Spectral Sets, Transformations, and Morphologies in Fedor Tau's Performance of *Xöömei*[1]

Lawrence Beaumont Shuster

Unlike the vast majority of global music traditions that privilege pitch and rhythmic relationships as the formative variables of musical production, Tuvan throat singing, or *Xöömei* (also – *Khöömei, Hoomei,* or *Choomeij),* prioritizes tone color, or timbre, as the focus of perceptual interest and appreciation. Tuvan throat singers (*Khoomeizhi*) engage sophisticated vocal techniques in order to generate a fundamental tone accompanied by one or more additional harmonics.[2] The fundamental functions as a drone, sustained throughout the performance while the performer selectively intensifies some harmonics whilst simultaneously dampening others, weaving together delicate melodic designs accompanied by a rich and shimmering kaleidoscopic tapestry of changing harmonic colors and spectral contexts.

Enculturated Tuvan listeners display a deep sensitivity to nuanced sonic detail and the subtle transformations of timbral elements in the performance and reception of traditional Tuvan throat singing. This perceptual focus on tone color is reflected in the sonic qualities associated with Tuvan vocal and instrumental music in general, as well as the design of musical instruments which are engineered to maximize and exploit a broad selection of timbral possibilities. A common characteristic of all throat singing styles is the use of a drone and the accompanying overtones it generates. Süzükei (1994) asserts that this drone-overtone complex results in the production of a specific type of timbral complexity that enculturated Tuvan listeners appreciate as a sonic ideal, iconic of musical meaning and culture at large, and perceived through a perceptual process she characterizes as *Timbre-Centered Listening* (Levin and Süzükei 2006).

Tuvan listeners have developed the ability to detect, interpret, and describe the timbral qualities of a sound with tremendous accuracy and have also evolved a highly complex descriptive language in which to do so. Levin and Süzükei (2018) have organized this descriptive vocabulary into a tripartite classification including: 1) idiophonic vocabulary comprising onomatopoeia and other forms of sound symbolism

1. I would like to express my deepest thanks and admiration for my two anonymous reviewers for their expertise, patience and numerous helpful suggestions.
2. Typically, the terms diphonic and biphonic are both used to designate the presence of a single harmonic sounding in tandem with the fundamental drone whereas the term multiphonic refers to the presence of two or more harmonics sounding above the fundamental. The fundamental pitch is the first harmonic; the second harmonic is the first overtone.

2) the use of cross-modal sensory associations; and 3) the use of affective words (i.e. – warm, soft, prickly, fuzzy, etc.) which are described as establishing a rich, descriptive lexical resource (Levin and Süzükei 2018).

The origins of Tuvan throat singing can be traced to the tribes of the Asian steppe in the vicinity of the Altai and Sayan mountains located in south-central Siberia and western Mongolia (Aksenov, 1973). These communities collectively establish the larger cultural area of Inner Asia including Mongolia; parts of Russia and its member federated states of Khakassia, Tuva, Altai, and Buryatia; and Inner Mongolia and Tibet, both of which are currently autonomous regions within China. The Tuvans are a distinct ethnic group whose cultural traditions such as language, belief systems, lifestyle, traditions, etc. are remarkably well-preserved especially when compared to other neighboring states which were also forcefully occupied as a result of Soviet colonial expansion.

Tuvan beliefs systems are syncretic, integrating elements of indigenous Animism and Shamanism with Tibetan Buddhism. For the Tuvan pastoralist, spirituality is rooted in nature; both in its physical and supernatural manifestations (Edgerton and Levin; 1999). Tuvans believe that natural objects (rivers, trees, rocks, animals); and natural events (rain, light, wind, lightning), have spiritual energies, or souls, that potentially manifest not only as physical objects and/or geographic locations but through sound and light as well. For the mobile Tuvan pastoralists of the steppe, musical utterances function as a sonic reflection and embodiment of their acoustic environment, or soundscape. The sonic models available to them are manifest in the ambient sounds generated by these surroundings; each landscape characterized by a distinctive sonic ambience unique to the specific location and time. Levin's theory of *sound mimesis* (Edgerton and Levin; 1999) asserts that the imitative sounds of Xöömei comprise a particular type of sonic discourse between man and nature on the one hand, and the natural and supernatural on the other.

Levin (1994) explains further that the vast diversity of imitative sounds themselves exist on a spectrum extending from the use of direct imitation, or mimicry, typically involving animal and other natural sounds such as wind, rain, waterfalls, and the turbulence of water on the one side; to more abstract imitative techniques, where imitative sounds are employed to depict visual images such as valleys, landscapes, or mountains. As a result of the importance placed on the sonic imitation and/or depiction of indigenous soundscapes, timbre, and other aspects of sonic design, acquire greater salience and importance, as sonic icons of musical and cultural meaning.

In order to inventory and compare the variety of timbral contexts available in a given performance, new interpretive strategies and listening hierarchies are required.

For enculturated listeners, the sonic features which identify a performance or piece consist primarily of the unique assortment or palette of available timbral colors and textures, and the manner in which they succeed one another in a given performance.

Whereas the majority of formal schemas and processes characteristic of western music can be characterized as primarily linear and goal directed, Tuvan *Xöömei* performances are largely nonlinear and non-processive in scope and purview. A performance starts and stops but does not begin or end. The dynamics of musical design are not characterized by large-scale teleological architectures but instead, by internal consistencies displaying a type of mosaic, nonlinear unity. As opposed to motivic development; there exists potential for constant motivic variation.

In terms of listening hierarchies, considerations involving pitch, intonation, rhythm, and melodic contour, which play a formative role in musical design in most western musical traditions, function more in a supportive, decorative capacity. No two performances by the same performer will ever be identical. Instead of a single definitive version, there is rather an endless series of potential variations based on a common timbral 'theme' and the unique constitution of the latter is invariably informed by sonic qualities and features of the soundscape at hand.

Alekseev, Kirgiz, and Levin (*Voices from the Center of Asia*, 1990) identify five styles of Tuvan Xöömei: *Sygyt*, *Xöömei*, *Kargyraa*, *Ezengileer*, *Borbangnadyr*. Other scholars such as Cosi and Tisato (2002) propose a tripartite stylistic classification and consider Ezengileer and Borbangnadyr as special effects applied to the remaining three styles but not distinct traditional styles in their own right. In general, the primary distinguishing features associated with each style include the registral placement of the fundamental pitch or 'drone'; the number of selectively intensified harmonics sounding simultaneously (biphonic or multiphonic); and other timbral features that collectively establish the unique spectral palette characteristic of each style.

The Sygyt ('whistle') style places the fundamental in the baritone register (100–140 Hz) in order to produce a piercing, clear, whistle-like melody accompanied by a thin, sparse harmonic texture. In the Sygyt style, the performer attempts to isolate and intensify a single harmonic whilst dampening all other harmonics excluding the drone pitch. In addition to serving as a general term for throat-singing, Xöömei also refers to a specific style of throat singing characterized by the positioning of the drone in the baritone range and is also accompanied by high harmonics featuring a more rounded and muted sound than the piercing whistle-like timbres associated with the Sygyt style. Also, whereas the latter attempts to isolate a single harmonic in order to achieve maximal clarity and intensification, the Xöömei style is often multiphonic with more than a single harmonic being targeted for intensification.

Characteristic features associated with the Kargyraa style include the relatively low registral position of the drone (ca. 60Hz) accompanied by a thick texture of densely packed, brassy harmonics. A peculiarity associated with the former style is that in addition to the vocal folds, it also employs the so-called 'false' vocal folds as well which provide two sources of periodic vibrations and thus twice the harmonic density as a result (Edgerton and Levin, 1999). The Ezengileer style features rapid lip movements that mimic the sound of a galloping horse and/or the sound of the rider's stirrups clanging rhythmically across the steppe. Finally, one of the most virtuosic techniques in Tuvan throat singing is the Borbangnadyr style which is similar in terms of its muted timbres to that of the Xöömei style combined with the sounds of chirping birds, gurgling, swirling sounds representative of the rapids of a shallow river, creek or stream. [3]

These same styles are also the general basis for categorization of Mongolian, Khakassian, Altai, and Khai style classifications as well, although each tradition interprets them with a number of regional variations involving language, narrative topic, instrumentation, and ornamentation. Whereas other cultures employ throat singing as a type of embellishment or ornamentation to narrative poetry, or liturgical chant as in the case of several sects of Tibetan Buddhists, it is only in the Tuvan and Mongolian traditions that throat singing emerges as a virtuosic genre of vocal performance to be appreciated in its own right as opposed to fulfilling any ancillary function or decorative role. Instrumental accompaniment and/or text may be included but not necessarily so.

Prior to release of the seminal 1990 Smithsonian Folkways recording, *Tuva: Voices from the Center of Asia*, throat singing had remained largely unknown to western audiences primarily due to factors involving geographic and later, political isolation, resulting from the Soviet colonial occupation (see O'Toole, 2015; Black, 2018). The 1990 recording included field samples collected by an international team of ethnomusicologists consisting of Eduard Alekseev, Zaya Kirgiz and Ted Levin and providing the first scholarly study of Tuvan Xöömei in English.

However, it was through the enthusiasm that followed the release of the highly acclaimed, award-winning documentary *Genghis Blues* (DVD, 1999) pairing San Francisco blues musician Paul Pena with Tuvan national throat-singing icon, Kondar ol Ondar, that the majority of the mainstream public received their initial exposure to the exotic sounds of Xöömei. Once considered rustic and unsophisticated, throat singing today has become a treasured cultural artifact throughout the region and

3. In addition to the aforementioned traditional styles there have been a number of newly developed traditional sub-styles as well over the last three decades; there is also a multitude of recently established hybrid styles that integrate features associated with traditional Tuvan throat singing with other regional and international genres and styles of popular music.

international sensation on the world music stage with skilled performers such as Alash, Huun-Huur-Tu, and Kondar ol Ondar, acquiring rock star status in their respective homelands and on the international stage abroad.

Scholarship since the 90s has further increased our understanding and appreciation of Tuvan throat singing, from a number of scholarly perspectives. The perception and cognition of throat singing has been the source of numerous important psychological studies (see: Bloothooft, Bringmann, van Cappellen, van Luipen and Thomassen, 1992; Barras and Gouiffes, 2008). Consideration of the anatomical and physiological mechanisms involved in the production of sound have also been the subject of rigorous study (Edgerton and Levin 1999; Grawunder, 2009; Bergevin, 2020); as have detailed acoustic studies analyzing various features of sonic and acoustic design (Adachi and Yamada, 1999; Edgerton, Bless, Thibeault, Faberholm, and Story, 1999; Kob, 2004; and Ruiz and Wilken 2018). Additionally, important ethnographic and historical studies such as Aksenov (1973); Van Tongeren (1995); Levin and Süzükei, (2006); Lukov (2008); and others have collectively established a broad framework of pertinent cultural and historical considerations and charted the various roles acquired by traditional Tuvan Xöömei within.

The significance and impact of colonialization on traditional Tuvan culture and music during both the Soviet Union, and subsequently as a federative subject of Russia, has been considered extensively in Süzükei (2007; 2010; 2021). Süzükei's unpublished 2021 paper "Epistemological Problems in Russian Ethnomusicology" reconsiders the significance and impact of research methods implemented by Russian and post-Soviet ethnomusicologists in the study of Tuvan traditional music and how the early disregard towards the Tuvan's unique use of the drone-overtone system resulted in epistemological differences between the Western and post-Soviet scientific approaches to Tuvan music in general. The history, development and influence of Xöömei on contemporary music has been examined by Black (2018); whereas the export and commodification of Tuvan music and culture has been addressed earlier by O'Toole (2015). More recently, complicated issues involving cultural heritage regimes in the post-Soviet era as well as the transregional politics which involve Xöömei as a common heritage shared with other regions in Russia as well as Kazakhstan, Mongolia, and China have been the topic of an important special issue of the *Journal of Asian Music* (52/2 (2021).

