frequency measurements. Indeed, it will be clear in the transcriptions below that conventional notation alone cannot express microtuning.

The staff notation in the transcriptions throughout this article also includes cent deviations (hundredths of a semitone) from the notated pitches in twelve-tone equal temperament. The reasoning for this choice is an extension of that for the use of staff notation: cent deviation is a practical way of relaying microtuning information in the context of an education system and preponderance of music technology (MIDI, digital Audio workstations, synthesizers, standardization of temperament in instrument design, etc.) that uses twelve-tone equal temperament and cents nearly ubiquitously. However, the inclusion of cent deviation may falsely suggest the twelve-tone equal temperament is a neutral musical standard, whereas I use it for my transcriptions as a tool for measurement. The alternative would be to impose another tuning standard, such as just intonation, which would likely further complicate the transcriptions by imposing a less familiar tuning standard. For this reason, the transcriptions throughout this article combine staff notation with spectrograms, pYin visualizations of frequency, and cent deviations to provide the details of microtuning as legibly as possible. Furthermore, the analysis and discussion does not arbitrarily attribute meaning to deviations from twelve-tone-equal temperament but instead focuses on pitch variability, contour, and in the case of the Busch/Beethoven case study, harmonic relationships and temperament. While this solution is imperfect, I treat twelve-tone equal temperament and cents as a fine grid to measure against, but do not place value on adherence to the grid or pitch precision.

Staff notation is not the *only* tool for sonic representation, but its benefits for the purpose of analysis are undeniable. This is evidenced by recent articles from the discipline of music theory that consider music outside score-based music making: analyzing popular music (Gardner and Shea 2022), hip-hop (Maler and Komaniecki 2021), metal (Lilja 2009), improvisation (Aaslid 2022), and Japanese shō performance (Momii 2020), while relying upon staff notation to convey their analysis. Turning once again to Michael Tenzer, who argues for the usefulness of staff notation:

> Assembling musics of the world together and juxtaposing them via staff notation asserts the present and future value, Eurocentric or not, of notational literacy as a potent means of imagining, knowing, comparing, and emulating sounds and sound-structures. Computers, recording, books, and scores are a mutually enhancing quartet. (Tenzer 2003, 4)

The goal of this article is to cover a breadth of musical examples, make claims across disparate traditions, test the flexibility of the analytic methodology, and suggest future paths of inquiry. The following recordings will be analyzed in this article:

1. Almeda Riddle, "Chick-a-la-li-o" (United States, folk song, solo voice, 1959)

2. Teru Nishizama, "Chisi-Sinotcha" (Japan, Ainu song, solo voice, 1978)

3. Alèmu Aga, "Abatatchen Hoy" (Ethiopia, Christian devotional, voice and begenna, 1995)

4. Busch String Quartet, Beethoven String Quartet no. 15 in A minor, op. 132, first movement (Germany, classical string quartet, 1938)

5. Tom Waits, "That Feel" (United States, popular song, voice, guitar, bass, and drums, 1992)

This group of recordings was selected to provide contrasting musical styles, production, and traditions. Additionally, all the recordings are of either exclusively or primarily untempered instruments, allowing for dynamic microtuning and a richness of tonal shading. While most studies that realize computer-aided pitch analysis are limited to in-depth analysis of a single recording or comparisons between multiple recordings of the same work, this article purposely juxtaposes not only of different musical works but also varied styles, traditions, and musical textures that are not commonly placed alongside one another. A notable exception to this analytic trend is Danielson et al., researchers who investigated microrhythm across several musical disciplines and genres (Danielson et al., 2024); they did not, however, discuss pitch or microtuning in their study.

Case 1: Almeda Riddle, "Chick-a-la-li-o"

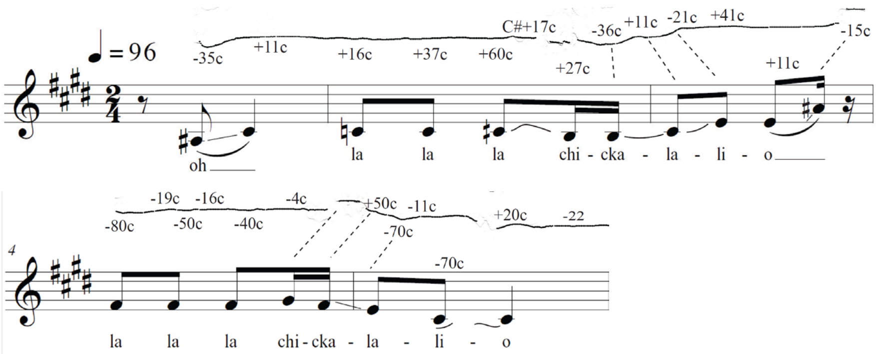
Almeda Riddle (1898–1986) was an American folk singer from Arkansas who became a prominent performer and folklorist after a recording session with musicologist Alan Lomax for the Library of Congress. In 1959, Lomax recorded Riddle singing a series of children's songs in her Arkansas home that had been passed down to Riddle orally. One of these songs was "Chick-a-la-li-o" (Riddle 1959), and this recording exemplifies the difference between performance and a standard staff transcription. "Chick-a-la-li-o" is a strophic song in C-sharp pentatonic minor that repeats a short musical phrase with different lyrics in the first half of the phrase upon on each repetition. Isolating the phrase of the song, traditional staff notation could be as shown in figure 1.

Figure 1. "Chick-a-la-li-o" phrase in standard Western notation

Adding the microtuning details of Riddle's performance with numbers indicating cent deviation from the note above the staff and a line generated in Sonic Visualizer using the pYIN pitch detection algorithm produces the transcription in figure 2.

In this recording, Riddle sings with a straight tone, unaccompanied, and displays a flexibility in microtuning. On the tonic note of C-sharp, there is a range of 84 cents flat to 20 cents sharp, or approximately a semitone, in variation. The two C-sharps that are tuned the highest occur on the "oh/o" syllable at the beginning and end of the phrase respectively, suggesting a possible relationship between phoneme timbre and tuning.

Figure 2. "Chick-a-la-li-o" opening phrase as sung by Riddle



The pYIN visualization above the staff shows both a constant wavering of pitch as well as sliding between most notes. Riddle also tends to scoop up into notes from below, a feature that stylizes and marks the American folk genre generally. Moving on to a longer transcription including the following few phrases of "Chick-a-la-li-o," in figure 3a and 3b.

The transcription in figures 3a and 3b confirms some of the previous observation regarding the opening phrase: sliding between notes is the norm, tuning variation of approximately a semitone and bending upward into notes at the beginning of words and phrases all persist. Building from the identification of the "o" syllable tending to be higher in pitch relative to the scale degree, we can expand this idea to a broader claim that the text, and how it might be spoken, influences microtuning in performance. Another stylistic tendency appears to be that on long notes and some repetitions of notes the pitch tends to drift slightly upwards. We can hear this at the end of each phrase on the final long note (the "o" of "chick-a-la-li-o").

