9 Mapping timbral surfaces in Alpine yodeling

New directions in the analysis of tone color for unaccompanied vocal music

Lawrence Beaumont Shuster and Yannick Wey

9.1 Introduction

Yodel is a way of singing often characterized by rapid alternations between chest voice and head voice, and a vocalization using syllables that lack lexical meanings.[1] While these singing techniques exist around the globe, use of the term "yodeling" (German: Jodeln) is sometimes limited to the Alpine region. While sometimes described as a national tradition (e.g., Swiss yodel, Austrian yodel, etc.), yodeling practices are not homogenous within national borders, and variances inside these countries are probably greater than those between them. Yodel, or the German word "Jodel," has become an umbrella term for many different traditions of vocal performance, each with its own local name; there is the Bernese "Jutz," the Central Swiss "Juiz," the Appenzell "Zäuerli" and "Rugguusseli," the "Dudler" in and around Vienna, and the "Johlar" in Vorarlberg. And this is not a complete list.

In the Swiss context, a terminological differentiation is made between yodel songs ("Jodellieder") and natural yodel ("Naturjodel").[2] A piece of natural yodel, often comprised of two-to-three melodic segments and sung by one-to-three lead singers and an accompanying choir, is sung entirely on syllables that have no lexical significance. Natural yodel is primarily an oral tradition, though transcriptions of most melodies do exist. Yodel songs, on the other hand, are usually composed and written in standard staff notation.

The syllables are sung in frequent alternation between chest and head voice. This change of register produces throat beats that are made audible depending on the aesthetics of the regional vocal tradition. The employed syllables are not arbitrary: despite the absence of explicit rules, the choice of syllables depends on regionally established aesthetics and shows patterns of changes between voice registers, small or large intervals, and pitch in general.

As mentioned earlier, yodel has been the subject of musical analysis in a small number of studies, for instance: the difference between classical singing and yodeling (Luchsinger and Arnold 1949), form and text of yodel songs (Hänggi 2011), and the relationship of yodeling to alphorn music (Ammann et al. 2019).

In order to demonstrate the general utility of the analytical tools and strategies developed herein, we have selected an excerpt from the classic 1965 performance

DOI: 10.4324/9781003033080-10

of *Höch Obe*—a well-known natural yodel melody from Central Switzerland—as our primary analytical sample. This excerpt is provided in the file titled Audio 9.1 (**<URL HERE for associated eResources>**). The performance is by Ruedi Rymann, a renowned yodeler and folk singer from the Canton of Obwalden (Central Switzerland).[3] The recording has been published in various formats; in this study, we refer to the CD *Die Jodelarten der Schweiz* (the yodeling genres of Switzerland), a sampler portraying the diversity of yodel in various regions of Switzerland published in 2010 (Bachmann 2010). While *Höch obe* has been arranged for a soloist with choral accompaniment, our study is based on Rymann's solo performance.

9.2 Formal shape and design

Figure 9.1 displays a transcription of *Höch Obe* (Rymann 2010) using conventional Western notation. This transcription is prescriptive in the sense that it represents the recording as a notation that could be used for its performance. It does not account for pitch drift, ekmelic intervals, or rhythmic deviations. Notations

Figure 9.1 Transcription of *Höch Obe* (Rymann 2010) with vowels.

of this kind are regularly written by yodel composers and used by conductors and instructors as a memory aid. The syllables are given in German, and the vowels coincide with phonetic script. The performance consists of two phrases designated phrase A (mm 1–7), and phrase B (mm 15–18). The two phrases present stark contrasts in terms of meter and rhythm, phrase length, melodic contour, intervals, and motivic design. Phrase A contains seven measures and uses a 3/4 meter whereas phrase B contains four measures and has a 4/4 meter. Each phrase is repeated; following the repetition of phrase B, the phrase A is performed one additional time establishing an abbreviated ternary form ([A1 (A2) – B1 (B2) – A3].

The repetition of each phrase is literal with subtle discrepancies in terms of rhythm, intonation, and articulation. Junctures between successive phrases are defined by cadences in measures 7, 14, 18, 22, and 29. The medial cadence positioned at the exact center of phrase A (measure 3.5) defines two symmetrical subphrases organized in the form of a parallel period designated subphrase A1a (mm. 1.0–3.5) and subphrase A1b (mm. 3.5–7.0). In phrase B, there are three subphrases resulting in a tripartite design: subphrase B1a (mm. 15.0–16.1); subphrase B1b (mm. 16.2–17.1); and subphrase B1c (mm. 17.2–18.0).

9.3 Tonal design and linear reduction

Figure 9.2 illustrates a foreground linear reduction of phrases A1 and B1. The overall melodic contour is informed by the distribution and positioning of chord tones associated with the B♭ major triad, which chord tones occur at the beginning, middle, and end of each phrase and provide a scaffolding for melodic design. Phrase A, subphrase 1 begins with the unfolding of the B♭ tonic triad: the chordal fifth (F) unfolds down to the chordal third (D) via the passing tone (E) in

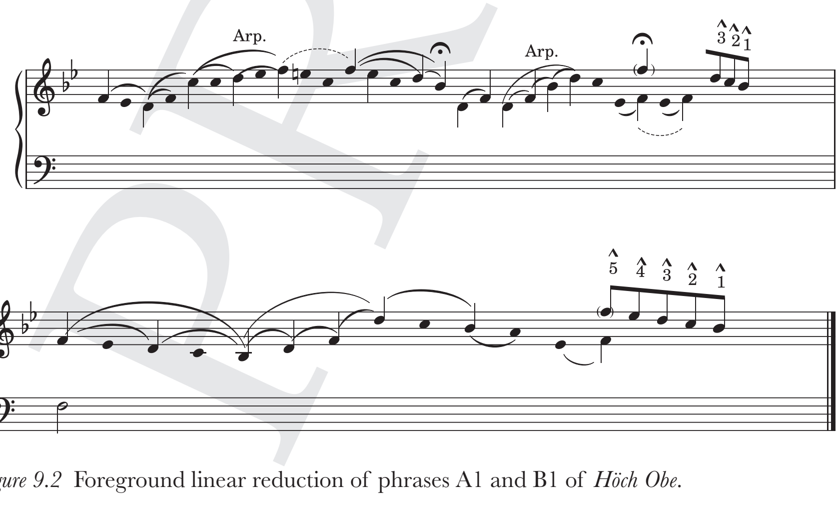


Figure 9.2 Foreground linear reduction of phrases A1 and B1 of *Höch Obe*.