Issues involving the conceptual complexities of authenticity, identity, and inclusion in relation to the musical culture surrounding Xöömei is explored by Alexander Glenfield in his 2003 article "The Pearl of Tuva: Authenticity and Tuvan Khorekteer (Throat Singing)" whereby he asserts that the primary markers of authenticity in Tuvan music are ethnicity and race. In his "Post-Soviet Tuvan Throat-Singing and the Circulation of Nomadic Sensibility" Oliver Beahrs (2014) argues how the revival of Xöömei via the world music stage prompted a revitalization of traditional nomadic sensibilities, which he asserts function as both an ideology and as a disposition of contemporary interpretations of Tuvan traditional music. Beahrs demonstrates how the selective use of history, cultural memory, and natural environments are appropriated as sources for constructions of nomadic sensibility embodied in Tuvan music and culture. Interrelationships involving ideology, aesthetics, circulation, and agency in performances of Tuvan traditional music is rigorously examined in Beahrs' seminal 2017 article "Nomads in the Global Soundscape: Aesthetics in Post-Soviet Tuva's Traditional Music Productions". Finally, Lamazhaa Chimiza's 2018 publication "Indigenous Methodology and Research of Tuvan Culture" offers a critical reappraisal of indigenous Tuvan research methodologies in comparison to that of Russian and other foreign scholars and examines their respective strengths and weaknesses from both etic and emic perspectives.

In addition to these important studies, what is needed is the development of new, complementary theoretical perspectives and analytical technologies, informed by indigenous theory and practice, and based on the accumulation of internal evidence gathered from the analysis of the sounds themselves, the results of which can then be contextualized within the existing broader purview of pertinent cultural, historic, and scientific considerations. Enculturated Tuvan listeners already have a highly developed sensitivity and understanding of the manifold varieties of vocal timbres available in Xöömei. What is missing are complementary methods of transcription or other means by which to visualize these timbral correspondences and features with scientific precision. The intention here is to provide analytical tools and strategies useful for exploring the wide diversity of timbral contexts available in Xöömei and a method in which to characterize and inventory them, and thus provide additional levels of technical detail to be combined with the Tuvans highly developed sensitivity to timbre, and indigenous methods of description and interpretation. Iconic timbral features associated with specific styles and performers of Xöömei can not only be identified and described but measured and inventoried in a technically explicit manner, thus providing additional levels of analytical engagement and facilitating nuanced comparison.

The objectives of this study are threefold: (1) to introduce several new analytical tools and strategies useful for characterizing the organization of timbral surfaces and other select features of sonic design in Tuvan throat singing (2) to revisit Robert Cogan's theories of spectral morphology from the 80s (Cogan 1982; 1998) and offer some methodological extensions and refinements facilitated by the availability of new software technologies, and (3) to demonstrate the general utility and implementation of both of the aforementioned considerations using an excerpt from Fedor Tau's legendary performance of *Xöömei* (Tuva: Voices from the Center of Asia, 1990; CD).

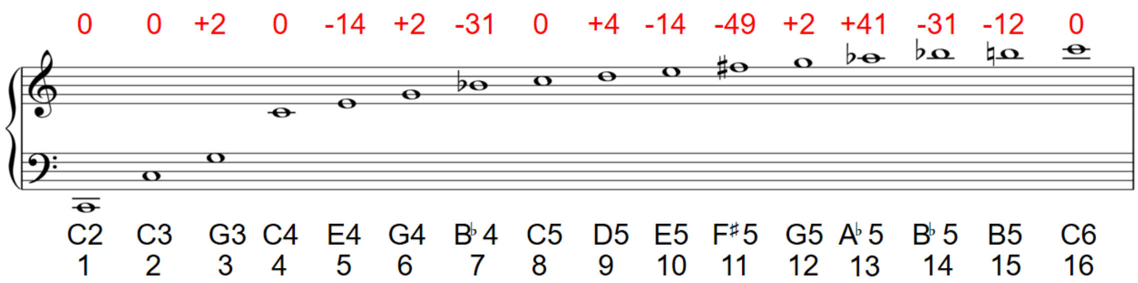
Audio 1. Excerpt from Fedor Tau's performance of Xöömei (Tuva: Voices from the Center of Asia, 1990; CD).

In vocal music, the source of the acoustic signal is the periodic vibrations of the vocal folds and in the case of the Kargyraa style of throat singing, the false vocal folds as well. Mathematically, when a fundamental pitch is produced, the number of harmonics generated by that fundamental is infinite; each successive order position in the harmonic series being an integral multiple of the former and growing every closer in frequency as the divergent series progresses. The approximate boundaries for average human audition are roughly 20 to 20000 Hz and within this audible range, Tuvan pentatonic melodies employ a relatively small spatial ambitus with the lower and upper boundaries defined by the 6^{th} and 12^{th} harmonics, respectively (Edgerton and Levin, 1999).

Once the vibrating column of air has been generated it is subsequently filtered as it passes through the vocal tract and oral cavities respectively, each of which produce their own intrinsic resonant frequency bandwidths, or formants, as informed by the dimensions and acoustic properties of each cavity. As opposed to an individual frequency, formants are comprised of bandwidths of closely positioned frequencies. When the frequency of a given harmonic sounds within the boundaries of the resonant frequency bandwidths, the harmonic consequently becomes intensified. Harmonics which fall outside of frequency bandwidths in general become dampened or suppressed. Once the sound exits the mouth, environmental and radiant sonic diffusion considerations further condition the sound.

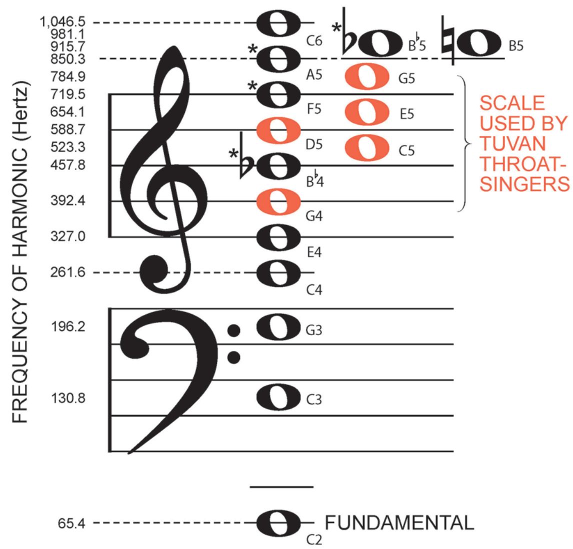
In the majority of vocal music traditions, the frequencies of the fundamental pitches that form the melody as well as their associated harmonics, are in a state of continual flux while the formant frequencies remain more or less static with small degrees of variance. In Tuvan, and central Asian throat singing more generally, the situation is reversed. It is the fundamental pitch, its corresponding harmonics, and associated vowel sounds that remain largely static while the performer shifts and merges the lower order formant frequency bandwidths to intensify selected harmonics. As a consequence, significant contributions to vocal tone color derive more from these shifting and combined formants and the specific harmonic(s) they intensify than with the fundamental pitch, which is usually only barely audible, and its corresponding vowel color.

Example 1 illustrates the harmonic series associated with the fundamental pitch, C2. The corresponding pitches appear on the bottom portion of the example. The respective order position and corresponding octave registrations for each pitch are displayed below.



Example 1. The harmonic series with pitch deviations from equal temperament.

Example 2 demonstrates the same harmonic series using a vertical orientation as reproduced from Levin and Edgerton's Scientific American article (1999). The notes appearing in red distinguish the Tuvan pentatonic scale consisting of order positions 6 through 12 of the harmonic series with their corresponding pitches G, Bb, C, D, E, F, and G. Note that although the 7^{th} and 11^{th} harmonics, corresponding to the pitches Bb4 and F5 respectively, occupy a position within the available musical space defined by the pentatonic scale they are treated as 'avoid tones' and typically not employed in Tuvan throat singing melodies.

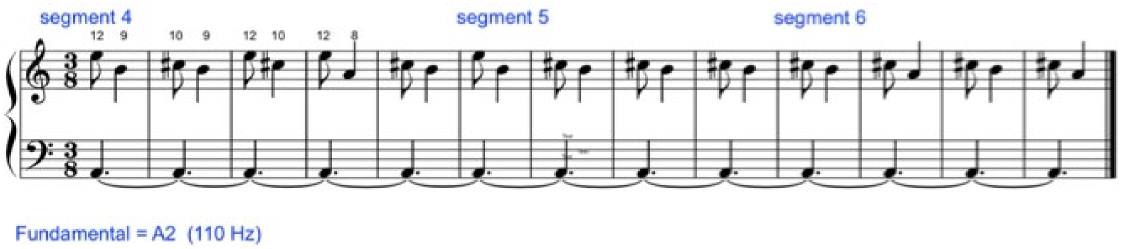


Example 2. The harmonic series using a vertical orientation.

Example 3 presents an analytical transcription of the latter half of the performance of Fedor Tau's *Xöömei*.[4] The primary line is comprised of those harmonics selectively

4. Fedor Tau was born in Ak-Erik (1929–2006) and is one of the most celebrated Tuvans of his generation, highly regarded for his consummate skills in throat singing, especially Xöömei and Kargyraa; and as a national wrestling and skiing champion. Tau was awarded the prestigious title of Peoples Khoomeizhi of the Tuva Republic in 1993 following release of the Smithsonian Folkways recording Tuva: Music from the Center of Asia (1990).

intensified by the performer and demonstrated on the top staff in phrase segment 4. The numbers appearing above the staff indicate the respective order position for each pitch within the harmonic series. The fundamental pitch, A2, appears on the bottom staff and is prolonged throughout the entirety of the performance. As an oral practice, Tuvan throat singing had no tradition of notation prior to the arrival of the Soviets. The transcription provided below is not intended as an accurate portrayal of the performance nor as a reference for performance, but rather, simply a visual map useful for indicating specific musical components and relationships among pitch classes.



Example 3. Transcription of Phrase Segments 4, 5, and 6.

FORMAL SHAPE AND ARTICULATION

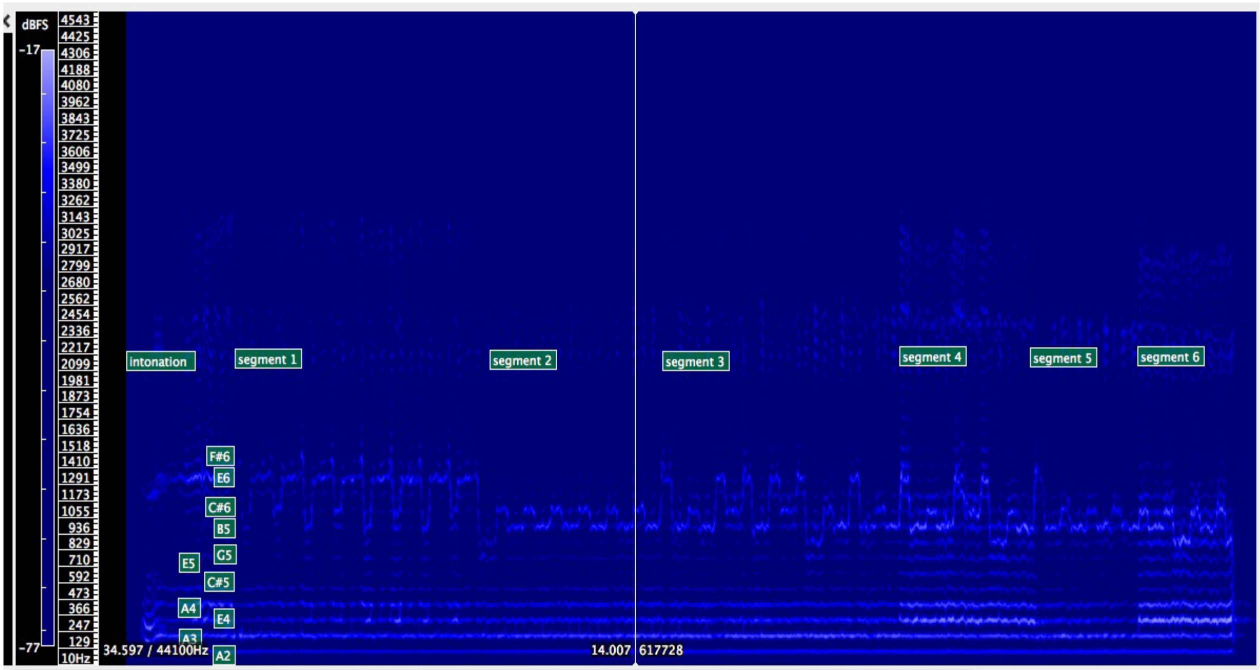
The performance consists of six phrase segments. Segments are distinguished by subtle and nuanced changes in melodic contour, pitch collection, as well as contrasting timbral features. The internal organization of each segment is characterized by a high degree of internal consistency engendered through the use of a small number of both pitch classes and durational values which are continually recycled with nuanced and varied repetitions within each segment. Phrase segments start and stop but do not begin and end. Successive segments are not connected by any type of functional transitions; nor do the segments combine to establish some type of large scale formal developmental scheme or ordering. Instead, junctures between successive segments are characterized by immediate, yet subtle contrasts achieved by small changes in melodic contour and surface rhythm. And, in the case of segments four, five, and six; through pronounced contrasts in spectral intensity as well.