There are moments when the contour of how one would likely speak the text is represented in the microtuning. For example, the end of the excerpt at measure 24 of the transcription, the syllable "pre" in "He's the prettiest boy" is accented with about a quartertone shift upward in pitch as compared to the syllables before and after. A speaker of this text would likely accent this syllable, and so the slight upward push of the tuning may be a musical translation of the speech accent.

I noted above that the pitch variation in the opening phrase of "Chick-a-la-li-o" was approximately a semitone, and we see this expand to approximately 1.6 semitones (-80c to +80c) when considering the first two verses transcribed in figure 3. This variation in pitch is significantly larger than what Mauch et al. (2014, p.410) identified in their study on intonation in unaccompanied singing of the Happy Birthday song. And in their computer-automated analysis of folk songs recordings from the Netherlands, Müller et al (2010) note that that "The singers often deviate significantly from the expected pitches and have serious problems with the intonation. Even worse, from a technical point of view, their voices often fluctuate by several semitones downwards or upwards across the various stanzas of the same recording" (Müller et al. 2010, 247). While the authors attribute this to the recorded songs being performed by elderly non-professional singers, the analysis of Riddle performing "Chick-a-la-li-o" provided above suggests that the malleability of intonation may be a persistent feature that contributes to musical expressivity.

Figure 3a. "Chick-a-la-li-o" transcription as sung by Riddle part 1

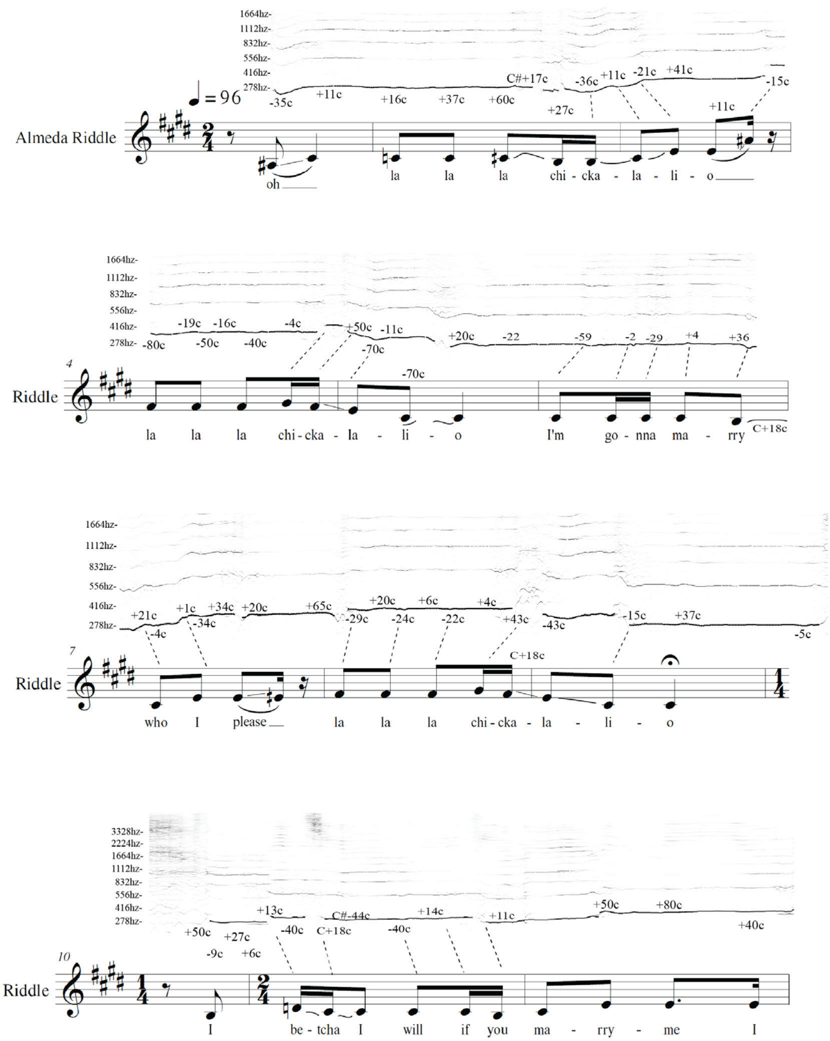
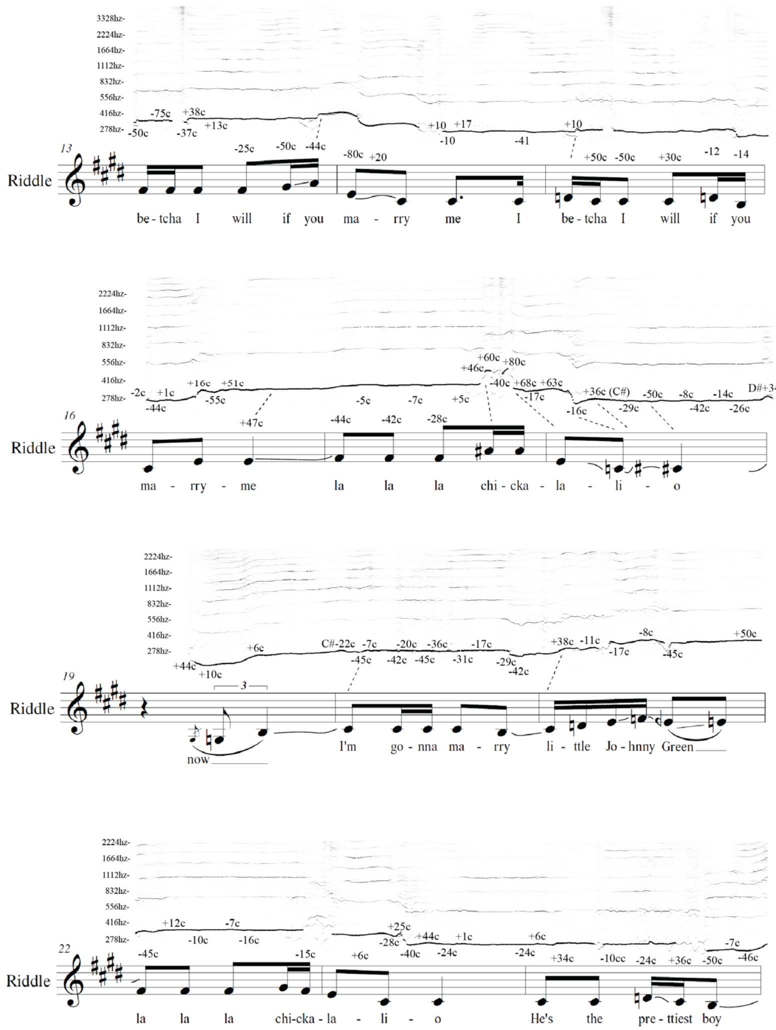


Figure 3b. "Chick-a-la-li-o" transcription as sung by Riddle part 2



Case 2: Nishijima Teru, "Chisi-Sinotcha"

The second case study is again a solo voice recording: Nishijima Teru's[1] (1896–1988) performance of the Ainu song "Chisi-Sinotcha" (crying song). The Ainu are an Indigenous people in Japan with a distinct language and cultural heritage, traditionally located on the lands surrounding the Sea of Okhotsk, an area that was not colonized by the Japanese until the latter half of the nineteenth century (Siddle 1999, 108). Sinotcha are lyric songs sung by a single person expressing their personal emotions (Tanimoto 1999, 787), and this particular sinotcha is a sorrowful love song in which a repeated melody is set to different words upon each repetition. Like most Ainu songs, the melody features an anhemitonic pentatonic scale with variations in microtuning (Tanimoto 2002, 205).