measure 1. From this point we observe an ascending arpeggiation spanning D4 to D5 (measure 1) and continuing upwards to the registral apex (F5) at the downbeat of measure 2. The remainder of subphrase 1 features a prolongation of the pitch F5 with E5 initially appearing as a lower neighbor tone, and subsequently as a passing tone, enroute to D5 in measure $\hat{3}$; the arrival on the latter being amplified through the use of a double neighbor note configuration (E5 and C5) prior to the final move to the tonic B♭4 that completes the unfolding of the B♭ major triad.

Despite small contrasts in surface detail, we observe that the two subphrases comprising phrase A involve nearly identical tonal architectures. For example, both subphrases feature a prolongation of the chordal fifth (F). In subphrase 1, F5 is prolonged via the lower chromatic neighbor E5, the latter changing function to that of a passing tone at the end of subphrase 1 and linking together F5 to the chordal third (D5) in measure 3, before reaching the medial cadence on the tonic B♭4 in measure 4. Subphrase 2 begins with a strong sense of starting over: both subphrases begin with the same motivic content featuring an ascending arpeggiation of the B♭ major triad. Moreover, both feature the prolongation of F with a lower E natural neighbor tone as well—the registral displacement in the second iteration notwithstanding. Phrase endings in Alpine yodel are often articulated by use of a common cadential pattern which features a stepwise descent from scale degree $\hat{3}$ to the tonic ($\hat{3}$–$\hat{2}$–$\hat{1}$). We note that this pattern is interrupted at the end of subphrase 1 (m. 3) where it pauses on scale degree $\hat{2}$ before the entire process is repeated again, only coming to a full cadential close on the tonic as evidenced by the linear progression $\hat{3}$–$\hat{2}$–$\hat{1}$ that marks the end of phrase A.

The tonal design of phrase B is also organized around the unfolding of the tonic B♭ major triad. Similar to phrase A, the melodic ambitus features the juxtaposition of two distinct tessitura; one associated with the low-pitched chest voice; the other, the higher-pitched head voice. The beginning (measure 15) features an unfolding of the tonic arpeggio from the fifth (F) down to the third (D) and, finally, to the tonic (B♭, measure 16). Positioned in between this succession of chord tones, the addition of diatonic passing notes creates a five-note stepwise linear descent extending from scale degree $\hat{5}$ (F) to scale degree $\hat{1}$(B♭) whilst sounding over a dominant pedal (F) that is performed in hocket-like alternation with the descending melodic line. In measure 16, a second subphrase consisting of an ascending arpeggio of the tonic triad unfolds from the root (B♭) to the fifth (F). The large leap and stark juxtaposition of registers further helps to distinguish the structural juncture that defines the boundary between respective subphrases. Whereas the previous cadence at the end of phrase A employed a linear descent from scale degree $\hat{3}$ to the tonic ($\hat{3}$–$\hat{2}$–$\hat{1}$), the cadence at the end of phrase B employs a five-note linear descent extending from scale degree $\hat{5}$ to the tonic ($\hat{5}$–$\hat{4}$–$\hat{3}$–$\hat{2}$–$\hat{1}$). Both cadential figures are commonplace; their ubiquity transcending stylistic and regional boundaries throughout the greater Alpine region.

9.4 Tuning, intonation, and pitch drift

Data on pitch were retrieved from a TCIF (time-corrected instantaneous frequency) spectrograph, whereby the fundamental frequency for each note was marked at a midpoint where the note is sounding stable. A spreadsheet with pitch measurements is available as a supplement. Figure 9.3 displays the distribution of frequencies (one measurement per note) and the distribution of adjacent intervals. The melody is characterized by the frequent succession of small intervals, seconds, and thirds. Large jumps, which in yodeling especially mark the alternation between chest and head voice, remain comparatively rare. Nonetheless, these play a significant role in creating the wide ambitus of two octaves, which is enabled using both chest and head voice.

While the histogram of pitch (Figure 9.3, left side) can be distorted by pitch drift, this effect is mitigated in the distribution of neighboring intervals (right