Segment #	Timings	Pitch Collection	C- segments[5]
Segment 4	21.046-25.114	[B, C#, E]	<3121323021>
Segment 5	25.114-27.838	[B, C#, E]	<20101010>
Segment 6	27.838-30.632	[A, B, C#]	<21202120>

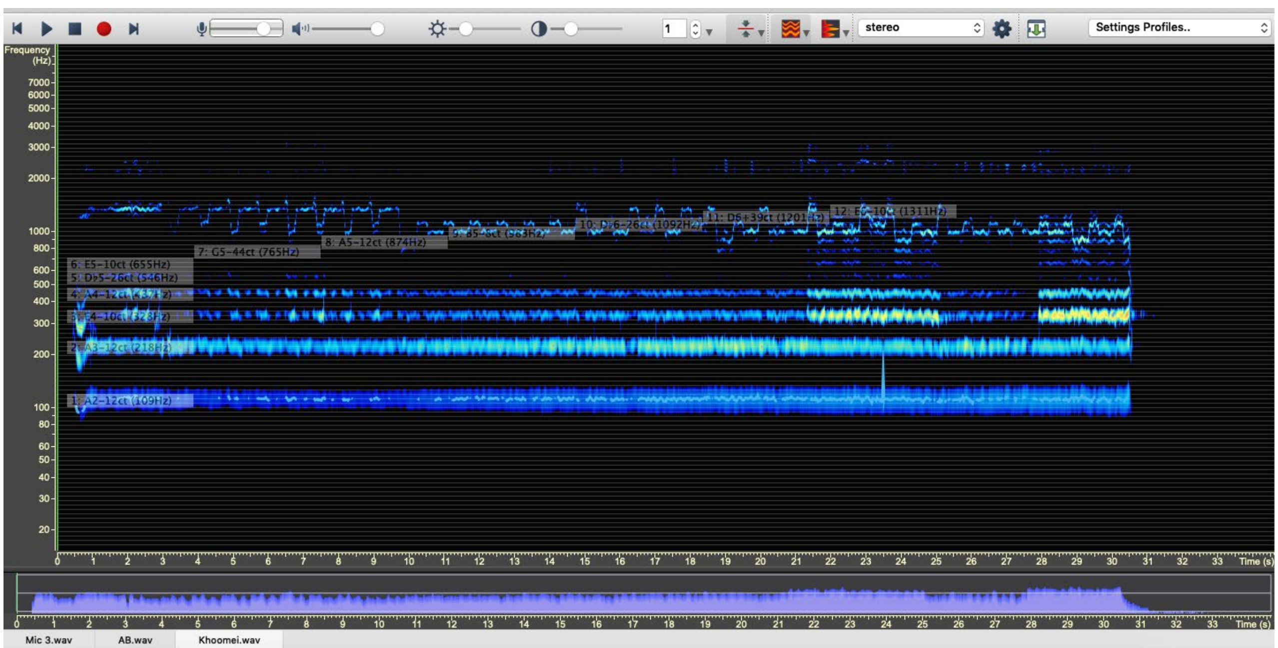
While successful in terms of illustrating the fundamental pitch and the primary melodic line, numerous other important sonic features are left unaccounted for, incapable of being adequately depicted using conventional western notation. In order to gain additional detail and perspective regarding the sonic design of Tau's performance of *Xöömei*, a spectrographic image of the performance is displayed in Example 4. In reading the graph, the following criteria apply. The horizontal axis indicates time, and the vertical axis represents frequency measured in Hertz (Hz). Located to the left of the frequency scale, the amplitude of each sound is indicated by the colored decibels relative to full scale or dBFS values which provides a contextual measurement for amplitude where the highest value on the recording is assigned the value '0' with all subordinate values indicated as negative values. Pitch and octave registrations are indicated adjacent to the spectral image located in the main body of the graph.

The brightest spectra indicate the selectively intensified harmonics that form the primary melodic line. The six phrase segments that comprise the performance are displayed in the central portion of the graph. While spectrographs provide an objective analysis of the acoustic signal in exacting detail and with tremendous technical precision, our subjective experience often results in a case of what you see is not what you hear but often far from it. The perceptual liabilities and constraints involved with interpreting spectrograms has been given detailed scrutiny by Cook and Leech-Wilkinson in *A Musicologist's Guide to Sonic Visualiser* (2009). The analytical method employed herein, while not a perceptual study, is supported by perceptual and cognitive considerations and through use of analytical methodologies and verification protocols established by Shuster and Wey (2021, 2022).

5. Contour segments or 'C-Segs' abstract and summarize the melodic contour of a given melody by eliminating precise intervallic measurements and replacing them with a contextual measurement of respective vertical order position. Accordingly, the lowest pitch in a contour segment is always 0; the next lowest pitch is 1 and so on.



Example 4a. Narrow-band spectrogram – linear perspective.



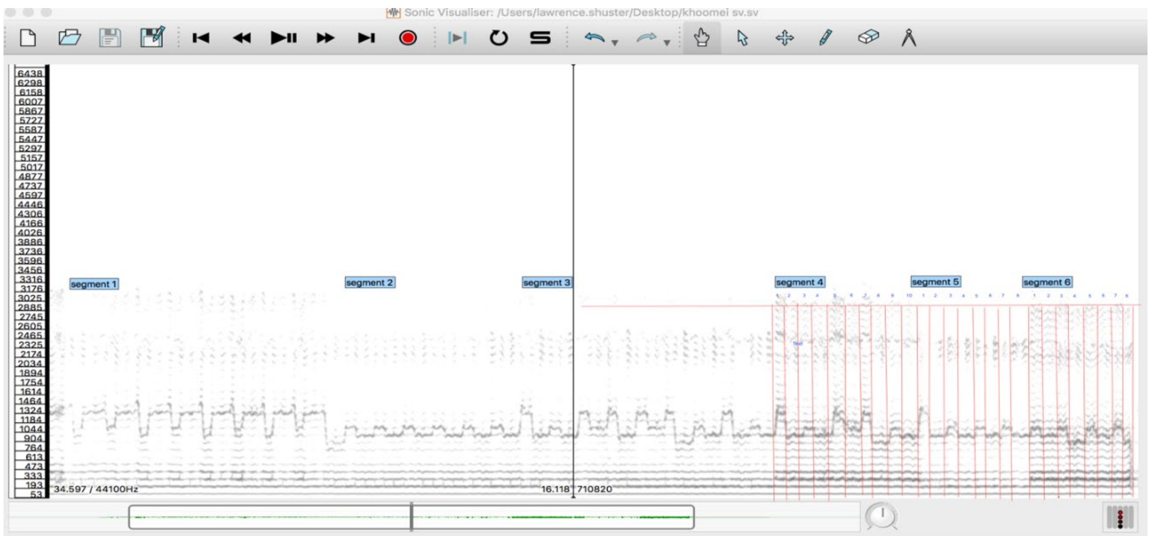
Example 4b. Narrow-band spectrogram – logarithmic perspective.

One characteristic quality in which the relative skill of throat singers is evaluated is to what extent they can successfully suppress or dampen some harmonics while selectively intensifying other harmonics simultaneously. In order to distinguish between harmonics that are targets of selective intensification – those which form the primary line – as opposed to other sounding harmonics within the texture, the former are designated as '*primary'* harmonics and the latter, '*associated*' harmonics. In *Xöömei*, these associated harmonics form a textural background against which the primary melodic line is unfolded by the more prominent primary harmonics. Observe that the brightest spectra, positioned in the upper registers, represent those harmonics targeted by the performer for selective intensification. As evident in the spectrographs, at the same time as these primary harmonics are being highlighted there are other harmonic details in play within our perceptual experience as well. These include the lower order positions within the harmonic series between f0–f4. As visualized within the vocal source-filter model of timbre, the lower order positions of the harmonic series are the most prominent in our perceptual experience and become increasingly softer as one ascends the series. It is for these reasons that these harmonics appear invariant throughout the performance. Irrespective of the technical skill of the performer, these lower order harmonics impact our perceptual discriminations in that they are incapable of being entirely dampened. Still others result from inflections associated with the unique 'sonic fingerprint' of the performer—those features which allow us to differentiate between two singers performing the same melody. Others may result as artefacts of the recording process and environment.

Vertical spectral segmentations

In order to explore, characterize, and interpret the timbral organization of *Xöömei* we need some way to sort and manage the overwhelming complexity of the timbral surface and reduce it to smaller, more basic elements or 'building blocks'. Once these basic units of analysis have been disclosed, we can search for correspondences and various forms of relationship between them. Example 5 illustrates the how phrase segments four through six are further subdivided to form a succession of 26 vertical spectral segmentations. The durational boundaries of each spectral segment are co-extensive with the duration of the corresponding selectively intensified 'target' harmonic whose successions inform the primary line: phrase segment 4 contains vertical spectral segments 1–10; phrase segment five contains vertical spectral segments 11–18; and phrase segment six contains spectral segments 19–26.

The succession of vertical spectral segmentations can be visualized as a series of adjacent vertical slices through the spectrogram; each segment including the fundamental pitch (drone) which appears as the bottommost harmonic; those harmonics selectively intensified by the performer that combine to form the primary line; as well as all additional sounding associated overtones. On occasion, sounds not associated with the harmonic series and its corresponding fundamental frequency are also present in the signal and thus are represented in the data sample. Some of these elements result from the previous sound via reverberation and delay; others are artefacts of the recording process; whereas others stem from acoustic and environmental sources. Because our intention is to isolate and model as much as possible the specific tone color properties and other sonic features associated with Fedor Tau's voice apart from any accompanying acoustic artefacts or other forms of noise and interference, they are not included in the subsequent analysis of the spectral sets.



Example 5. Spectrogram with vertical spectral segmentations (phrase segments 4, 5, and 6).

Measuring harmonics

Because the spectrogram displays the objective data accomplished by the computer as opposed to the human ear, much of the visual information presented in the spectrograph is far beyond the range of human perception. Therefore, in order to gauge an approximate threshold for salience that we can use to distinguish those sonic features we can hear from those we cannot, several measurements are taken, the results of which then provide the requisite perceptual foundation for the analytical assertions and observations that follow whilst simultaneously offering some means, however limited, of phenomenological accountability.