The repeated melody in Western staff notation could be simplified as shown in figure 4.

Figure 4. "Chisi-Sinotcha" (crying song) in standard Western notation

A detailed transcription that includes microtuning in the form of sonogram, pYIN algorithm, and cent deviations from equal temperament of the first three iterations of the melody as recorded by Nishijima Teru is provided in figure 5. I have transcribed the words of the song using American IPA.
The lyrics of "Chisi-Sinotcha" translate to:

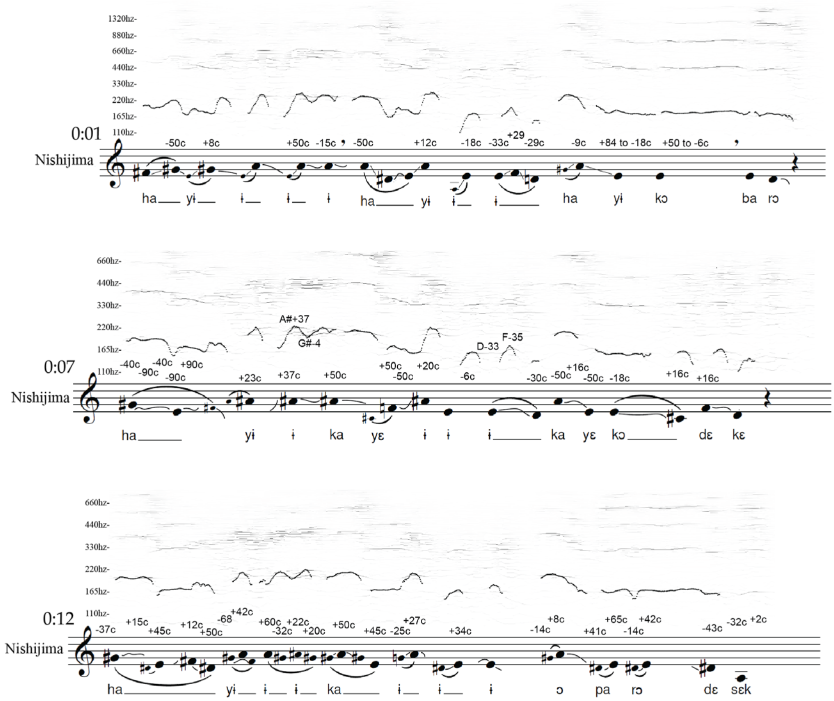
> I am so sad / ku sanpe wen na,
> embrace my mouth / ku paro kisma,
> clasp my hand / ku teke kisma,
> give me water / wakka en kore (Okitsu 2022)

Before the words are sung, each line begins with a moan (notated as hayɨ in my transcriptions). Although there is some sense of the anhemitonic scale, the tuning is fluid throughout this performance. The lines generated by the pYIN algorithm in the transcription above show this fluidity clearly as the pitch is constantly sliding. Indeed, the shape of the line is recognizably similar in each

1 Although the liner notes (Nishijima 1980) of the album that includes this recording identifies the singer as Nishizama Teru, this name is not known among Ainu denshosha (transmitters of tradition/culture) and the singer is more likely Nishijima Teru, a renowned denshosha from Biratori. This information originates from personal correspondence with Terachi Goichi (Professor at Tokyo School of Economics) and Okitsu Tsubasa (Ainu denshosha). Information on Nishijima Teru may be found here: https://ainugo.nam.go.jp/siror/denshosha/?denshosha_id=D010

iteration of the melody and easier to keep track of visually than with grace notes and glissando lines on the staff.

Figure 5. "Chisi-Sinotcha" (crying song) opening two phrases as sung by Nishijima Teru

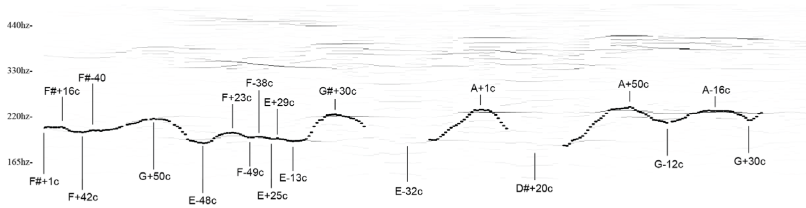


"Chisi-Sinotcha" is a crying song. The fluidity of pitch, along with the dynamic variations and vocal timbre, is used to express the sorrowful feelings of the singer. The melody begins in the upper register, wavering in a constrained range before falling down dejectedly to the low register toward the end of each phrase. The notes D and A are emphasized in this recording, although the tuning tolerance is greater than a semitone in both cases. More central than these two discrete pitches of D and A is the recognizable contour of the melody and how variations in pitch during the trill-like passages evoke weeping.

Below is a detailed transcription of the first two seconds of the vocal melody with the pYIN fundamental frequency in bold and sonogram bins from Sonic Visualizer shown with the addition of pitch annotations:

Figure 6 provides a zoomed-in transcription of the first "hayɨ" moan that begins the recording. This figure represents the gesture that defines the vocal performance: bending up and then returning back down with a slight accent at the peak. This gesture can be heard as a slow, measured vibrato that evokes the sound of weeping and is reflected both within this single moan as well as the larger contour of the repeated phrase throughout the performance.

Figure 6. "Chisi-Sinotcha" (crying song) opening moan gesture as sung by Nishijima Teru



Case 3: Alèmu Aga, Abatatchen Hoy

The third case study is a 1994 recording by Alèmu Aga performing the Amharic devotional "Abatatchen Hoy" (Aga 1994), which translates from Amharic as "Our Father's Prayer." The Amharic people are Indigenous to the northwest highlands of Ethiopia and are primarily Orthodox Christians. This recording is a devotional in the Ethiopian Orthodox Tewahedo Church, an African form of Christianity that predates the European colonization of the continent, originating in the Kingdom of Aksum in the fourth century (Moore 1936, 271–284).

While the first two recordings discussed are solo vocal performances, this recording is a voice and begena duo. The begena is an Ethiopian lyre with ten strings, as shown in figure 7.