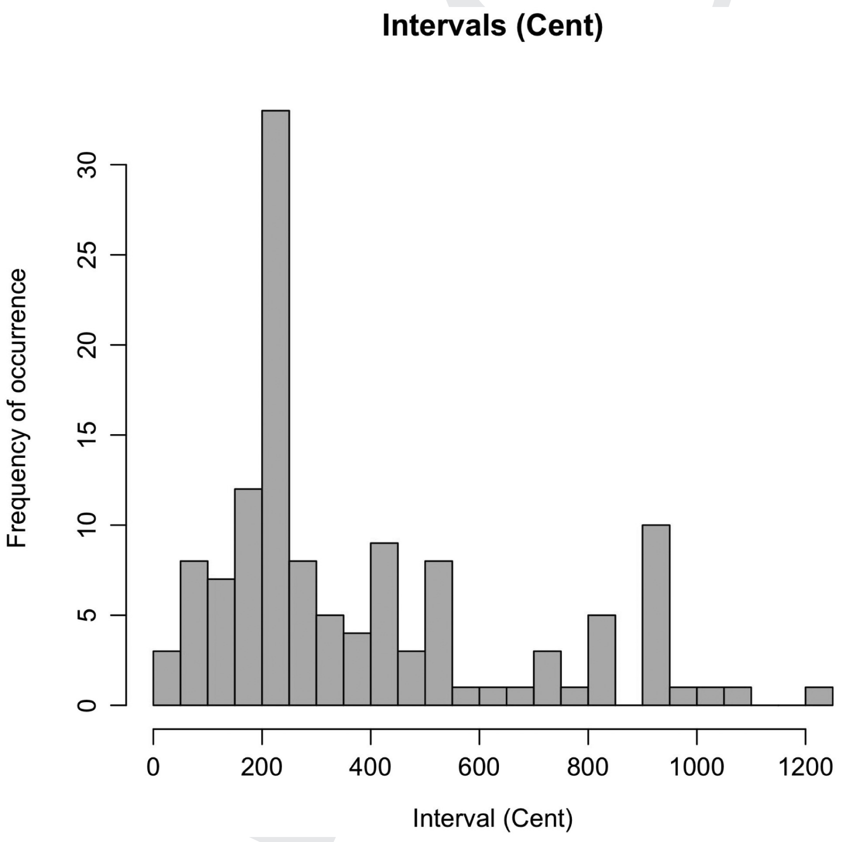


Figure 9.3 Distribution of frequencies and adjacent intervals.

side). Pitch drift is a known phenomenon in vocal performance (Mauch et al. 2014; Ambrazevičius 2014; Seaton et al. 2014). In yodeling, pitch drift can be construed either as a feature of performance practice or as an unintended phenomenon that occurs spontaneously. In multipart group performance, a certain stability of pitch and the resulting harmonic accord is preferred. In solo or duet singing, however, there is no compelling need for holding on to a given tuning. By tracing pitch drift, we aim to decipher whether a shift to another scale happens at a certain point or throughout the performance, and whether the drift is large enough to cross the lines between the pitch classes. Furthermore, the question arises as to whether drift occurs in a single direction only, or whether it oscillates in both directions. Figure 9.4 demonstrates pitch drift for the notes B♭, C, D, E♭, E, and F. The note A is omitted as it occurs only twice throughout the piece. Linear models of pitch drift for each note are available as supplementary data. From beginning to end, the drift

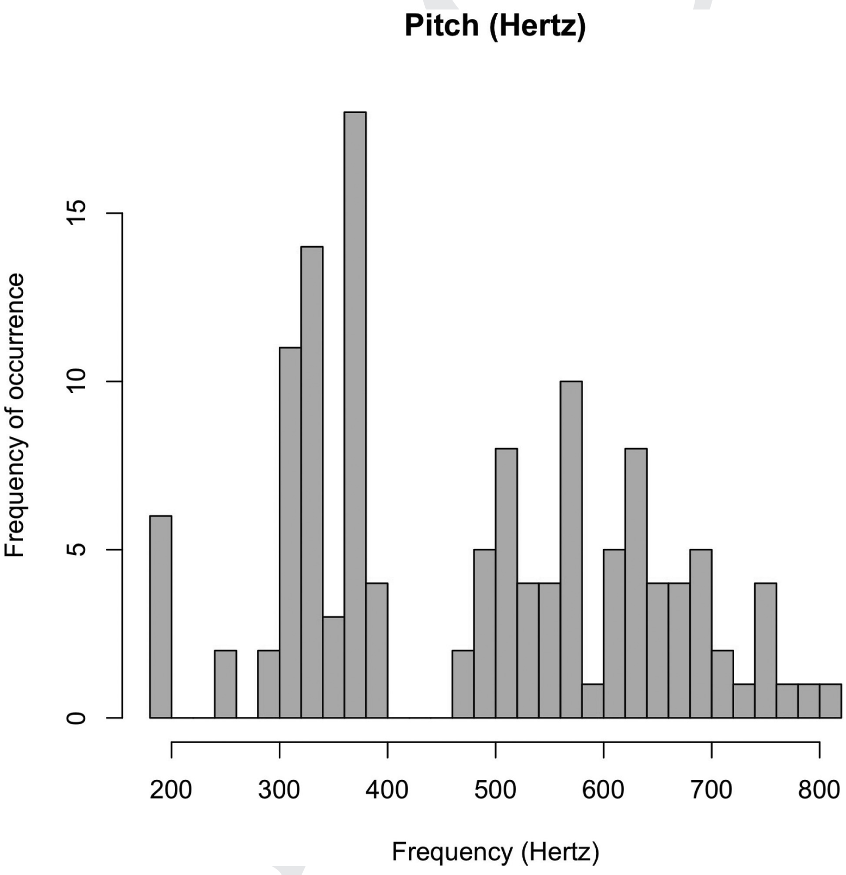


Figure 9.4 Pitch drift distance for degrees of the scale.

spans approximately 170 cents (more than three quarters of a tone), therefore reaching across different semitones.

Figure 9.4 reveals that pitch drift is both linear and one-directional. This comes as a surprise: when listening to and transcribing the piece by ear, a shift to a higher semitone was perceived over a designated, short period. We subscribe this discrepancy between the subjective perception and the measurements of acoustical data to the phenomenon of categorical perception (Goldstone and Hendrickson 2010: 69): Hearing is not linear but categorical, with the perception of pitch switching from one category to the next at a certain threshold point.

As pitch drift can be accurately described with linear models, a hypothetical, drift-corrected tonal scale can be created. The coefficient for each note is subtracted (multiplied for the rank-order of the note), then the frequencies are compared to the median values of the root (B♭) in order to receive a set of intervals, from whence the pitch classes emerge. Again, the note A was omitted. A pitch histogram shows discrete pitch classes, save for adjacent semitones which overlap. A problem, however, remained in the explanation of the disappearance of the note E, which should spike at 600 cents. Consulting the data, the frequencies seem to coincide largely with those of E♭ after the drift-correcting measure.[4] The intonation approximates equal-tempered tuning. To test this, we used the median value of the note B♭ as a baseline and calculated the distance of every note while correcting for drift based on the linear models. Subsequently, we ran a one-sample t-test. The result demonstrates that the present tonality is not significantly different from equal-tempered tuning (p=0.6875).