Three types of measurements are employed in order to distinguish and characterize the approximate spectral intensity for each harmonic in each vertical spectral segment. The results were then contrasted, compared and sifted and all harmonics deemed perceptible were included as elements within the spectral set. Two different types of measurements were used in measuring spectral intensity. The former, made possible by the VoceVista Video software, measures the corresponding *amplitude* for each harmonic in each spectral set using two types of values: *peak* intensity and *average* intensity, both of which are provided in terms of decibels relative to full scale (dBFS) and are objective measurements of sound pressure or amplitude. The latter, using WAVANAL software initially created for the analysis and tuning of church bells, provides an approximate perceptual value for *loudness* as measured in *phons*.

Example 6 illustrates the respective peak amplitude as measured in decibels (dBFS) for each harmonic in each vertical spectral segment of *Xöömei*. Observe that the fundamental pitch, or drone, appears in highlighted blue squares along the bottom horizontal row of the chart. The squares highlighted in gray indicate those harmonics targeted for selective intensification and whose succession comprise the primary melodic line. All remaining harmonics are generated by and thus associated with the fundamental pitch (drone) but not necessarily salient, or perceptible as such.

			Segment 4										Segment 5								Segment 6								
Freq. (Hz)	Harmonic Series	Pitch / Octave																											PE
1318	12	E6	-38.4	-65.6	-60.8	-66.8	-41.3	-64.2	-40.9	-65.5	-65.2	-65.4	-54.4		-68.7	-74.9	-61.7	-74.8	-73.9	-78.3	-67.7	-68.9	-60.9	-71.3	-60.2	-71.5	-61.8	-66.3	
1244	11	D#6*	-46.3	-58.4	-54	-60.5	-51.2	-49.9	-52.3	-69.8	-60	-63.6	-65.7	-67.6	-67	-69.1	-75.5	-70	-61.7	-68.9	-53.3	-60.1	-51.3	-62.2	-60.4	-57.9	-55.4	-61.7	
1108	10	C#6	-55.9	-56.1	-42.4	-55.6	-61	-46.7	-60.3	-62.3	-45	-54.4	-73	-64.3	-50.3	-63.3	-51.3	-66.3	-51.1	-66.9	-42	-53.8	-39.4	-64.1	-40.5	-54.3	-39.9	-43.5	
987	9	B5	-60	-39	-53.1	-38.2	-54.5	-54.8	-59.6	-52.3	-43	-39.8	-61.2	-42.6	-51.7	-44.5	-56.1	-43.9	-55.1	-43.3	-41.4	-46.5	-54.1	-43.9	-50.1	-37.1	-56.4	-43.8	
880	8	A5	-62.5	-53.4	-60.9	-54.7	-63.5	-63.9	-69.3	-47.8	-60	-59.7	-74.4	-64.5	-63.9	-64	-63.2	-63.9	-62.8	-64.6	-54.1	-48.3	-55.4	-41.1	-53.2	-52.9	-56.9	-41	
781	7	G5*	-57.5	-57.3	-58.6	-55.7	-59.9	-59.8	-60.2	-48.5	-58	-56.7	-70.5	-67.9	-65.3	-67.1	-66.2	-65.9	-67.4	-68.2	-60.9	-57.9	-61.2	-53.5	-58.4	-56.2	-64.1	-60.7	
659	6	E5	-59.3	-55.6	-58.5	-56	-62.4	-60.9	-63.6	-56.7	-60.6	-59.3	-66.1	-71.8	-68.9	-70.1	-72.9	-69.7	-70	-70.6	-58.3	-54.8	-58.2	-54.4	-56	-57.7	-61.5	-55.1	
554	5	C#5	-57.1	-60.2	-60.9	-60.2	-58	-61.4	-56.7	-55.7	-60.1	-58	-72.9	-65.4	-64.1	-68.3	-66	-69.1	-70.9		-61.6	-62.2	-60.6	-61.7	-56.4	-61	-59.7	-63.7	
440	4	A4	-41.2	-40	-45	-40.6	-41.1	-44.3	-41.6	-41.1	-46.4	-45.6	-55.5	-57.1	-55.4	-54.8	-51.6	-56.8	-55.7	-60	-42.2	-39.1	-42.5	-42.8	-41.1	-43.2	-42.5	-40.1	
329	3	E4	-43.2	-43.3	-39.9	-43	-42.2	-36.3	-41.7	-35.9	-38.2	-38.3	-46	-51.2	-48.3	-54.3	-47.5	-53.7	-53.4	-53.8	-34.8	-32.1	-36	-33.6	-34	-33.3	-30.9	-27.8	
220	2	A3	-43.1	-42.5	-42.1	-42.2	-43.8	-44.1	-43.7	-41.8	-43.1	-42.7	-42.5	-44.4	-39.2	-43.3	-43	-42.9	-43.1	-41	-42.8	-42.8	-45.3	-42.6	-45.9	-44.3	-45.5	-43.5	
110	1	A2	-50.9	-51.5	-50.4	-50.1	-50.7	-50.8	-50.9	-50.9	-51.4	-51.9	-50.5	-51.4	-51.2	-50.9	-49.7	-50	-50	-48.9	-49.9	-49.9	-50	-49.2	-50	-49.7	-49.5	-47.9	
			1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	

Example 6. Peak amplitude measured using the decibels relative to full scale (dBFS).

Example 7 illustrates the respective *average* as opposed to *peak* amplitude also measured in decibels (DBFS) for each harmonic, in each spectral in phrase segments 4, 5, and 6. Each type of measurement has implicit advantages and combining data from both can often help in deciding to what extent a given harmonic is perceptually salient. Within the boundaries of a defined spectral segment, the peak amplitude reveals the loudest instantiation of the corresponding harmonic; the average reveals the mean of all sounding instantiations of the same harmonic.

Both considerations prove useful depending upon the scope of insights and motivations intended at hand. Regardless of whether the intensity values are assumed at peak or average decibels the problem remains that amplitude is an objective measure of sound pressure in general, and the dBFS scale, in particular, is a contextual measurement where the loudest sound in the signal is assigned a value of '0' and all remaining sounds are measured in comparison and labeled with negative integers accordingly. It is for this reason that implementing some sort of perceptual threshold is desirable in order to make some approximate distinction between that which we can hear and that which we cannot. Unfortunately, given the complexities involved in the perception and cognition of musical sounds, for now approximations seem the best way forward while awaiting the eventual arrival of more sophisticated and accurate measurements of perceptual experience.

Freq. (Hz)	Harmonic Series	Pitch / Octave	Segment 4										Segment 5								Segment 6								AV
1318	12	E6	-46.2	-65.6	-60.8	-66.8	-45	-64.9	-40.9	-72.2	-65.2	-74.3	-54.4		-72.5	-74.9	-61.7	-74.8	-73.9	-79.9	-67.7	-68.9	-64.6	-71.3	-60.2	-71.5	-61.8	-70.1	
1244	11	D#6	-56	-61.7	-54	-62.2	-51.2	-53.9	-55	-69.8	-60	-67.9	-65.7	-73.2	-67	-69.1	-75.5	-71.8	-61.7	-68.9	-53.3	-61.3	-55.3	-68.8	-60.4	-66.3	-55.4	-69.6	
1108	10	C#6	-64.9	-58.3	-43.5	-55.6	-63.3	-46.7	-67.7	-64.7	-45	-65.7	-73.6	-64.3	-51.3	-63.3	-51.3	-76.1	-51.1	-66.9	-42	-53.8	-39.4	-64.1	-46.7	-54.3	-39.9	-43.5	
987	9	B5	-61.6	-41.2	-53.1	-42.4	-59.8	-54.8	-59.6	-54.1	-52.8	-41.2	-71.7	-42.6	-57.8	-45.5	-60.7	-45.3	-60.5	-46.2	-53.6	-46.5	-54.1	-43.9	-51.8	-44.8	-57.5	-56.6	
880	8	A5	-65.3	-55	-63	-57.5	-63.3	-63.9	-69.3	-47.8	-65.8	-59.7	-74.4	-64.5	-71	-64	-68.6	-63.9	-68.9	-68	-57.4	-51.9	-57.9	-41.1	-55.6	-52.9	-58.4	-41	
781	7	G5	-63.1	-59.2	-61.1	-58.3	-64.8	-60.8	-62.8	-53.1	-59.4	59.8	-70.5	-67.9	-68.5	-68.8	-66.2	-68.2	-67.4	-71.9	-60.9	-57.9	-64.2	-53.5	-64.7	-62.9	-64.1	-60.7	
659	6	E5	-62.3	-57.5	-60.6	-59.2	-62.4	-60.9	-64.2	-57.9	-62.2	-61.6	-71.4	-74.4	-68.9	-70.1	-72.9	-69.7	-70	-70.6	-58.3	-56.4	-60.6	-55.7	-61.9	-59.3	-61.5	-55.1	
554	5	C#5	-61.9	-61.5	-63.4	-60.9	-67.8	-61.4	-60.6	-58.4	-60.9	-62.2	-72.9	-72.4	-74.8	-68.3	-70.9	-74.2	-72.7		-61.6	-62.2	-63.8	-62.6	-60.2	-61	-65.6	-66.1	
440	4	A4	-44.9	-42.8	-46	-42.8	-45.6	-44.3	-45.7	-44.6	-47.4	-46.1	-59.8	-57.1	-60.2	-54.8	-58.4	-59.5	-58.7	-64.3	-46.6	-42.1	-44.4	-44.6	-46.9	-44	-46.5	-40.1	
329	3	E4	-39.2	-36.3	-36.1	-34.8	-36.6	-36.3	-41.7	-37.2	-38.3	-38.3	-53	-51.2	-53.1	-54.3	-54.2	-54.6	-53.4	-58.3	-34.8	-37.8	-39.7	-33.6	-34	-33.3	-33.1	-30.2	
220	2	A3	-43.1	-42.5	-43.8	-43	-43.8	-44.1	-43.7	-41.8	-44.3	-45.8	-45.6	-45.9	-40.9	-43.3	-43	-42.9	-45.8	-47.1	-42.8	-43.6	-45.3	-43.8	-46.9	-46	-47.9	-43.5	
110	1	A2	-50.9	-51	-51.4	-50.1	-50.7	-51.3	-50.9	-50.9	-51.4	-51.3	-51.8	-52.4	-51.2	-50.9	-49.7	-50	-50	-49.9	-49	-49	-50	-49.2	-51.3	-49.7	-51.5	-49.3	
			1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	

Example 7. Average amplitude measured using the decibels relative to full scale (dBFS).

Edit Partials

Nominal Frequency (Hz) 110.5 Note A(-1) +8

OK

Recalc

A=440.0Hz

Freq. (Hz)	Splash Amp.	Tail Amp	Attack	Decay	Cents	Partial	Note
110.5	0	0.399307364	0	0	0	nominal	A(-1) +8
221.5	0	0.990556726	0	0	1203.9	oct. nom.	A(0) +12
332	0	1.459197128	0	0	1904.5		E(1) +12
441.5	0	0.542113444	0	0	2398		A(1) +6
550.5	0	0.077921310	0	0	2780		Db(2) -11
660.5	0	0.082882373	0	0	3095.4		E(2) +3
781	0	0.092584920	0	0	3385.5		G(2) -6
889.5	0	0.083850373	0	0	3610.7		A(2) +19
1000	0	0.227552709	0	0	3813.4		B(2) +21
1074	0	0.249076281	0	0	3937		C(3) +45
1195	0	0.093358401	0	0	4121.8		D(3) +30
1211.5	0	0.178726585	0	0	4145.6		Eb(3) -46
1319.5	0	0.214371440	0	0	4293.4		E(3) +1
1425	0	0.063626747	0	0	4426.6		F(3) +34
1545.5	0	0.052567474	0	0	4567.1		G(3) -24
2092.5	0	0.030940065	0	0	5091.7		C(4) +0
2206	0	0.039883654	0	0	5183.1		Db(4) -8
2435	0	0.075863809	0	0	5354.1		Eb(4) -37
2493	0	0.075272819	0	0	5394.9		Eb(4) +3
2546.5	0	0.060903055	0	0	5431.6		Eb(4) +40
2696.5	0	0.032739346	0	0	5530.7		E(4) +39
2765	0	0.031123973	0	0	5574.1		F(4) -17
2875	0	0.037008694	0	0	5641.7		F(4) +50
2982	0	0.039887113	0	0	5704.9		F#(4) +13
3061	0	0.039697428	0	0	5750.2		G(4) -41
Noise	0	0	0	0			

Clear

Example 8. This table illustrates the respective *average* loudness using the perceptual measurement of phons for each harmonic within each vertical spectral segment of *Xöömei*.