Figure 7. Begena, instrument of the Ethiopian Orthodox Tewahedo Church. Credit: Wikipedia.

Of the ten strings of the begena, only five are plucked while the other five catch the fingers after a through-stroke pluck and do not themselves resonate. The begena is in a low register, and in this recording it is tuned between C1 and A2 (assuming C4 as middle C), and only the open strings are played. The begena strings are somewhat slack compared to orchestral stringed instruments or guitars, and as a result of this relative slackness the pitch wavers, especially at the moment of attack. This is most notable on the lowest string, which sounds approximately 50 cents (1/2 semitone) higher on attack compared to during the sustain.

Figure 8. Begena tuning in Aga's "Abatatchen Hoy"

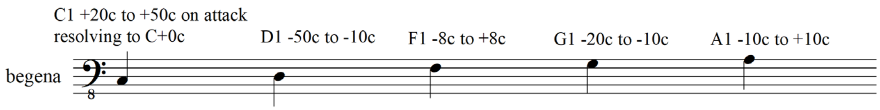


Figure 8 provides the tuning of the begena strings played in this recording of "Abatatchen Hoy," including the pitch range in cents as the string resonates. The upper three strings have a much smaller tuning deviation than the lower two strings as the greater tension of these strings produces the higher pitches even though the length of all five strings are identical. While the pentatonic scale is clearly perceivable as these strings are plucked, there is a wavering in pitch and a continuous warping of the tonal center because of the unsettled pitch of the lower strings. Alèmu Aga alludes to this phenomenon in a 2018 interview:

> Well, the Harp of King David [begena] is used for prayer. It's as well used to praise God. Quite the reverse, the other Ethiopian traditional instruments are mostly used during occasions of dancing and singing. But we do not take advantage of the Harp for other purposes as the instrument is very different.
>
> The instrument takes people far deep in thought apart from creating an electrifying effect in people from head to foot. At the time when the instrument was on the brink of extinction, it was tried to orchestrate the Harp of King David with other traditional instruments targeting at revitalizing it. But the deep sound it produces and the vibration it releases could not be harmonized with other musical instruments. As such the orchestration effort did not get acceptance. (Alèmu Aga, Mulat 2018)

Alèmu Aga's recording of "Abatatchen Hoy" begins with a begena introduction (known as a *Derdera*), which gradually introduces the melody with added ornamentation before finally presenting the melody plainly just before the first entrance of the voice. A simplified transcription of the repeating melody could be written as shown in figure 9.

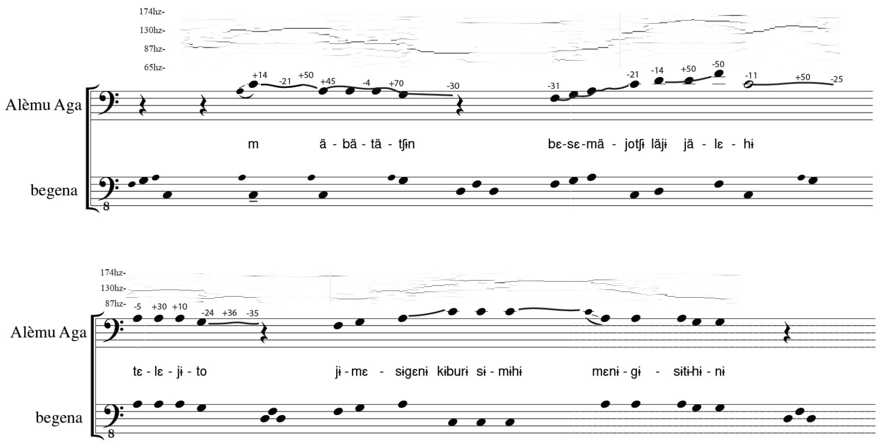
A transcription of the begena introduction and the first two iterations of the vocal melody are shown in figure 10 with the words written in American IPA.

Figure 9. Simplified melody of "Abatatchen Hoy"

Figure 10. Transcription of opening section of Aga's "Abatatchen Hoy"

The wavering of pitch in the low C that undergirds the perception of a pitch center is emphasized in the begena introduction as well as being played and sung as the final note of each phrase. Thus, musical phrases begin and end with the microtuning feature of the lowest note settling gradually downward by about 50 cents. Turning to the microtuning of the vocal line, the first instance of the sung melody along with the sonogram, cent deviations, and a manually drawn line showing the sliding of the pitch is shown in the transcription in figure 11.

Figure 11. Vocal entrance in Aga's "Abatatchen Hoy"



The voice maintains a soft dynamic without vibrato throughout the recording of "Abatatchen Hoy." The microtuning is expressive, with subtle bends that interact with the ever-present harmonic field produced by the begena. The vocal pitches have approximately 50 cents (a quartertone) of variance, while the constant bending from one note to the next challenges the perception of discrete notes and the clarity of the mode. The notes are always sliding from one to the next with the exception of the final notes of each sub-phrase, which alternate between a sustained G and C. These terminal notes tend to bend slightly lower by approximately 50 cents as they are held, imitating the plucked strings of the begena settling in pitch as they resonate: a higher pitch on the attack and settling slightly downward as they sustain and decay. The downward bend on held notes in the vocal line may originate from the settling of the open strings of the begena, but this feature also stylizes the singing in the same way that Riddle bent long notes upwards at the end of each phrase in her performance of "Chick-a-la-li-o," as discussed in the first case study. These stylizations have an expressive effect, however subjectively interpreted and described, and the begena and voice consistently bending slightly downward could be felt by the listener as calming and relaxing, reflecting the social context of this music as prayer.

Matching the variation in pitch, the tempo of this recording also varies greatly, moving dynamically between about 40 and 100 beats per minute throughout the performance. A thorough discussion of the temporal aspects, as well as the musical relationship with the devotional text, are beyond the scope of this article; instead, I note here only that there is likely a meaningful relationship between microtuning and these other aspects of this music.

Indeed, the settling downward of pitch of the lowest string of the begena is a gradual process that can only be heard clearly in the context of a slow and spacious musical texture.