Figure 9.5 shows the five degrees of the drift-corrected tonal scale, which correspond to equal-tempered tuning. A few neutral intervals can be measured, but these measurements are rare enough not to change the overall proximity to equal temperament. The use of the harmonic scale is evident in part A; however, this was not mirrored by intonation. The restriction to the harmonic series is not uncommon in Central Swiss yodel (Leuthold 1981: 27; Ammann et al. 2019: 153), but, the prevalence of the augmented fourth degree in the present recording deserves special attention, as it concurs with the hypothesis that the intonation of harmonic 11 has been transferred from the alphorn to vocal performance. Colloquially known as "alphorn-fa," harmonic 11 is situated a quartertone between its equal-tempered nearest semitones. The exact ekmelic intonation of this note in yodel, however, remains very sparsely documented (e.g., by Ammann et al. 2019: 185). More often the phenomenon of a fourth augmented by a semitone can be observed, without being justified by underlying harmonics. One interpretation of this phenomenon is an adjustment of harmonic 11 to the equal-tempered intonation due to written transmission of transcribed melodies.

Vibrato is not regarded as a common feature of yodeling. Rather, long tones are sustained and the volume of the sound results from the abundance of overtones, rather than from the use of vibrato (Räss and Wigger 2010). In the present recording, however, vibrato appears when tones are sustained. To

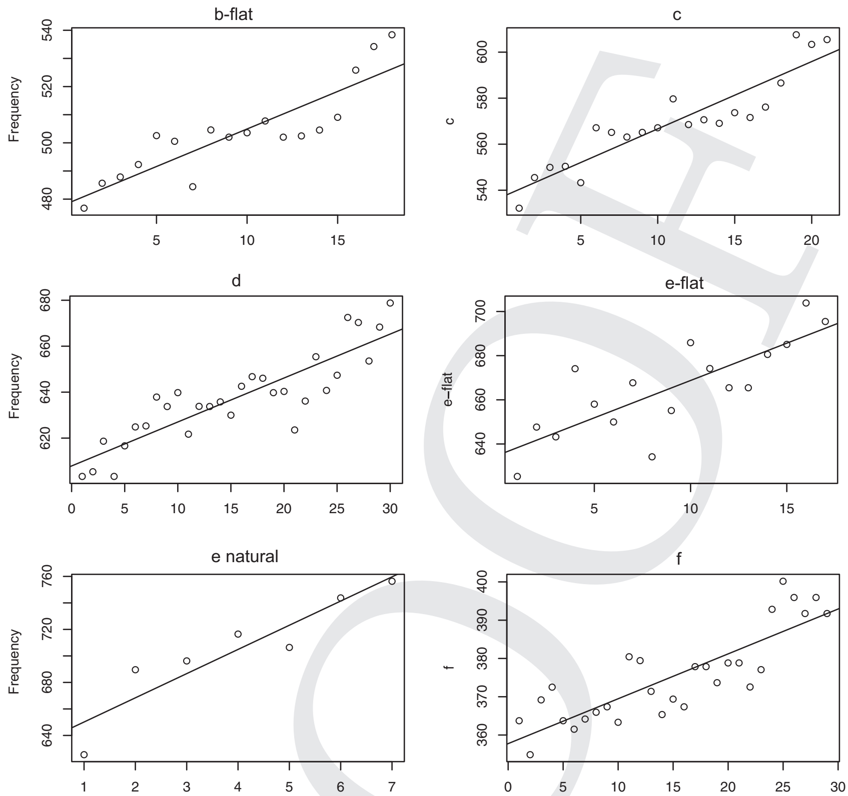


Figure 9.5 Drift-corrected tonal scale.

determine the speed of vibrato and its distribution, every "turning point" during a given note was stamped with a time marker. Five notes in segment A include more than ten of these points and are represented in the analysis. Time intervals between those markers are available in the supplementary data. Vibrato oscillates in mean rates of 0.09 s. The maximal intra-tone pitch interval caused by vibrato amounts to approximately one semitone (89–114 cent). The mean amplitude of vibrato is 70 cents, with a standard deviation of 18 cents. For comparison, in his study of vibrato in selections of *Orfeo*, performed by a trained opera singer, Kharuto (2005: 298) measured amplitudes of 32–93 cents. In a study of ten classical singers, Prame (1998: 619) reports mean amplitudes of 57–86 and maxima of 71–123 cents. As mentioned earlier, vibrato is not a commonly assumed feature of yodeling, yet in Rymann's case it resulted in even larger amplitudes than measured in the studies of Western art music just cited. Nevertheless, vibrato occurred only on sustained high notes, where it serves to achieve greater resonance in the very high register.

9.5 Mapping timbral surfaces

Whereas the previous discussion centered on issues of pitch and frequency, we now shift attention to tone color, or timbre. Of all parameters of musical design, timbre remains the least studied and least understood yet along with pitch and rhythm, timbre features at the forefront of our musical experience. The Acoustical Society of America defines timbre as "that multidimensional attribute of auditory sensation which enables a listener to judge that two non-identical sounds, similarly presented and having the same loudness, pitch, spatial location, and duration, are dissimilar." Timbre is determined both by the harmonic content of the sound, as well as the dynamic states in which these contents are manifest.[5] Harmonic content refers to the combination of fundamentals, overtones, and harmonics present in the sound, and the respective intensity of each of these elements in relation to the others.[6] Conversely, dynamic qualities of timbre include consideration of amplitude, vibrato, intonation, and features associated with the attack-decay envelope of a sound. For a very long time, the primary obstacle in the analysis of timbre was simply the lack of the necessary technologies capable of exploring the acoustic structure of a sound in sufficient detail. Today, the problem is how to reconcile the vast amounts of objective data generated through computer-assisted acoustic analysis with our subjective experience of sound, which is ultimately conditioned and informed through perceptual and cognitive processing.