The perceptual threshold for salience for this recording is set at an amplitude of -51 dBFS for either peak or average intensity values, as measured in decibels. The verification protocols used to provide perceptual support for these assertions were adopted from Shuster and Wey's "Mapping Timbral Surfaces in Alpine Yodeling: Towards A New Theory of Tone Color for Unaccompanied Vocal Music" (2021). Shifting the salience threshold is akin to adjusting the level of magnification with more or less spectral detail emerging whether one zooms in or out. The complete data set for each of the three aforementioned measurement considerations are included in the appendix.

VERTICAL SPECTRAL SETS

In an unpublished paper from the 2012 Annual Meeting of the Society for Music Theory, I introduced the concept of spectral sets as an analytical tool useful for characterizing the relationships between a fundamental pitch, or drone, and its accompanying harmonics, and as means by which to characterize a basic profile of tone color in unaccompanied vocal contexts.

Spectral sets represent the basic building blocks of timbral surfaces and summarize and encapsulate many important qualities regarding the harmonic content and timbre associated with a given vertical spectral segment. In *Xöömei*, each spectral set contains a fundamental pitch which functions as a drone, (A2) and thus determines the palette of available harmonics and their unique timbral qualities. The drone is prolonged throughout the performance using a single, sustained breath. In addition, each set also contains both the target harmonic(s) selectively intensified by the performer to establish the primary melodic line as well as all additional sounding, or associated harmonics.

The analytical notation for the spectral set is <H1, H2, H3, H4, **H5**> where the use of angle brackets denotes the ordered collection of the harmonics series. The first harmonic, H1, is also the fundamental pitch. Individual order positions for each harmonic within the series are designated accordingly alongside their respective pitch class and octave registrations. Observe that the use of bold-faced font with underline is reserved for designating the target harmonic(s) selected for intensification by the performer. The analytical notation provides an efficient means by which to inventory, catalogue, and compare the wide diversity of sonic contexts and experiences available in Tuvan biphonic and multiphonic throat singing with archival accuracy and technical precision. We don't perceive the individual sonic elements that comprise the membership of the set on a one to one, individual basis. Rather, the harmonics and other sonic features blend, merge, and fuse into a singularity perceived as a composite timbre, or tone color. By reducing the complexity of a given sound into its basic constituent elements and then reverse engineering the process via reassembly we gain additional insight and nuanced perspective regarding the internal organization of a sound, despite the fact that we never perceive these individual elements outright as such, but only their composite exterior, as it were.

Example 9 demonstrates the succession of the 26 vertical spectral sets that collectively partition phrase segments 4, 5, and 6. Observe that phrase segment 4 contains 10 vertical spectral sets whereas phrase segments 5 and 6 contain each contain 8. Each vertical spectral set contains the same fundamental pitch (A2); the same collection of lower order associated harmonics (harmonics nos. 1–4); and the selectively intensified target harmonics appearing in shaded squares in Example 9. Observe that many of the features which comprise the timbral surface remain largely constant and unchanging such as the continually recycled statements of the drone, the presence of the lower order harmonics in order positions 1–4, as well as the nearly constant rate of the vibrato. These combine to create a strong sense of uniformity and consistency across the performance. The one thing that does change however is which harmonic within the palette of available colors generated by the fundamental drone will be targeted for selective intensification at any given point.

Shuster and Wey (2021) assert that in unaccompanied vocal music it is the corresponding vowels and the ratio of the formants which disambiguate them that ultimately determines vocal tone color, with variations as a result of changes in set cardinality, the respective order positions of member harmonics, and the consistency of the fundamental pitch, resulting in a tremendously broad assortment of available shades and hues to the primary vowel color. In *Xöömei*, the drone as well as the invariant, lower order position harmonics (H1–H5) create a single tone color that informs the entire performance; distinctions regarding which harmonic is being selectively intensified at any given time within the primary line offer a series of available variant shadings, hues, and intensities.

Accordingly, vertical spectral sets 1, 4, 6, and 9 feature the intensification of the twelfth harmonic, E6; vertical spectral sets 3, 6, 9, 13, 15, 17, 19, 21, 23, and 25 highlight the 10th harmonic, C#6; whereas vertical spectral sets 2, 4, 10, 12, 14, 16, 18, 20, and 24 prioritize the 9th harmonic, B5; and, vertical spectral sets 8, 22, and 26 accenting the 8th harmonic, A5.

Segment 4

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A2	A3	E4	A4	- -	-	-	-	-	-	-	**E6**
A2	A3	E4	A4	-	-	-	A5	**B5**			
A2	A3	E4	A4	-	-	-	-	B5	**C#6**		
A2	A3	E4	A4	-	-	-	-	**B5**			
A2	A3	E4	A4	-	-	-	-	-	-	D#6*	**E6**
A2	A3	E4	A4	-	-	-	-	-	**C#6**		
A2	A3	E4	A4	-	-	-	-	-	-	D#6*	**E6**
A2	A3	E4	A4	-	-	G5*	**A5**				
A2	A3	E4	A4	-	-	-	-	B5	**C#6**		
A2	A3	E4	A4	-	-	-	-	**B5**			

Segment 5

A2	A3	E4	-	-	-	-	-	-	-	-	**E6**
A2	A3	E4	-	-	-	-	-	**B5**			
A2	A3	E4	-	-	-	-	-	-	**C#6**		

A2	A3	E4	-	-	-	-	-	**B5**	
A2	A3	E4	-	-	-	-	-	-	**C#6**
A2	A3	E4	-	-	-	-	-	**B5**	
A2	A3	E4	-	-	-	-	-	-	**C#6**
A2	A3	E4	-	-	-	-	-	**B5**	

Segment 6

A2	A3	E4	A4	-	-	-	A5	B5	**C#6**
A2	A3	E4	A4	-	-	-	A5	**B5**	
A2	A3	E4	A4	-	-	-	-	B5	**C#6**
A2	A3	E4	A4	-	-	G5*	**A5**		
A2	A3	E4	A4	-	-	-	A5	B5	**C#6**
A2	A3	E4	A4	-	-	-	A5	**B5**	
A2	A3	E4	A4	-	-	-	-	-	**C#6**
A2	A3	E4	A4	-	-	-	**A5**		

Example 9. Vertical Spectral Sets.

The relatively small assortment of selectively intensified harmonics employed in the primary line (E6, C#6, B5, and A5) are organized and varied so that they create a succession of simple patterns in pitch space.

In phrase segment 4, the pattern begins on E6, moves down to B5; up to C#6 and back down to B5 as apparent in spectral segments 4–6. Observe that a slight variation of the same pattern occurs in spectral sets 7–10. The replication of simple cellular units establishes a sense of unity amongst pitch resources. Whereas timbre might emerge as a featured parameter of the ideal sonic experience it does not mean Tuvans were unaware of the linear, motivic dimensions of pitch design.

Additional forms of nonlinear uniformities and consistencies derive from the symmetrically distributed pitch classes (the harmonics) in pitch space. Observe that vertical spectral sets 1 and 11 both contain the twelfth harmonic, E6, which delineates the starting and ending points for what would be a palindrome excepting that the harmonic A5 has now replaced B5 as it appeared in spectral segment 5. Additional instances of pitch symmetry include phrase segment 5, segments 12–19; and bipartite symmetry (segments 19–22; 23–26) implicit in phrase segment 6. Note that each phrase segment emphasizes a different portion of the pitch space defined by the harmonic series: phrase segment 4 uses the upper region denoted by the 9th, 10th, and 12th harmonics respectively; phrase segment 5 presents the simple oscillation of B5 and C#6 employing the middle range associated with the 9th and 10th harmonics; whereas phrase segment 6 extends the range down another step to include the 8th harmonics as well.

In Shuster/Wey (2021) a number of potential forms of relationship available between spectral sets are defined. These are based on two categories of correspondence, pitch and timbre, and facilitated by adapting tools and techniques from traditional set class theory. Potential forms of correspondence based on pitch relationships include the following: the cardinality (aka. – density) of the spectral and the respective order positions of the harmonic series in play. Sets sharing the cardinality and fundamental pitch will be related by the identity function or T_0. If they share the same cardinality but have different fundamental pitches, they are related by T_x where x is equal to the number of half steps in the transposition. Moreover, sets related by T_0 or T_x can also display an inclusion relation. Consequently, one collection of pitches will form a subset or superset of the other (literal subset/superset); or alternatively, be a transposed subset/superset of another (abstract subset/superset).

Timbral correspondences in unaccompanied vocal music are typically based first and foremost on the quality of the vowel associated within the spectral set. Phonologists have established that it is the particular arrangement of the first three formant frequency bandwidths (F1–F2–F3) that permit the vowel to be recognized or disambiguated. If two sets share a common vowel but have different fundamental pitches, they are related by transposition shown as T_x, with x equal to the number of semitones between respective fundamental pitches. In the case of discrepant cardinalities, the degree of offset is shown in parenthesis $T_x(_y)$ as before. We can conceptualize the various types of relationship available in the form of a timbral hierarchy extending from those which are most salient (identity, T_0; same vowel, fundamental, and cardinality); and salient (same vowel, same fundamental, literal subsets).

Whereas this approach is successful in mapping vocal timbres in conventional unaccompanied vocal music it proves insufficient for explaining the timbre and sonic design characteristic of Tuvan throat singing. The central constraint regards the presence of the drone. The drone is the ultimate aural reference point. In harmonic singing the drone not only generates the other pitches; the harmonics and the manner of their configuration ultimately derive their meaning and significance in relation to the drone. Like the sun in our model of the solar system, the drone occupies the central position of radial symmetry, and exerts varying degrees of gravitational attraction on the various harmonics it generates. The execution and articulation of the drone determines how many harmonics are included in the palette of available tone colors and how 'salient' each harmonic is within the boundaries of its respective vertical spectral set. In short, the drone and the harmonics form an indivisible sonic unity, or singularity. To consider either one independently would only distort the underlying phenomenological foundations of the music.

Unlike conventional singing where the physical dimensions of a particular vowel shape determine the resultant formant frequencies and accompanying harmonics associated with a given fundamental pitch, in Fedor Tau's performance of *Xöömei*, the fundamental pitch is suppressed and only barely audible. Instead, it is the harmonics generated by this barely discernible harmonic which are amplified and emphasized. While Tuvan singers use specific vowel shapes and articulatory methods to generate specific sounds, the formants responsible for vowel disambiguation are not static but shifted, merged, and otherwise manipulated in order to highlight and amplify specific order positions within the harmonic series. Instead, two things that do change appreciably are the succession of selectively intensified harmonics which form the primary line, as well as the comparative levels of sonic intensification apparent between phrase segments 4 and 6 in contrast to that of phrase segment 5.