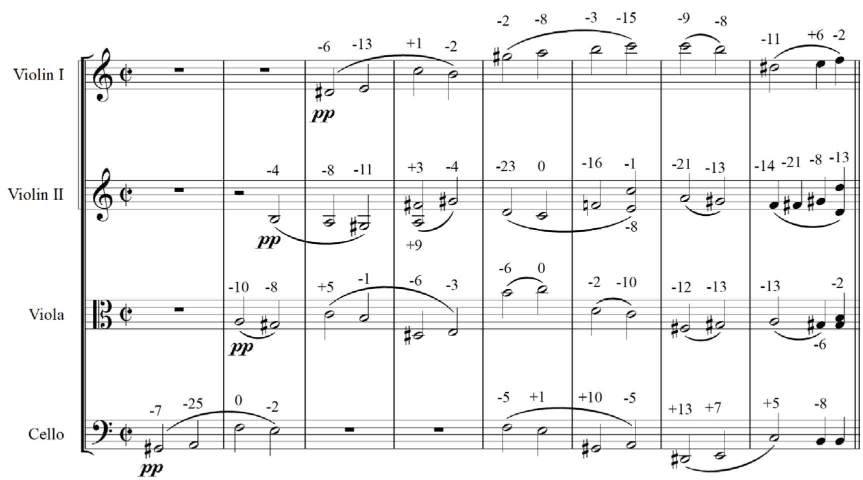
Case 4: Busch String Quartet, Beethoven String Quartet no. 15, first movement

The fourth case study considers the 1938 Busch String Quartet recording of the slow introduction of the first movement of Beethoven's String Quartet no. 15 in A minor, op. 132. This recording is notable among the case studies in two ways: a written score was used to produce the performance and the musical texture is polyphonic. It is undeniable that the music of Beethoven is already well represented in music theory literature and pedagogy. However, in seeking to open up new lines of inquiry and place score-based music making in a global context, it is essential that the methodology found in this article also be applied to recordings of music that are already situated within the canon of music theory analysis so that we may discover how microtuning operates in such recordings. Polyphony brings some challenges to the measurement of microtuning using algorithms. Melodyne parses the polyphonic texture into its constituent notes, but the algorithm is proprietary and so the analysis, while it seems to be accurate when tested against the ear, is a black-box function. FFT analysis using sonic visualizer enables verification and sometimes more detail by identifying peak frequency bins; however, when an overtone from one instrument interacts with the fundamental from another or when two instruments play in unison, it becomes challenging to disaggregate the lines. For the purposes of this paper and its broad goals, the Melodyne and FFT-based analysis is relied upon and provides reasonable accuracy when tested against each other and the ear, despite these limitations.

The slow polyphonic texture of the recording provides an opportunity to analyze the harmonic dimension of the microtuning in this recording. In 2012, Papiotis et al. noted the absence of literature "involving a computational approach to intonation adaptations" (Papiotis et al 2012, p.2) in string quartet and ensemble performance; they also noted in their study, which focused on temporal synchronization, that in the context of the string quartet, "after being provided with the score annotations and practicing, the musicians tended to move away from an equal temperament tuning system" (Papiotis et al. 2012, p.4). Considering the common-practice tonal syntax of this music, we can identify triadic structures and measure the relationship between chord tones and compare them with idealized tuning systems beyond 12-tone equal temperament such as just intonation and Pythagorean, meantone, and well temperaments. From the melodic perspective, the slow introduction features a G-sharp–A–F–E cell, which is transformed, repeated, and transposed. Figure 12 provides a reduction of this section with averaged cent deviations for each note.

The first triad of the piece (second chord of measure 2) is a root-position dominant triad comprised from the bottom up as E, G-sharp, and B. The perfect fifth between the E and B are at a distance of 698 cents, which is 2 cents smaller than equal temperament and 4 cents smaller than a just or Pythagorean fifth. The relationship between the E and G-sharp is 394 cents, which is 6 cents smaller than equal temperament and 8 cents larger than a just major third. However, this chord conforms perfectly with 1/6-comma meantone temperament, which has the fifth at 698 cents and the major third at 394 cents. 1/6th-tone meantone was a common temperament in Beethoven's time, as well as the temperament advocated for by Leopold and Wolfgang Amadeus Mozart (Chestnut 1977). While it would be unrealistic to expect untempered instruments to perform perfectly in an idealized temperament, it is striking how precisely this chord adheres to meantone temperament and the historical appropriateness of the tuning. Not all chords in this sequence follow 1/6th-tone meantone temperament so closely, but there is a tendency toward meantone tunings, via slightly contracting perfect fifths and thirds, thus approaching just intonation.

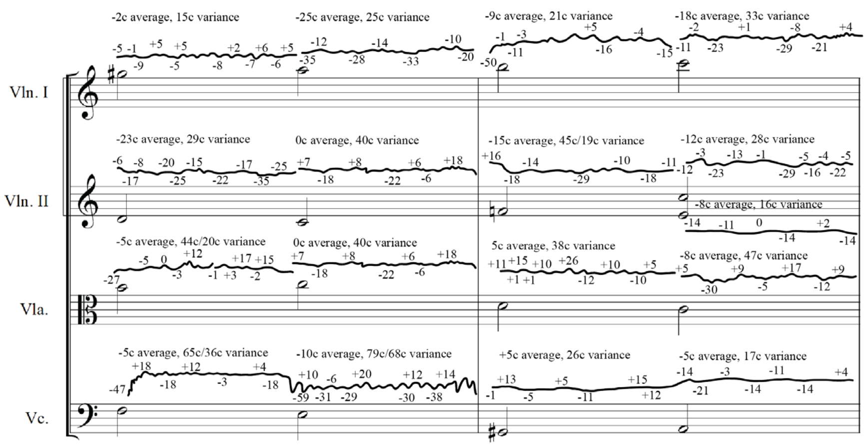
Figure 12. Measures 1–8, Beethoven String Quartet no. 15 in A minor, op. 132, first movement, slow introduction, performed by the Busch String Quartet



Focusing in on measures 5–6 (figure 13), we can track some of the microtuning within held notes. Along with the cent deviations and a manually drawn line that visualizes pitch, there is also the average pitch and pitch variance for each note. When there is a slide between notes, the slide is omitted from the pitch average and variance. This analysis was derived from peak bin frequencies in the sonogram created by Sonic Visualizer.

Figure 13 shows the variance in pitch of the vibrato, which is around 30–40 cents throughout this passage. This level of variance within a note is much more constrained than any of the vocal lines discussed in the recordings of the three previous case studies. This constraint may be related to the performance practice of Beethoven's music, rooted in an effort to produce a performance that adheres to the written score as much as possible. Another explanation may be that the perception of harmony in the polyphonic texture could be blurred if vibrato is too wide, although Geringer et al. found that the use of vibrato increased the perception of harmonic relationships being correctly tuned (Geringer et al. 2015, 680). We also can discern that the vibrato is not perfectly even in the Busch performance, but usually increases in the middle of each note, with the exception of the second note in the cello, where the vibrato is used to accent the beginning of the note. This accented cello note is the dominant scale degree of E, played with much faster and somewhat wider vibrato than the other notes, emphasizing the tension of the dominant.

Figure 13. Measures 5–6 of Beethoven String Quartet no. 15, first movement, performed by the Busch String Quartet



Case 5: Tom Waits, "That Feel"

The fifth and final recording contrasts the others in instrumentation, style, and genre. "That Feel" by Tom Waits is a song from his eleventh studio album, *Bone Machine* (1992). This recording includes drums (Waits), upright bass (Larry Taylor), guitars (Waits, Keith Richards, David Phelps, Waddel Wachtel), and voices (Waits, Richards) and was co-written by Tom Waits and Keith Richards. The lyrics of "That Feel" relate to themes of addiction, loss, and death. Speaking about the album as a whole, Waits states that "some of the songs deal with dying and . . . the idea that we're all hurling through space and eventually the earth will open up and swallow us all" (Waits 1992b).