In vocal music, the raw sound is produced by the periodic vibrations of the vocal cords resulting in the simultaneous generation of multiple harmonics sounding as integer multiples of the frequency associated with the fundamental pitch.[7] The higher the harmonic, the softer the vibration and consequently, the lower the volume. Theoretically, there is no limit to the number of potential harmonics that can be generated, but the lower the frequency of the vibrations, the more overtones produced are within our range of hearing than high frequency sounds. In order to characterize the sonic designs of Alpine yodeling performances and identify and inventory the diverse assortment of timbral contexts and experiences contained therein, we need a way to sort the individual components of sonic design in order to map their organization and the manner in which they interact and co-function in the formation of timbre. In order to accomplish this, we will engage a number of computer software technologies, and in particular, the use of spectrographic images, which enable us to tease apart and inspect the individual elements that inform sonic design.

9.6 Spectrographic analysis

The use of spectrographic analysis techniques in music was primarily innovated by Robert Cogan in the early 1980s (Cogan 1984). In light of developments over the nearly four decades since, spectrographs—which make possible the examination of the frequency, pitch, intonation, amplitude, and duration of sounds with a high degree of technical precision—provide us with our most powerful tool for examining timbral surfaces. There is also a downside to spectrography, given that it is capable of producing great quantities of information, much of which extends beyond the boundaries of human perception. Thus, while the spectrograph provides a powerful tool that can help guide, shape, and focus our listening experience, and which we can use to verify our analytical intuitions and demonstrate the phenomenological basis of our analytical assertions, the visualized data produced should be understood as a starting point for the analytical process, and not be confused as an accurate model of our perceptual experience.

Figure 9.6 demonstrates a narrow-band spectrograph of Rymann's performance of "*Höch Obe*" using the VoceVista Video Pro software set to a logarithmic display. In reading this graph, the following criteria apply: the horizontal axis indicates time, measured in milliseconds; the left-side vertical axis indicates frequency, measured in hertz. Within the main frame of the graph, the various overtones are displayed. The lowest sounding horizontal line usually represents the succession of fundamental pitches whose succession forms what listeners typically identify as the melody. Subsequent horizontal lines positioned above represent the combination of various harmonics and other overtones embedded on the timbral surface. The respective intensity of each element is indicated by the dBFS (decibel full-scale) appearing adjacent to the frequency scale located to the left of the graph.

9.7 Parsing the timbral surface: defining spectral segmentation

The initial step of the analytical method involves parsing the timbral surface into a succession of spectral segments. Each segment appears as a vertical slice of the available spectrum as demonstrated in the main display of the spectrograph in Figure 9.6. While the horizontal boundaries of each respective vertical spectral segment are coextensive with the duration of the fundamental pitches that form the melody, the vertical boundaries are defined by its lowest (the fundamental) and highest sounding frequencies. Each spectral segment contains: (a) the fundamental pitch, (b) all perceptible overtones, and (c) the corresponding vowel type. Please note that section A1 (mm. 1–7) contains a total of 30 spectral segments whereas section B1 (mm. 15–18) contains 20.

9.8 Measuring harmonics within spectral segments

Having defined the spectral segments that provide the basic units of analysis, we now explore ways in which to characterize their internal organization and discover correspondences between them. The next step is to filter out those features that play a role in our perceptual discriminations from those that do not. The relative intensity—or amplitude—of the individual harmonics contained in each spectral segment can be determined through fast fourier transform analysis and displayed as a function of frequency. Each fundamental pitch and its accompanying harmonics are measured in terms of their respective amplitudes using the Decibel Full-Scale.