COGAN'S THEORY OF SONIC OPPOSITIONS

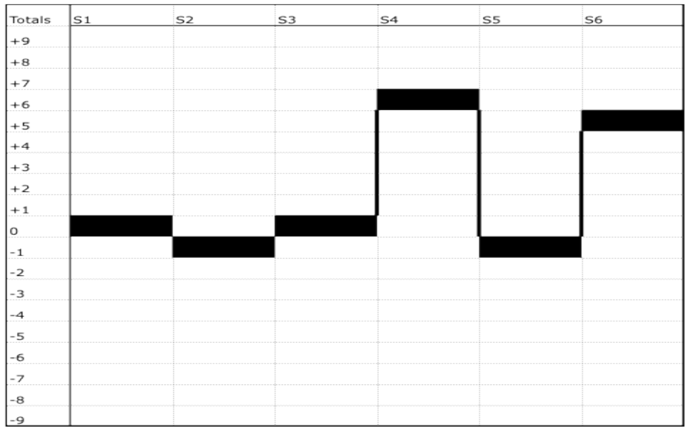
In his seminal text, *New Images of Musical Sound* (1984), Robert Cogan innovated the analysis of musical spectra with a methodology fabricated in alignment with structuralist currents of the time as initially espoused in Levi-Straus's groundbreaking research in mythology, characterized through the lens of binary oppositions and signification. Cogan's use of sonic oppositions or *features* was adapted directly from linguist Roman Jakobson's pioneering research in phonology. Collectively, these oppositions are understood to define a feature space with instantiations of various oppositional pairings resulting in the formation of feature-complexes or vectors. Cogan's approach begins with parsing the timbral surface into a succession of discrete segments which he then describes as a combination of oppositions, each of which is assigned a corresponding contextual value of either '+' positive, '-'negative, +/- mixed, or 'o' neutral. These results are inventoried in scorecard fashion and the total values obtained are understood as representative of the general level of spectral intensity characteristic of a given segment. Cogan considers the changing intensity levels expressed between successive segments to be descriptive of spectral morphology.

Example 10 demonstrates a Cogan style analysis where 13 sonic oppositions are employed to characterize the spectral organization of the Fedor Tau's performance. The left-side vertical column indicates the 13 oppositional pairings whereas the six phrase segments are identified along the top horizontal row. The southwest corner of the chart identifies the various intensity levels as positive, negative, neutral, and mixed and the actual values for each appear in the central portion of chart. The total level of spectral intensity for each phrase segment appears highlighted in yellow along the bottom.

negative (-)	positive (+)	1	2	3	4	5	6
grave	acute	-/+	-/+	-/+	-/+	-/+	-/+
centered	extreme	+	-/+	+	+	-/+	-/+
narrow	wide	+	+	+	+	+	+
compact	diffuse	-	-	-	-	-	-
non-spaced	spaced	+	+	+	+	+	+
sparse	rich	-	-	-	+	-	+
soft	loud	-	-	-	+	-	+
level	oblique	-	-	-	-	-	-
steady	wavering	+	+	+	+	+	+
no-attack	attack	+	+	+	+	+	+
sustained	clipped	-	-	-	-	-	-
beatless	beating	+	+	+	+	+	+
slow beats	fast beats	-	-	-	+	-	+
neutral (0)		-	-	-	-	-	-
negative (-)		6	6	6	3	6	3
mixed (-/+)		1	2	1	1	2	2
positive (+)		6	5	6	1	5	8
Total (-)		7	8	7	4	8	5
Total (+)		7	7	7	10	7	10
Final total		0	-1	0	+6	-1	+5

Example 10. Cogan's table of binary oppositions.

This information is then represented visually in Cogan's spectral intensity chart as shown in Example 11. The chart reveals that spectral intensity hovers around 0 (zero) over the first three phrase segments. In segment 4 we observe a sharp increase in intensity level from 0 to +6 apparent between segments 3 and 4, followed by a subsequent decrease in intensity between segments 4 and 5 where levels shift from +6 to -1 respectively. This, in turn, is followed by yet another marked intensification between segments 5 and 6 to conclude the performance.



Example 11. Overall sonic intensity fluctuations (segments 1–6).

The use of sonic oppositions to examine sounds and plot changing levels of spectral intensity is a convenient tool with which to characterize, inventory, and communicate our perceptual experiences and the analytical intuitions they give rise to. While Cogan's method indicates general progressions in spectral intensity between larger segmental units of formal design, my desire is to explore in detail the interiors of these segments, and to model not only the organization of sounding harmonics within each spectral segment but their micro spectral constituents and how they change over time.

Along the way, several methodological extensions and refinements will be made to Cogan's initial theoretical formulations in order to accommodate recent advances in computer software technologies, which make much more nuanced and detailed consideration of timbre and other features of sonic design possible. These include: (1) reducing the number of binary oppositions to develop fewer, more efficient categories; (2) replacing the use of binary values (+/-) with actual measurements & values; and (3) shifting focus and interpretation away from the visual inspection of spectrograms to an approach at least in part informed by pertinent perceptual and cognitive considerations, and include application of some mechanism for auditory 'screening' in order to verify the aural salience of visual data.

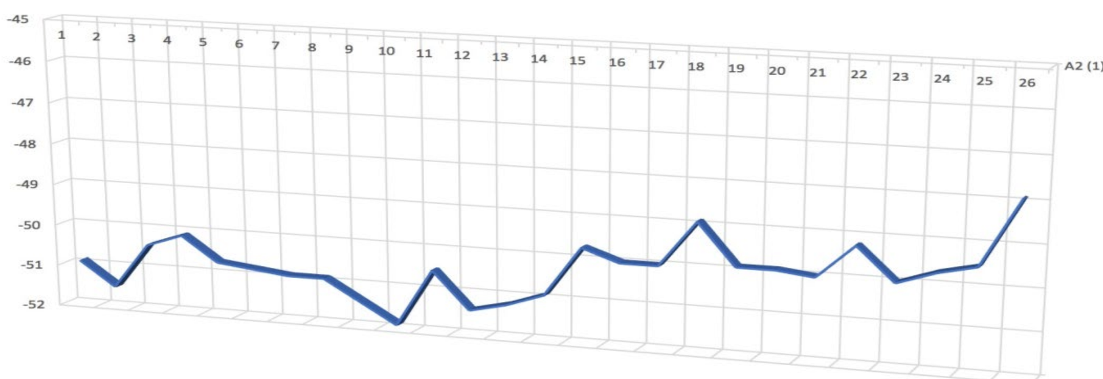
In this model of spectral morphology, we consider the following analytical parameters: spectral density (cardinality), spectral compass (i.e., height) as indicated by the number of semitones between the fundamental and the uppermost harmonic, intensity fluctuations of individual harmonics between successive vertical spectral sets, and changes in the respective speed and width of vibrato. While these represent only a cluster of numerous other contributing factors involved in our perception and experience of spectral morphology, these four analytical parameters in particular are distinguished as important constituents that are universal to sonic design and morphology in all styles and traditions of unaccompanied vocal music, and therefore they provide a foundation on which more sophisticated future studies can build.

Technological advances over the past two decades, however, have now made accessible various tools and resources through which the elements and features descripted in Cogan's inventory of sonic oppositions can be evaluated and measured in explicit terms and represented using specific values, thus facilitating a much more nuanced and detailed perspective of organization on the timbral surface. Consideration of set cardinality subsumes Cogan's binary opposition 'sparse-rich'. Similarly, measuring the respective speed and width of vibrato renders the oppositions 'slow beats – fast beats'; 'beatless – beating'; and 'steady – wavering' obsolete. Consideration of the respective spectral height of the set includes the oppositions 'narrow – wide' whereas implicit in the analytical notation for the spectral set the following oppositions are included: observation of the respective order positions designated in the spectral set notation reveals the degree of 'non-spaced – spaced' oppositions as well as the 'compact – diffuse' oppositions. Whether the spectrum is 'grave – acute' and 'centered – extreme' is indicated by the pitch class and respective octave registrations included in the spectral set data.

Thus far, we have delineated a series of vertical spectral segmentations which represent vertical slices within the spectrographic image; established a basic threshold for salience akin to a selected level of magnification in which we view the timbral surface; and gleaned those harmonics whose intensity values are greater than this perceptual boundary and characterized their organization as spectral sets. Spectral sets afford a profile or cross-section of those features which inform the sonic design of a given spectral segment. Now let us shift perspective from consideration of set membership to how the constituent features within each set change over the course of the performance. The relative presence of each harmonic within each spectral set has been approximated using the average dBFS scale so as to create an approximate gauge of aural salience for each sounding harmonic. In order to distinguish and compare sets in terms of the particular harmonics contained therein, we can now index the relative intensity each member harmonic contributes to its corresponding spectral set and view their changes over the course of the performance. The spectral sets occurring in phrase segments four through six can be divided into two collections. The first is the lower spectrum; this collection remains largely invariant and includes order positions 1–5 of the harmonic series; The second collection contains the 6^{th}–12^{th} harmonics whose respective boundaries define the spatial ambitus corresponding to the Tuvan scale. So far, we have adopted a detailed, taxonomic approach to the analysis of tone color, one that examines the vertical organization of harmonics above a common fundamental and organizes these into sets. The focus is on aspects of spectral morphology and developing ways in which we can illustrate how the contents of each spectral set change over time.

The drone

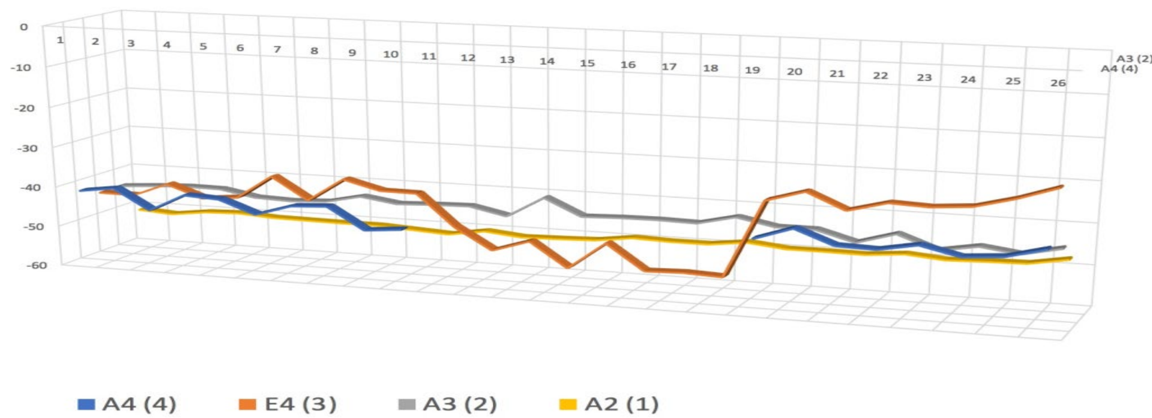
In Example 12 we observe a profile of changing intensity values associated with the fundamental pitch A2 across phrase segments 4–6. The left side column indicates the peak intensity dBFS values whereas the top horizontal row indicates the 26 spectral sets contained in the excerpt. The chart reveals that the intensity of the fundamental pitch remains quite consistent. In general, we observe a slight decrease in intensity values across spectral sets 1–10, beginning at -51db in spectral set 1 and after a slight increase to -50 db in spectral set 4, recessive fluctuations continue downward to -52 dBFS at spectral set 10, which coincides with the end of phrase segment 4. Between spectral sets 10–26 we observe progressive increases in intensity fluctuation for the remainder of the performance. In phrase segment 5, which includes spectral sets 11–18, we note incremental increases as dBFS values shift upwards from -52 to -49. Phrase segment 3 which includes spectral sets 19–26 presents a different intensity contour involving a slight decrease in levels at the onset of phrase segment 5 and making its way to the loudest value -48 at the end.



Example 12. The drone pitch and fundamental (f0).

The lower strata of harmonics (f1 – f5)

Undoubtedly, the most salient timbral transformation throughout the course of the performance occurs in with the large scale timbral intensification of phrase segments four (sets 1–10) and six (sets 19–26), in contrast to phrase segment five (sets 11–18) which represents the 'normal' state of affairs for phrase segments one through three as well. As observed, while there is some variance in intensity level for the fundamental pitch, the primary contributions are created by the lower order position harmonics H2–H3 as observed in Example 13.