The music has a drunken lilt and Waits's vocals are particularly varied in terms of timing, pitch, and timbre. This lilt and instability of the voice is a purposeful quality of the song that is part of the singer's practice of embodying the lyrics. Referring to the process of writing music Waits states:

> In the past couple of years, I'm more physical about the music. I used to sit at a piano, which was almost like, it hurts your back and then you have to put your head down on the keyboard after a while . . . it's just . . . it's crazy.

> I think most people that play the piano want to see that piano thrown off a building and they want to be there when it hits the street . . . writing with your fists is better sometimes . . . and that music is a physical thing, you can't see it and in it there's a spirit. (Waits 1992b)

This quote suggests a rejection of creating music as a combination of pitches and rhythms and instead building a song in what Waits calls a more "physical" way, emphasizing the dimensions of sound that cannot be captured by the musical score (represented in the quote above by the piano).

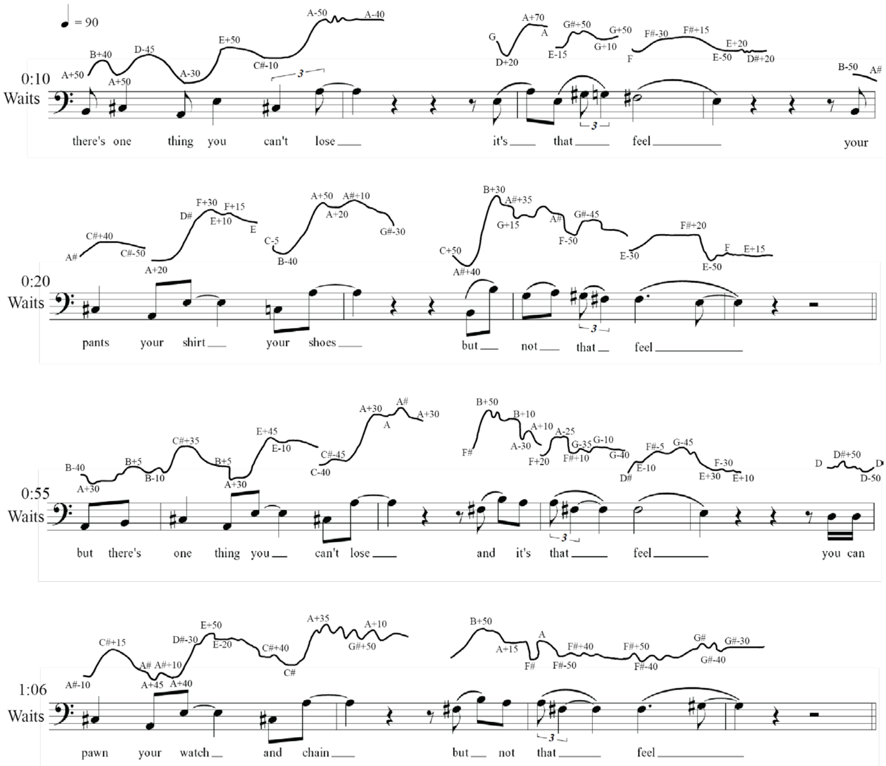
Transcribing Waits's singing in this recording was challenging because the timbre of his voice is particularly rich, variable, and inharmonic. Heidemann (2016) identifies the notable use of growl in Waits' vocal timbe and Thomas (2016) also notes the challenge of unreliable pitch analysis on Waits' voice as well as the preponderance of microtonal pitch variations. Even after applying source separation to isolate Waits' vocals, computer analysis of pitch, both with Sonic Visualizer and Melodyne, did not provide clear results that aligned with my perception of pitch. As a consequence, the following transcription relies more on my ear and less on computer analysis than in the previous case studies. Figure 14 is a transcription of the first two verses of the song.

Figure 14 is a transcription of the same melody sung four times. With the exception of the final note of the last melody shown in the transcription and the pickup notes, a traditional score for this song would likely notate this melody in the exact same way each time it appears. Waits's performance, in contrast, delivers each iteration of the melody in a unique way, varying the microtuning and microtiming while maintaining the overall shape of the melody. This may be related to what Waits means when he calls music a "physical thing," as conventional notation does not capture many aspects of the music—not only in terms of the microtuning but also the subtleties of the timing and timbre.

"That Feel" features several guitars, plucked upright bass, and drums performing a sparse accompaniment to the voice. There is a slow and even pulse that is marked by the drums and bass at the whole-note level, while a guitar pick pattern creates a 4/4 metrical feel. This steady rhythmic context allows Waits to perform the vocals with a great deal of rhythmic ebb and flow in relation to the steady accompaniment. Some of rhythmic context this is captured in the transcription above (figure 14) in the shifting rhythmic placement of the final note of each phrase as well as in other moments within the phrases. The accompaniment also provides a clear tonal grounding with the instruments playing in tune with 12-tone equal temperament and with three-note harmonies of a tonic, subdominant, dominant chord progression that repeats. While the tonic and subdominant are conventional triads, the dominant harmony is quartal, including a suspended fourth that does not resolve. This quartal harmony on the dominant undercuts the directionality of the harmonic progression as the tonic note is present within the dominant chord and we do not hear a "leading tone" resolution as the harmony returns to the tonic chord. The result is a progression with minimal changes in chord tones as the three harmonies cycle through, creating a feeling of harmonic stasis. Like the rhythmic/

metrical context, this harmonic context undergirds Waits's vocal performance, which is rich and varied in terms of the microtuning and timbral variation.

Figure 14. First four occurrences of "That Feel" verse melody as sung by Tom Waits



Conclusion

Harry Partch likens microtuning to shades of color when he writes that "there are no shades of C-sharp, no shades of red, for [the composer]. The one shade that his gods will allow him to use is before him. He is taught that that is enough; it is good, traditional, and proper, and he feels a vague sense of immorality in even wondering about those possible bastard C-sharps" (1940, 159–160). Indeed, musicians have always used these "shades" in performance, but the theorist and composer struggle to account for them. Performers of non-tempered instruments and voice, whether they are creating music from a score, improvising, or working in an oral tradition, all develop techniques related to microtuning in their training and must reckon with microtuning whenever they perform. Mira Benjamin's doctoral thesis explores microtuning from a performer's perspective and shows how developing a knowledge in this realm enriches their practice (Benjamin 2019) and cellist Pablo Cassals believed that most of a performer's expressivity is derived from intonation (O'Malley 1983/2017).