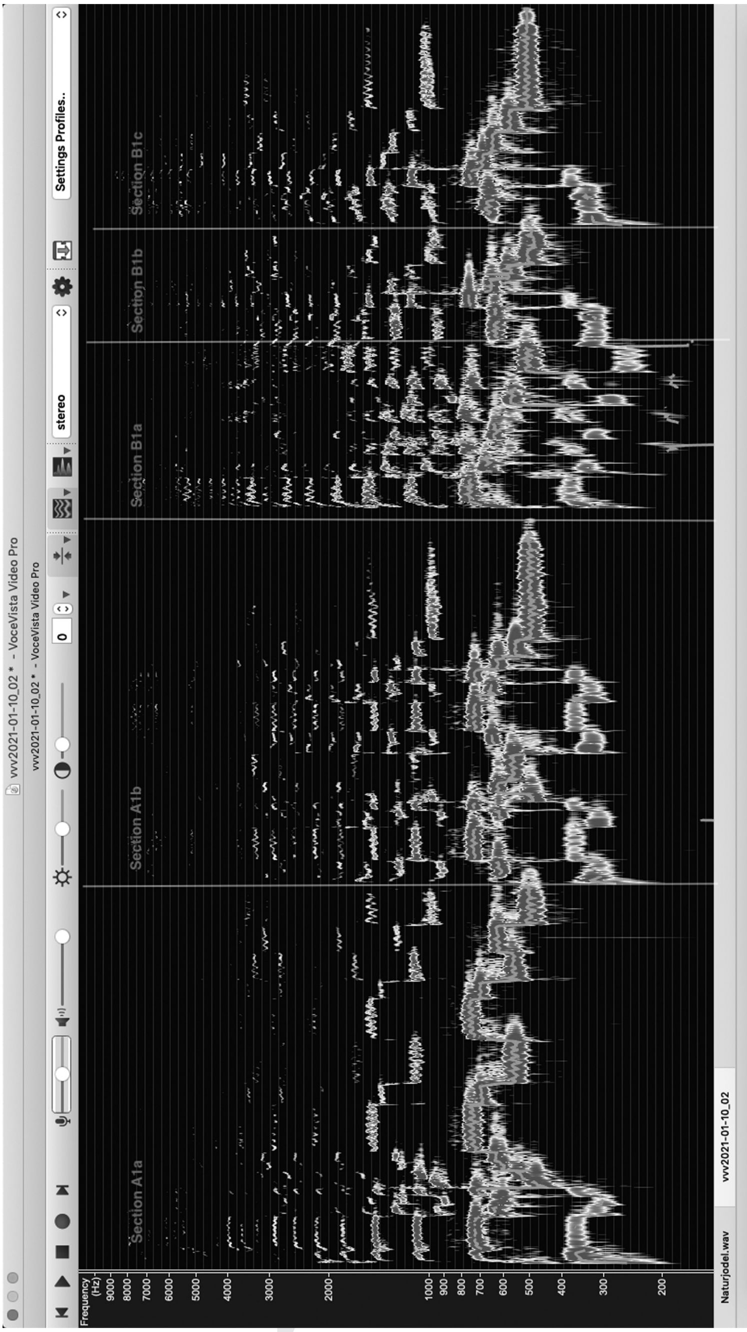


Figure 9.6 Narrow-band spectrograph of Rymann's performance of *Höch Obe*, sections A1 and B1.

Decibel Full-Scale, or dBFS, is a digital measurement of amplitude in which a value of zero is assigned to the loudest sound in a given signal and all other measurements are shown as negative integers in comparison. As such, dBFS values are context dependent. Once the user has defined a specific spectral segment, the software will reveal both the average and peak amplitudes, frequency, pitch class, and corresponding octave registration for each harmonic within the segment.

We note, however, that the acoustic measurements provided by the software are entirely objective. The sounds as seen in the spectrograph are not the same as those that we perceive once the acoustic stimulus has been conditioned as a result of auditory and cognitive processing. What's more, it is not simply a matter regarding the respective amplitude of the acoustic signal, given that our auditory discriminations and perceptions are not equal. As the equal-loudness curves and bark scales clearly indicate, our perceptual experience of some frequencies is actually louder than the acoustic data would suggest. And the reverse is also true. The frequency bandwidth between 1–4 khz is amplified, and sounds in this frequency range seem louder to our ears than their dBFS values would suggest.[8] Conversely, we hear sounds on either side of this bandwidth as weaker than their dBFS values would suggest.

From another perspective, spectrographs usually display extremely broad bandwidths of frequencies—sometimes upwards of 20 khz, yet for human ears, harmonics above 3,500 hz become increasingly difficult to perceive. This is due to the fact that as the respective order positions of the harmonic series become increasingly closer in frequency, they become grouped within the same critical frequency bandwidth and thus become indistinguishable.[9]

Figure 9.7 demonstrates a sample of the automated average dBFS values for each spectral segment in sections A1 and B1. The vertical spectral segment appears highlighted within the main display of the graph; the corresponding timeline marker indicated along the horizontal axis at the bottom of the example. To the right, the display provides the inventory of average dBFS values for each harmonic, its corresponding frequency as measured in Hz, and the respective pitch class, octave registration, and intonation for each harmonic. We note that in addition to those harmonics which form part of the harmonic series associated with the fundamental D♯5, there are many other sounding components as well. These are not included in the analysis, and must be the result of either noise, or other sounds blending into the specific segmentation.

In order to make the dBFS values in Figure 9.7 meaningful, we need some way to separate those features that play a direct role in our perceptual experience from those that do not. Given that dBFS values are context specific, we need to establish a general threshold for salience while at the same time keeping in mind the fact that these values are at best approximations from which to orient our analysis. And though we take perceptual and cognitive considerations strongly into account when assessing the dBFS data, it is important to remember that our results are shaped—but not constrained—by these considerations. Instead, the various perceptual and cognitive considerations provide a kind of filtering process that assists in identifying those features that play a prominent role in our