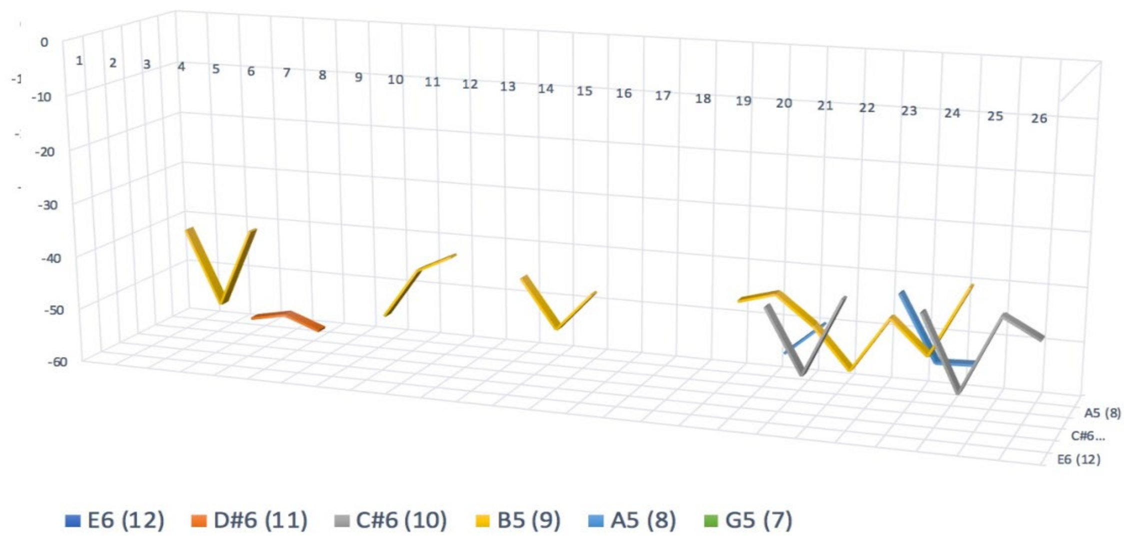


Example 13. The lower harmonic spectrum (f0 – f5).

We observe that the fundamental A2 appears as a more or less straight line resulting from the relatively small degree of variance involved at 4 decibels. The second harmonic A3 represented by the gray line also displays low levels of variance in phrase segments 4 and 6 represented by spectral sets 1–10 and 11–18 respectively. Notice that it is in phrase segment 4 that we observe marked, perceptual increases in intensity. The third and fourth harmonics (E4 and A4) carry the bulk of spectral energy and it is through coordinated shifts in their respective intensity levels that the striking textural contrasts evident between the three phrase segments are manifest. For instance, note that the fourth harmonic A4, while prominent in phrase segments 4 and 6, is altogether absent in phrase segment 5. In addition, observe that the loudest harmonic in the excerpt (E4) is coordinated with intensity shifts in the remaining voices to achieve the striking timbral contrasts evident between the three phrase segments. In phrase segments 4 and 6, the third harmonic E4 dominates the spectral texture whereas in phrase segment 4 it becomes the weakest sounding harmonic.

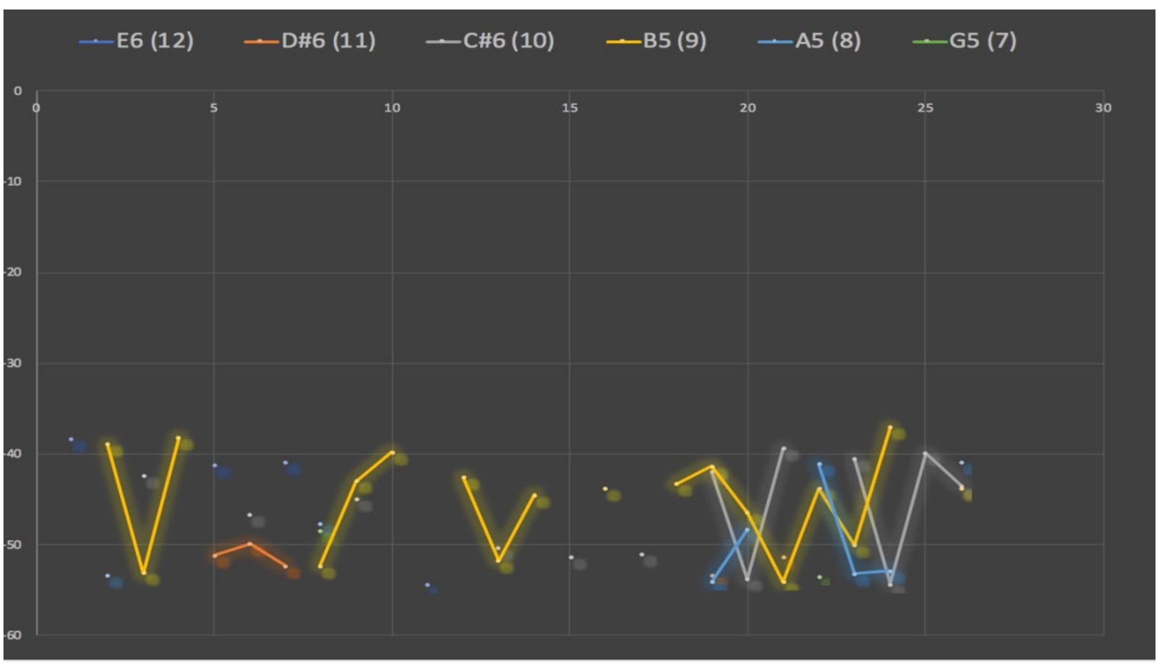
The upper strata of harmonics (f6 – f12)

Example 14 demonstrates corresponding intensity fluctuations for the upper spectrum voices that unfold the primary line or 'melody'. Note that several harmonics are apparently absent from the chart; this has to do with the type of visual display I have selected, namely 3D line segments.

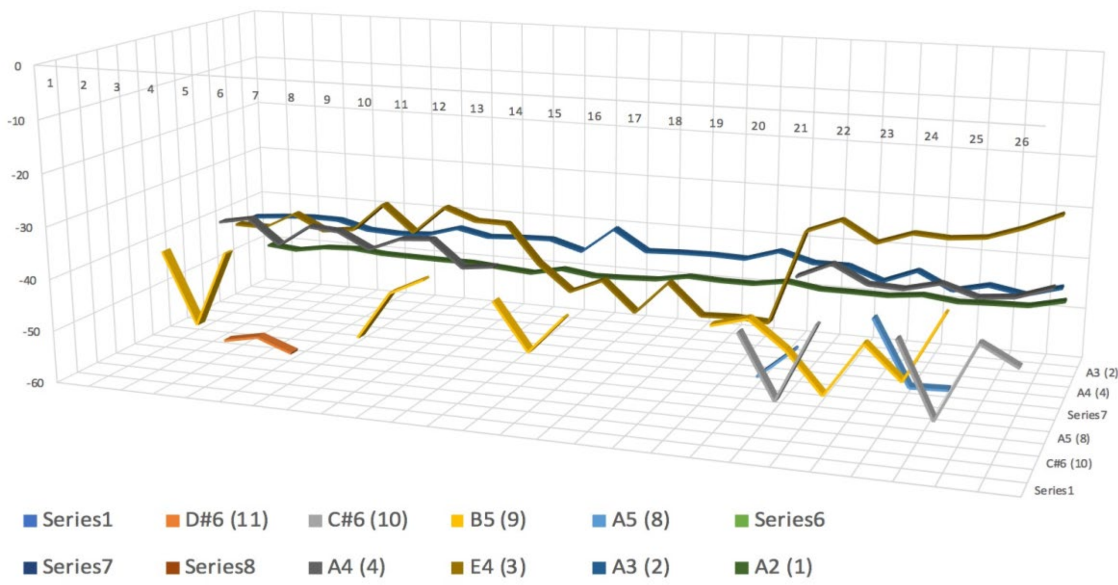


Example 14. The upper strata harmonics (f6–f12)

In order to reveal instances where a particular harmonic is sounded only once we can use a 'scatter graph' to fill in the missing pitches and values as shown in Example 15 below.



Example 15: Graph with points indicating harmonics that are sounded only once.



Example 16. The combined harmonic spectrum (f0–f12).

Spectral set cardinality and set height

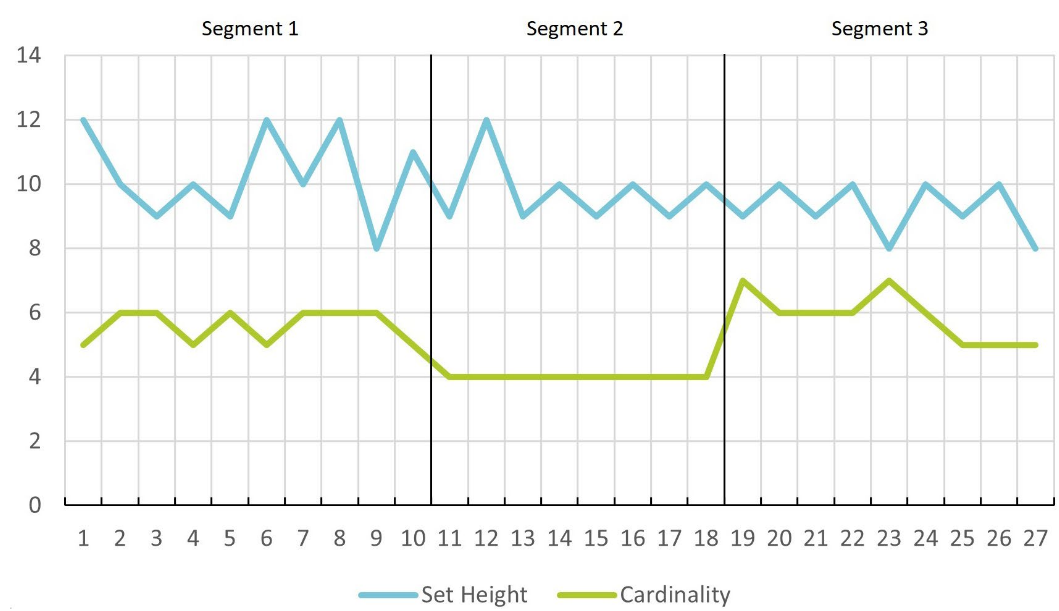
In addition to mapping the changing intensity values of harmonics, other important considerations involving spectral morphology concern the number of perceptible harmonics contained in each set. Cumulative increases or decreases in set cardinality represent progressive or increasing fluctuations in overall intensity values, while the reverse represents decreasing, or recessive intensity fluctuations.

Mapping successive changes in set cardinality across the performance reveals the organization of an important shaping force on spectral morphology. In general, fundamental pitches with lower frequencies generate more harmonics within the range of human audition than higher pitched fundamentals. Moreover, the loudest harmonics are found at the bottom of the harmonic series with each successive ascending step becoming increasingly softer. As characterized in the familiar source-filter model of vocal production, as the cardinality of the set increases or decreases, it does so in alignment with the respective ordering of harmonics within the harmonic series. For instance, a spectral set with a cardinality of 7 harmonics typically would include the harmonics in order positions 1–7, and so on. Consequently, there is typically a direct correlation between the density of a set and the respective height of a spectral set.

However, due to the Tuvan performers' ability to shift and merge formants in order to spotlight and intensify specific harmonics while dampening others, the aforementioned considerations are no longer in play. Instead, as observed, the spectral sets in Xöömei contain numerous empty order positions within the harmonic series.

Cogan would characterize these types of spectral segments as being 'spaced' as opposed to 'non-spaced', whereas the respective cardinality of the set would indicate whether the set is 'sparse' or 'rich'. As the number of empty positions within a given spectral set increases, the sound becoming increasingly hollow. Example 17 demonstrates the changes in set cardinality across in phrase segments 4, 5, and 6. Observe that the two phrase segments subject to general intensification, phrase segments 4 and 5, have average set cardinalities of 6 and 7 harmonics respectively whereas phrase segment 6 displays an average of only four harmonics per spectral set. The sense of general intensification resulting from increased loudness of the second and especially the third harmonic across phrase segments 4 and 6 at large.