A refinement of the methodology used in the five case studies presented in this article could be applied in music theory classrooms, which would allow for the meaningful inclusion in analysis of diverse musics. Indeed, building from conventional notation and iterating on a foundation of Western music theory while moving beyond traditional limitations situates the approach as one that could be integrated into Western conservatory-style pedagogy. Conventional notation already captures some aspects of microtuning, such as portamento, glissando, or expressive indications that suggest how to perform vibrato. Notation has been expanded by many composers over the past hundred years to capture microtuning in more detail, such as microtonal accidentals, parametric notation, or the many forms of graphical notation and text scores—although these notational expansions rarely become standardized.

The transcriptions created for this article, while hopefully providing clarity for analysis, are more descriptive than prescriptive, and there are myriad ways to approach microtuning as a composer. To provide one particularly elegant and economical example, below is Oliveros's *The Tuning Meditation* (1971) from *Four Meditations for Orchestra* (1996):

> Begin by playing a pitch that you hear in your imagination. After contributing your pitch, listen for another player's pitch and tune in unison to the pitch as exactly as possible. Listen again and play a pitch that no one else is playing. The duration of pitches is determined by the duration of a comfortable breath or bow. (Oliveros 1996)

The case studies in this article suggest many directions that microtuning analysis could be taken in. The first case study (Riddle) could be expanded to a broader study of how microtuning relates to stylistic markers as well as the relationship between the intonation of spoken word and singing. The second case study (Nishijima) could be expanded to draw connections between microtuning with emotional expression. The third case study (Aga) hints at a relationship between vocal timbre with microtuning and the details of pitch-matching between voice and instrument. The fourth case study (Busch/Beethoven) begins to explore the dimension of harmony as well as the score-to-stage transformation of a notated musical work. The fifth and final case study (Waits), touches on many of the themes of the earlier studies, but also suggests potential for research on the perceptual limits of interpreting a musical phrase as an iteration or a variation. Furthermore, all of the possible findings of the case studies could be refined and tested in studies that compare other recordings in the same tradition and genre or even expanded and contrasted through comparison with the music of unrelated traditions.

Microtuning analysis also presents many possible future directions for inquiry: How might the analysis of microtuning be systemized? What role could microtuning analysis have in a musical education, particularly in music theory and composition classes? How might microtuning analysis contribute to improved humanization functions in digital audio workstations (DAWs) and music notation software? How is microtuning treated in the context of music performance training? What are some of the implications of tuning tolerance for composers who write precisely tuned microtonal music? With all of these potential directions, this conclusion offers many more questions than answers. Nevertheless, this methodology and focus on microtuning proposes one path toward enriching musical pedagogy by both diversifying the musical traditions included in the discourse as well as focusing on an often-overlooked aspect of musical performance and analysis.

Bibliography

Aaslid, Vilde. 2022. "Sociable Musicopoetics in Vijay Iyer and Mike Ladd's *In What Language*." *Music Theory Online* 28 (3).

Aga, Alèmu. 1994. Ethiopiques, Vol. 11: Bèguèna (The Harp of King David). 2002 CD reissue by Buda Musique 82232–2.

Benjamin, Mira. 2019. "Thick Relationality: Microtonality and the Technique of Intonation in 21st Century String Performance." PhD diss., University of Huddersfield.

Bozkurt, Barış, Ruhi Ayangil, and André Holzapfel. 2014. "Computational Analysis of Turkish Makam Music: Review of State-of-the-Art and Challenges." *Journal of New Music Research* 43(1): 3–23.

Cannam, C., C. Landone, and M. Sandler. 2010. "Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files." Proceedings of the ACM Multimedia 2010 International Conference, Firenze, Italy, 1467–1468.

Celemony. (n.d.). Melodyne 5. Computer Software.

Chen, Kainan. 2013. "Characterization of Pitch Intonation of Beijing Opera." Master's thesis, Music Technology Group, Universitat Pompeu Fabra, Barcelona.

Chesnut, J.H. 1977. "Mozart's Teaching of Intonation." *Journal of the American Musicological Society* 30 (2): 254–271.

Cook, Nicholas. 2014. *Beyond the Score: Music as Performance*. New York: Oxford University Press.

Cooper, D., and I. Sapiro. 2006. "Ethnomusicology in the Laboratory: From the Tonometer to the Digital Melograph." *Ethnomusicology Forum* 15 (2): 301–313.

Danielsen, Anne, Ragnhild Brøvig, Kjetil Klette Bøhler, Guilherme Schmidt Câmara, Mari Romarheim Haugen, Eirik Jacobsen, Mats S Johansson, et al. 2024. "There's More to Timing than Time: Investigating Musical Microrhythm Across Disciplines and Cultures." *Music Perception* 41 (3).

Devaney, Johanna. 2019. "Eugenics and Musical Talent: Exploring Carl Seashore's Work on Talent Testing and Performance." *American Music Review* 48 (2).

Duinker, Ben. 2019. "Plateau Loops and Hybrid Tonics in Recent Pop Music." *Music Theory* Online 25 (4).

Ewell, Philip. 2020. "Music Theory and the White Racial Frame." *Music Theory Online* 26 (2).

Frühauf, Jan, Reinhard Kopiez, and Friedrich Platz. 2013. "Music on the Timing Grid: The Influence of Micro-Timing on the Perceived Groove Quality of a Simple Drum Pattern Performance." *Music Perception* 17: 246–260.

Gardner, Samuel, and Nicholas J. Shea. 2022. "Gestural Perspectives on Popular-Music Performance." *Music Theory Online* 28 (3).

Geringer, John M, Rebecca B MacLeod, Clifford K Madsen, and Jessica Napoles. 2015. "Perception of Melodic Intonation in Performances with and without Vibrato." *Psychology of Music* 43 (5): .

Haugen, M. R., & Danielsen, A.. 2020. "Effect of tempo on relative note durations in a performed samba groove." *Journal of New Music Research* 49 (4): 349–361.

Heidemann, Kate. 2016. "A System for Describing Vocal Timbre in Popular Song." *Music Theory Online* 22 (1).

Iyer, Vijay. 2002. "Embodied Mind, Situated Cognition, and Expressive Micro-Timing in African-American Music." *Music Perception* 19 (3): 387–414.

Lilja, Esa. 2009. *Theory and Analysis of Classic Heavy Metal Harmony*. The Finnish Music Library Association.

Madison, Guy, and George Sioros. 2014. "What Musicians Do to Induce the Sensation of Groove in Simple and Complex Melodies, and How Listeners Perceive It." *Frontiers in Psychology* 5: 894.

Maler, Anabel, and Robert Komaniecki. 2021. "Rhythmic Techniques in Deaf Hip Hop." *Music Theory Online* 27 (1).

Marchini, Marco, Rafael Ramirez, Panos Papiotis, and Esteban Maestre.2017 "The Sense of Ensemble: A Machine Learning Approach to Expressive Performance Modelling in String Quartets." *Journal of New Music Research* 43 (3): 303–17.