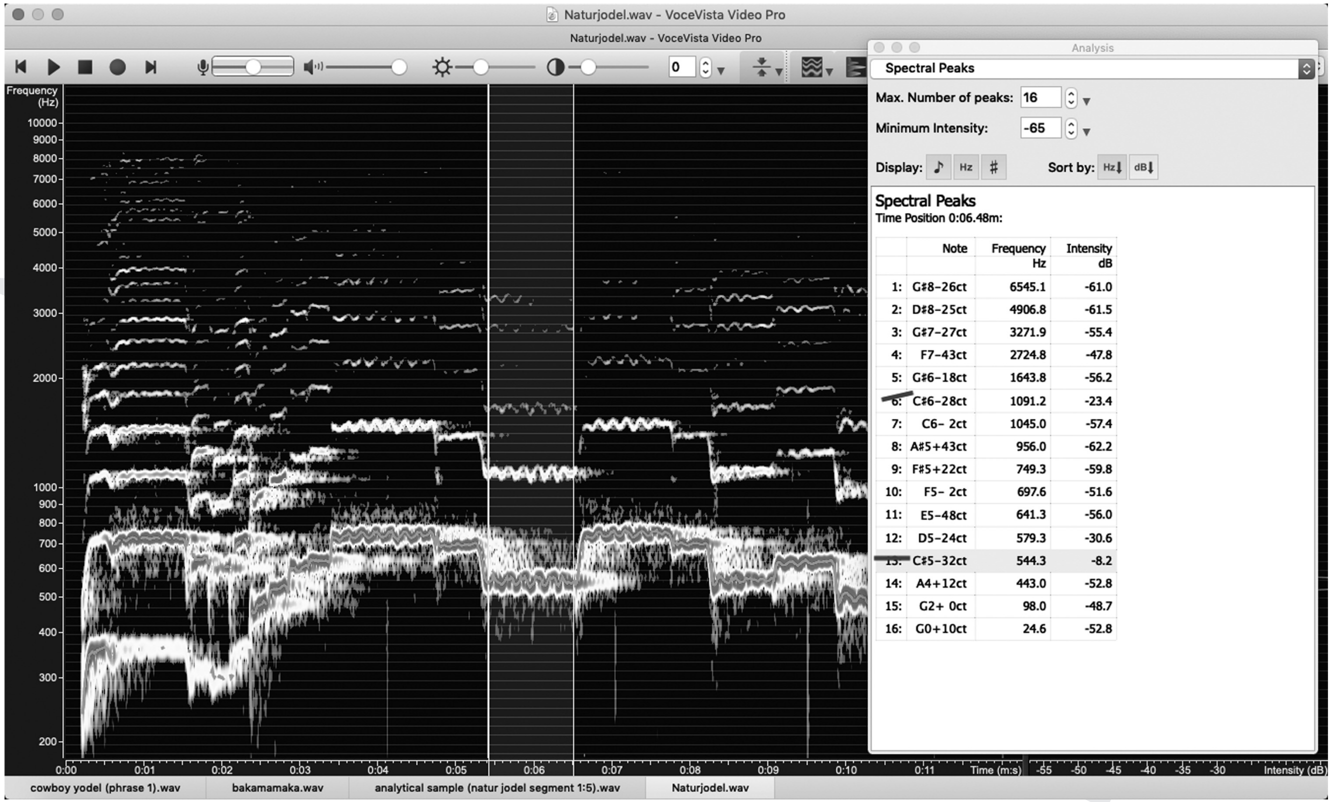


Figure 9.7 Spectral segmentations and measuring amplitude, in dBFS values.

experience and thus support our perceptual constructs and analytical assertions. These considerations preclude any type of formal quantification and limit our use of measurements in this study to asserting general organizational tendencies and approximations, some of which impact our perceptual experience directly. From others we can "learn" to hear. Still others remain concealed, inhabiting the deeper levels of structure and informing the timbral surface from a distance.

On the basis of our verification protocols, we have established −55 dBFS as the threshold of salience for this particular recording. Adjusting this salience threshold (−55 dBFS) is tantamount to adjusting the level of magnification in which the timbral surface is viewed. By using the filters available in the software, we are able to isolate and confirm individual frequency bandwidths associated with the delineation of pitch classes and registers of the harmonic series. And in order to account for the possibility of masking, the various harmonics are then assembled into groups: those whose aural presence is confirmed are kept; those deemed imperceptible are removed.[10]

9.9 Spectral sets and classes

Spectral sets and classes represent the basic building blocks of timbral surfaces. Within each spectral segment those harmonics whose average intensity is −55 dBFS or above and whose aural presence has been confirmed by the salience verification protocols established earlier, are considered members of the corresponding spectral set. Figure 9.8 demonstrates the inventory of spectral sets for sections A1 and B1. The analytical notation for spectral set is [x/y] where the x value indicates the respective pitch class, and the y value the respective octave registration for each harmonic. The succession of harmonics within each set conforms their distribution within the harmonic series with the fundamental pitch always appearing in the first order position. In rare instances where the respective harmonic assigned to an order position is absent, a dash (–) serves as a placeholder (see spectral sets 4, 10, and 15). More generally, a spectral class simply abstracts the spectral set by indicating only the respective order position of the harmonic within the harmonic series and not the specific pitch class nor the corresponding octave registration. Consideration of spectral class allows us to compare sonic contexts that have different fundamental pitches. Note that while only sounds above the salience threshold are included as members of a corresponding spectral set, the perceptual salience of member harmonics within a set differs significantly in terms of perceptual impact, with the louder, lower order positions of the harmonic series exerting much more influence on the resulting timbre than the softer, upper level harmonics.

As observed earlier, pitch drift and variable intonation are considered one of Alpine yodeling's most novel and idiosyncratic features, and represent important expressive devices characteristic of traditional performance practice. As a result of these considerations there emerges a significant deviation from the pitch classes indicated in the transcription and their actual acoustical counterparts as evidenced in the recording and acoustical measurements. As a consequence, and given that

SECTION A1 (-55 DBFS)

	SY	1	2	3	4	5	6	7	8	9	10	11
1	JO (1234567TE)	F4	F5	C6	F6	A6	C7	D#7	-	-	A7	B7
2	LO (123456]	Eb4	Eb5	Bb5	Eb6	G6	Bb6					
3	LO (1234]	D4	D5	A5	D6							
4	LO (1234-678- -E)	F4	F5	C6	F6	-	C7	D#7	F7	-	-	B7
5	U (123-56)	Bb4	Bb5	F6	-	D7	F7					
6	LU (123456)	C5	C6	G6	C7	E7	G7					
7	LU (12345)	D5	D6	A6	D7	F#7						
8	LU (1234)	Eb5	Eb6	Bb6	Eb7							
9	LU (123)	F5	F6	C7								
10	U (12345)	E5	E6	B7	E7	G#7						
11	U (12)	C5	C6									
12	LU (123)	F5	F6	C7								
13	U (12345)	E5	E6	B7	E7	G#7						
14	U (123456)	C5	C6	G6	C7	E7	G7					
15	U (12345)	D5	D6	A6	D7	F#7						
16	U (123)	Bb4	Bb5	F6								
17	JO (123456)	D4	D5	A5	D6	F#6	A6					
18	LO (123456789T)	F4	F5	C6	F6	A6	C7	D#7	F7	G7	A7	
19	JO (123456789)	D4	D5	A5	D6	F#6	A6	C7	D7	E7		
20	LO (123456789)	F4	F5	C6	F6	A6	C7	D#7	F7	G7		
21	U (1234-6)	Bb4	Bb5	F6	Bb6	-	F7					
22	LU (123456)	D5	D6	A6	D7	F#7	A7					
23	U (123-567)	C5	C6	G6	-	E7	G7	Bb7				
24	JO (123456789-E)	Eb4	Eb5	Bb5	Eb6	G6	Bb6	Db7	Eb7	F7	-	A7

Figure 9.8 Spectral sets and classes inventory.