Whereas the cardinality of the spectral set can be visualized as equivalent to its density and weight, spectral compass refers to the height of the spectral set as measured from the lowest sounding frequency (i.e., the fundamental) to the highest sound harmonic frequency in the spectral set. While we don't hear perceive the harmonics within a spectral set on an individual basis, we do, however, perceive the fundamental pitch (A2) as well as the selectively intensified harmonics which form the primary line and which are also always positioned as the upper boundary of each set, making them particularly prominent. Consequently, we do perceive the respective expansions and constrictions of the musical space whose upper and lower boundaries are defined by the selectively intensified harmonic and its corresponding fundamental. Example 17 demonstrates the changes in set cardinality as they occur in phrase segments 4, 5, and 6 as indicated by the blue line segment whereas corresponding changes in the spectral height of each set appear in red line segments.



Example 17. Changes in set cardinality (green) and set height (blue).

Concluding discussion

The analytical tools and strategies developed herein provide a method with which to explore and characterize the organization of timbral surfaces and other select features of sonic design in Tuvan, and by extension, other regional traditions of throat singing as well. As such, these tools and their accompanying methodology are intended to complement well-defined indigenous theories and descriptions of timbre, alongside the numerous aforementioned studies which have explored the physiological and biomechanical mechanisms involved in the production of overtone singing, and other ethnographic and anthropological studies which have characterized the relationships between music, culture, society, ecology, and belief systems. The results and insights afforded by etic forms of analysis such as those proposed herein can only acquire meaning and significance when viewed within a broader matrix of pertinent cultural and historical considerations. Analytical and theoretical assertions of the former might provide a complement and offer additional details and perspective when coupled alongside highly developed, emic forms of indigenous descriptions and method.

For those interested in exploring the organization of the sounds, their unique timbral features and acoustic properties, as well as the full array of sonic experiences made available therein when we listen using various analytical perspectives or 'angles of hearing', this method offers a systematic way in which to inventory, characterize, and compare the diversity of sonic and timbral contexts available in Tuvan throat singing with archival accuracy and descriptive precision employing free, and easy to use computer software applications available at no cost.

The use of computer-generated visualizations combined with various types of automation available enables a much more detailed and nuanced description of timbral components: the drone (fundamental); the selectively intensified harmonics; and accompanying associated harmonics. It also sheds light on how they merge, interact, and cofunction in the formation of the timbral surface. By doing so, it offers an extension and refinement to Cogan's previous use of binary oppositions to model and interpret similar considerations regarding sonic design and spectral morphology, the implementation of which was demonstrated in the expository analysis featuring an excerpt from Fedor Tau's *Xöömei*. While not a perceptual or cognitive study, such conditions remained a primary consideration when attempting to devise a methodology that has some phenomenological foundation and is supported at least in part by pertinent perceptual and cognitive considerations.

REFERENCES

Adachi S, Yamada M. 1999. “An acoustical study of sound production in biphonic singing, Xöömij.” In *The Journal of the Acoustical Society of America* 105: 2920–2932.

Aksenov, Alexei N. 1973. “Tuvin Folk Music.” In *Asian Music* 4(2): 7–18.

Alekseev, Eduard and Kirgiz, Zoya and Levin, Ted. 1990. *Tuva Voices from the Center of Asia*. Smithsonian/Folkways, Rec. 1987, CD.

Beahrs, Robert. 2017. “Nomads in the Global Soundscape: Negotiating Aesthetics in Post-Soviet Tuva’s Traditional Music Productions.” In *New Research of Tuva* 2: 59–10. https://doi.org/10.25178/nit.2017.2.3.

_______. 2014. *Post-Soviet Tuvan Throat Singing (Xöömei) and the Circulation of Nomadic Sensibility*. Dissertation, University of California, Berkeley.

Barras, Marie-Cecile, and Gouiffes, Anne-Marie. 2008. “The Reception of Overtone Singing by Uninformed Listeners”. *Journal of Interdisciplinary Music Studies* 162(2): 59–70. http://musicstudies.org/wp-content/uploads/2017/01/Barras_JIMS_0821204.pdf.

Belic, Roko (director). 2000. *Genghis Blues*. DVD, Wadi Rum Films.

Bergevin, Christopher. 2020. *Overtone Focusing in Biphonic Tuvan Throat Singing*. Dryad, Dataset. https://doi.org/10.5061/dryad.cvdncjt14.

Black, Jocelyn. 2018. Overtone Singing: History, Development, and Influence in Contemporary Music. Master’s Thesis, California State University, Monterey Bay. https://digitalcommons.csumb.edu/caps_thes_all/402/.

Bloothooft, Gerrit, Eldrid Bringmann, Marieke van Cappellen , Jolanda van Luipen, and Koen Thomassen. 1992. “Acoustics and perception of overtone singing”. In *The Journal of the Acoustical Society of America.* 92 (4): 1827–36. https://doi.org/10.1121/1.403839.

Bubley, Jean. *Alash – About Tuvan Throat Singing*. https://www.alashensemble.com/about_tts.htm [09.08.2023].

Chimiza, Lamazhaa. 2018. “Indigenous Methodology and Research of Tuvan Culture”. *Sribica*, 2018, 68–82. https://doi.org/10.3167/sib.2018.170307.

Cogan, Robert. 1976. *Sonic Design*. Cambridge MA: Publication Contact International.

_______. 1984. *New Images of Musical Sound*. Cambridge MA: Harvard University Press.

_______. 1998. *Music Seen, Music Heard*. Cambridge MA: Publication Contact.

Cook, Nicholas, and Daniel Leech-Wilkinson. 2009. *A musicologist's guide to Sonic Visualiser*. London: Royal Holloway University of London, The AHRC Research

Centre for the History and Analysis of Recorded Music. https://www.charm.rhul.ac.uk/analysing/p9_1.html.

Edgerton, Michael. Diane Bless, Susan Thibeault, Margaret Fagerholm, and Brad Story. 1999. "The acoustic analysis of reinforced harmonics." In *The Journal of the Acoustical Society of America* 105: 1329. https://doi.org/10.1121/1.426220.

Glenfield, Alexander. 2003. "The Pearl of Tuva: Authenticity and Tuvan Khorekteer (Throat Singing)", *Music Cultures* 30. https://journals.lib.unb.ca/index.php/MC/article/view/21614.

Grawunder, Sven. 2009. *On the Physiology of Voice Production in South-Siberian Throat Singing: Analysis of Acoustic and Electrophysiological Evidences*. Berlin: Frank & Timme.

Karelina, Eikaterina. 2018. "The Problematic Aspects of Cultural Policy in Modern Tuva". In *Journal of the Siberian Federal University. Humanities and Social Sciences* 2: 218–226. http://journal.sfu-kras.ru/en/article/70369.

Kob, Malte. 2004. "Analysis and modelling of overtone singing in the sygyt style." In *Applied Acoustics* 65: 1249–1259.

Matrenitsky, Vladislav, and Harris Friedman. 2012. "Transpersonal Effects of Exposure to Shamanic Use of Xöömei." In *International Journal of Transpersonal Studies* 31(2): 111–117. https://doi.org/10.24972/ijts.2012.31.2.111.

O'Toole, Caitlin. 2015. *Tuvan Throat Singing: The Globalization of the Tuvan Spirit*. Unpublished paper, Georgia State University. https://scholarworks.gsu.edu/cgi/viewcontent.cgi?article=1549&context=gsurc.

Levin, Theodore, Eduard Alekseyev, and Zoya Kyrgys. 1990. *Tuva: Voices from the Center of Asia*. Compact disc and booklet notes. Smithsonian Folkways Recordings.

Levin, Theodore, and Michael Edgerton. 1999. "The Throat Singers of Tuva". *Scientific American.* 281 (3): 80–7. https://www.scientificamerican.com/article/the-throat-singers-of-tuv/.

Levin, Theodore Craig, and Valentina Süzükei. 2006. *Where Rivers and Mountains Sing: Sound, Music, and Nomadism in Tuva and Beyond*. Bloomington: Indiana University Press.

Levin, Theodore Craig, and Valentina Süzükei. 2016. "Timbre-Centered Listening in the Soundscape of Tuva." In Dolan, Emily, and Alexander Rehding (eds.): *The Oxford Handbook of Timbre*. Oxford University Press, 205–228. https://doi.org/10.1093/oxfordhb/9780190637224.013.15.

Ruiz, Michael J., and David Wilken. 2018."Tuvan Throat Singing and Harmonics," *Physics Education* 53(3). https://doi.org/10.1088/1361-6552/aaa921.

Shuster, Lawrence Beaumont and Yannick Wey. 2022. "Mapping Timbral Surfaces in Alpine Yodeling: Towards A New Theory of Tone Color for Unaccompanied Vocal Music." In Shuster, Lawrence, Somangshu Mukherji, and Noe Dinnerstein (eds.): *Trends in World Music Analysis*. New York: Routledge, 159–195.

Shuster, Lawrence Beaumont, and Yannick Wey. 2021. "Mapping Vowel Color and Morphology: A Cross-Cultural Analysis of Vocal Timbres in Four Yodeling Traditions." In *AAWM Music and Nature* 1: 99–132. https://doi.org/10.5281/zenodo.5607754.

Süzükei, Valentina Yurievna. 2007. *Muzykal' naia kul' tura tuvy v 20 stoletii* [The musical culture of Tuva in the twentieth century]. Moscow: Kompozitor [in Russian, with English abstract].

Süzükei, Valentina Yurievna. 1993. *Burdonno-obertonovaia osnova traditsionnogo instrumental' nogo muzitsirovaniia tuvintsev* [The drone-overtone basis of traditional instrumental musicianship of the Tuvans]. Kyzyl: Tuvinskii nauchno-issledovatel'skii institut iazyka, literatury i istorii respublikanskii tsentr traditsionnoi kul'tury.

Süzükei, Valentina Yurievna. 2010. *Problema kontseptual' nogo edinstva teorii i praktiki (na primere tuvinskoi instrumental' noi kul' tury)* [The problem of the conceptual unity of theory and practice (based on the example of Tuvan instrumental culture)]. Kyzyl: Tyvapoligraf.

Süzükei, Valentina Yurievna. 2021. *Epistemological Problems in Russian Ethnomusicology*. Unpublished paper.

Süzükei, Valentina Yurievna. 2010. "Tuvan Music and the Twentieth Century." In *Journal of Siberian Federal University* 3(1): 80–93. https://cyberleninka.ru/article/n/tuvan-music-and-the-twentieth-century

Tongeren, Mark van. 1995. "A Tuvan Perspective on Throat Singing". In *Oideion* 2: 293–312.

_______. 2004. *Overtone Singing: Physics and Metaphysics of Harmonics in East and West*. Amsterdam: Fusica.

Computer Software Applications

Boersma, Paul and David Weenink. 2020. *Audacity* (Version 2.3.3) [computer software]. www.audacityteam.org.

PRAAT [computer software] (Version 6.0.46). Amsterdam: University of Amsterdam. www.fon.hum.uva.nl/praat/.

Maass, Bodo, Wolfgang Saus, Don Miller, and Krzysztof Jaros. 2020. Overtone Analyzer (Version 5.4.2.5435) [computer software]. Sygyt Software. www.sygyt.com/en/.

Maass, Bodo, Wolfgang Saus, Don Miller, and Krzysztof Jaros. 2020. *VoceVista Video Pro* (Version 5.4.2.5435) [computer software]. Sygyt Software. www.sygyt.com.

Sonic Visualiser (Version 4.3) [computer software]. 2020. London: Centre for Digital Music, Queen Mary University. www.sonicvisualiser.org.