Mauch, Matthias, Klaus Frieler, and Simon Dixon. 2014. "Intonation in Unaccompanied Singing: Accuracy, Drift, and a Model of Reference Pitch Memory." *The Journal of the Acoustical Society of America* 136 (1).

Momii, Toru. 2020. "A Transformational Approach to Gesture in Shō Performance." *Music Theory Online* 26 (4).

Moore, Dale H. 1936. "Christianity in Ethiopia." *Church History* 5 (3): 271–284.

Müller, Meinard, Peter Grosche, and Frans Wiering. 2010. "Automated Analysis of Performance Variations in Folk Song Recordings." In *Proceedings of the International Conference on Multimedia Information Retrieval*. ACM.

Mulat, Addisalem. 2018. "Ethiopia: Gifted in Engrossing Hearts with Harp (Interview with Alèmu Aga)." *The Ethiopian Herald*, AllAfrica Global Media.

Nishijima, Teru. 1980. "Chisi-Sinotcha." On Japan: Ainu Songs. UNESCO Collection of Traditional Music.

Oliveros, Pauline. 1996. Four Meditations for Orchestra. Deep Listening Publications – ASCAP.

O'Malley, P.H. (1983/2017). Pablo Casals on expressive intonation. *The Strad*. https://www.thestrad.com/cellist-pablo-casals-on-expressive-intonation/1434.article.

Papiotis, Panagiotis, Marco Marchini, and Esteban Maestre Gómez. 2012. "Computational analysis of solo versus ensemble performance in string quartets: intonation and dynamics." In Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music. Thessaloniki, Greece, July 23–28, 2012.

Partch, Harry. 1991. *Bitter Music*. Chicago: University of Illinois Press.

Riddle, Almeda. 1959. "Chick-a-la-li-o." Recordist, Alan Lomax. Lomax Digital Archive.

Robinson, Dylan. 2019. "To All Who Should Be Concerned." *Intersections* 39 (1): 137–144.

Shea, Nicholas, Lindsey Reymore, Christopher Wm. White, Ben Duinker, Leigh VanHandel, Matthew Zeller, and Nicole Biamonte. 2024. "Diversity in Music Corpus Studies." *Music Theory Online* 30 (1).

Siddle, Richard. 1999. "From Assimilation to Indigenous Rights: Ainu Resistance since 1869." In *Ainu: Spirit of a Northern People*, edited by William Fitzhugh and C.O. Dubrueil. University of Washington Press.

Six, Joren, Olmo Cornelis, and Marc Leman. 2013. "Tarsos, a Modular Platform for Precise Pitch Analysis of Western and Non-Western Music." *Journal of New Music Research* 42 (2): 113–129.

Smith, Eron. 2024. "Prosodic Dissonance." *Music Theory Online* 30 (2).

Tanimoto, Kazuyuki. 1999. "To Live Is to Sing." In *Ainu: Spirit of a Northern People*, edited by William Fitzhugh and C.O. Dubrueil. University of Washington Press.

Tanimoto, Kazuyuki. 2002. "Music of the Ainu, Nivkhi, and Uilta." In *The Garland Encyclopedia of World Music*, edited by R.C. Provine, Y. Tokumaru, and J.L. Witzleben. Routledge.

Tavani, Nicholas John, II. 2021. "Quantifying Dynamic Pitch Adjustment Decision Structures in String Quartet Performance." Order No. 28540623, University of Maryland.

Tenzer, Michael. 2006. "Introduction: Analysis, Categorization, and Theory of Musics of the World." In *Analytical Studies in World Music: Analytical Studies in World Music*. Oxford University Press.

Thomas, Margaret E. 2016. "Text and Temporality: Toward an Understanding of Rhythmic Irregularities in the Music of Tom Waits." *Music Theory Online* 22 (4).

Timmers, Renee, Satoshi Endo, Adrian Bradbury, and Alan M Wing. 2014. "Synchronization and Leadership in String Quartet Performance: A Case Study of Auditory and Visual Cues." *Frontiers in Psychology* 5: 645–645.

Waits, Tom. 1992a. "That Feel." *Bone Machine*. Island Records.

Waits, Tom. 1992b. *Bone Machine Operator's Manual*. Island Records.

Yang, Yang, Graham Welch, Johan Sundberg, and Evangelos Himonides. 2015. "Tuning Features of Chinese Folk Song Singing: A Case Study of Hua'er Music." *Journal of Voice* 29 (4): 426–32.

Yust, Jason. 2024. "Tonality and Racism." *Journal of Music Theory* 68 (1): 59–88.

ABSTRACT

This article presents five computer-assisted analyses of musical recordings with a focus on microtuning, describing the details of pitch at sub-semitone level. Microtuning encompasses intonation, tuning systems, pitch inflection, vibrato, and sliding tones. As such, the article focuses on the analysis of recorded music and considers a wide variety of musical traditions: an American folk song, an Ainu lament, a Christian Ethiopian devotional, a Beethoven string quartet, and a Tom Waits song. This article asks how music notation might most effectively capture microtuning in transcriptions and how the analysis of recordings, rather than musical scores, might enrich pedagogy and practice in music theory, performance, and composition.

Keywords: tuning, microtuning, intonation, theory, composition, curriculum reform, ethnomusicology, computer-assisted analysis

RÉSUMÉ

Cet article présente cinq analyses assistées par ordinateur d'enregistrements musicaux, en mettant l'accent sur le microaccord, décrivant les détails de la hauteur au niveau inférieur au demi-ton. Le microaccord englobe l'intonation, les systèmes d'accord, l'inflexion de la hauteur, le vibrato et le glissando. En tant que tel, l'article se concentre sur l'analyse de la musique enregistrée et examine une grande variété de traditions musicales : une chanson folklorique américaine, une complainte Ainu, une dévotion chrétienne éthiopienne, un quatuor à cordes de Beethoven et une chanson de Tom Waits. Cet article s'interroge sur la manière dont la notation musicale peut rendre compte le plus efficacement possible du microaccord dans les transcriptions et sur la manière dont l'analyse des enregistrements, plutôt que des partitions, peut enrichir la pédagogie et la pratique de la théorie, de l'interprétation et de la composition musicales.

Mots clés : accordage, microaccordage, intonation, théorie, composition, réforme des programmes scolaires, ethnomusicologie, analyse assistée par ordinateur

BIOGRAPHY

Taylor Brook writes music for the concert stage, electronic music, music for robotic instruments, generative music software, and music for video, theatre, and dance. His music is often concerned with finely tuned microtonal sonorities as well as unique approaches to the integration of electronic sound and digital media. Brook enjoys collaborating with performers and ensembles on the development of new works, including Mira Benjamin, Jeffrey Gavett, Corey Hamm, Dana Jessen, Andy Kozar, Vicki Ray, Quatuor Bozzini, Del Sol Quartet, JACK quartet, PARTCH ensemble, and many others.