25	LO (1234-6789)	F4	F5	C6	F6	-	C7	D#7	F7	G7
26	JO (123456789)	Eb4	Eb5	Bb5	Eb6	G6	Bb6	Db7	Eb7	F7
27	LO (12345678)	F4	F5	C6	F6	A6	C7	D#7	F7	
28	U (123456)	D5	D6	A6	D7	F#7	A7			
29	LU (123456)	C5	C6	G6	C7	E7	G7			
30	LU (123)	Bb4	Bb5	F6						

SECTION B1 (-55 DBFS)

	SY	1	2	3	4	5	6	7	8	9	10	11
1	hch (1234567-9TE)	F4	F5	C6	F6	A6	C7	D#7	-	G7	A7	B7
2	O (12345)	E4	E5	B6	E6	G#6	-	-				
3	Be (1234)	F3	F4	C5	F5							
4	O (123456)	D4	D5	A5	D6	F#6	A6					
5	Be (123)	F3	F4	C5								
6	Do (123456)	C4	C5	G5	C6	E6	G6					
7	Be (123456789TE)	F3	F4	C5	F5	A5	C6	Eb6	F6	G6	A6	B6
8	De (12345678)	Bb3	Bb4	F4	Bb5	D6	F6	Ab6	Bb6			
9	Jo/Lo (123456789T)	D4	D5	A5	D6	F#6	A6	C7	D7	E7	F#7	
10	Lo (1234567-9T)	F4	F5	C6	F6	A6	C7	D#7	-	G7	A7	
11	U (123-56)	D5	D6	A6	-	F#6	A6					
12	Du (123456)	C5	C6	G6	C7	E7	G#7					
13	Du (123)	Bb4	Bb5	F6								
14	Du (123- - 6)	A4	A5	E6	-	-	E7					
15	Jo/Lo (123456)	Eb5	Eb6	Bb6	Eb7	G7	Bb					
16	Lo (123456)	F5	F6	C7	F7	A7	C8					
17	U (1234567)	E5	E6	B6	E7	G#7	B8	D8				
18	Du (12345)	D5	D6	A6	D7	F#7						
19	Du (123456)	C5	C6	G6	C7	E7	G7					
20	Du (123)	Bb4	Bb5	F6								

Figure 9.8 Continued.

our concern is with the acoustical structure of the performance, we will refer to each fundamental pitch and their associated harmonics by their actual frequency measurements so as to more accurately characterize the succession of frequencies on the musical surface.

9.10 Available forms of relationship between spectral sets

Spectral sets have a variety of characteristic features that are useful for making comparisons and establishing equivalencies. In this study we distinguish the potential forms of relationship on the basis of the intersection of four analytical considerations: (1) the fundamental pitch; (2) the number of elements in the set—that is, its cardinality; (3) the respective order positions of the set's constituents within the harmonic series; and most importantly (4) the set's associated vowel.

One way of sorting spectral sets begins with consideration of the fundamental pitch and its associated octave registration. When two sets share a common fundamental, we experience them as equivalent in terms of frequency and pitch class but not necessarily in terms of tone color unless performed by the same instrument with the same formant frequency bandwidths. Sets that contain the same number of harmonics are equivalent in terms of spectral density. When two sets have different fundamentals but the same cardinality—as well as the same order positions of the harmonic series—they are related by transposition.

Unlike pitch-class sets in which we can hear: (a) the individual elements that comprise the set, and (b) the transformations that map the members of one set onto another, we cannot really hear the individual elements that comprise a spectral set with the same degree of precision as we do the sets of the fundamental. Instead of perceiving the individual elements, what we hear is the composite timbre they combine to generate. As such, conceiving of spectral contents and establishing correspondences based on the same analytical approaches as employed in traditional set class theory or motivic analysis does not work.

In general, the relative intensity (perceptibility) of a given harmonic is determined largely by its order position within the harmonic series, with each successive order position becoming increasingly softer. Consequently, as the cardinality of a spectral set expands or contracts, it tends to do so in accord with the order positions of the harmonic series. As a result of this phenomenon, in terms of pitch class and registral considerations, all sets sharing the same fundamental but involving different cardinalities will be related by some type of inclusion relation (literal subset and superset). Sets generated on different fundamental pitches will be related by transposition, and in cases where discrepant cardinalities are involved, they will exhibit an inclusion relation based on abstract subsets and superset correspondences. In rare circumstances, other acoustic factors may result in the harmonic from a particular order position within the harmonic series being masked or otherwise concealed, resulting in an empty node within the set. It is for this reason that several spectral sets have empty order positions.

While these various considerations all play a significant role in contributing to the formation of a spectral segment's unique timbral profile, they do so in a type of supporting, secondary capacity. Of the various shaping forces of timbre, it is the set's corresponding vowel type and the corresponding formant frequency bandwidths that play the most significant role in terms of timbral shaping. In singing, the harmonics generated by the vocal cords are further shaped by the resonant frequency bandwidths—also known as formants—of the vocal tract and oral cavity. These structures act as filters, intensifying some harmonics whilst dampening others. Phonologists have established that vowels are distinguished by the ratios expressed between the first three resonant frequency bandwidths, or formants, typically abbreviated as: F1, F2, and F3. In most cases, however, only the first two formants are required to disambiguate the vowel.[11] This is due to the fact that while vowels do have intrinsic frequency, intensity, and duration, the formant frequencies themselves are not directly related to the frequency of the fundamental pitch. Formant bandwidths may remain more or less constant even as the fundamental changes. In short, selection of a specific vowel sound is tantamount to choosing a particular tone color. Figure 9.9 demonstrates the ratios of formant frequencies for assorted sample vowels of the International Phonetic Alphabet.

Harmonic frequencies that align with formant frequency bandwidths appear shaded with the loudest dBFS values in each spectral set as displayed in the narrow-band spectrograph. Despite changes in corresponding frequencies, the ratio between formant frequencies remains largely invariant across males, females, adults, and children.[12] Our perceptual discriminations have evolved such that our auditory system targets the position and relative distance of these formant peaks in a sound in order to identify the corresponding vowel. Vowels and other phonemes, the constituents of speech sound, are disambiguated by the unique positioning of their associated spectral peaks (formants). By extension, in unaccompanied vocal music, the distribution and positioning of different vowels thus plays a primary role in the production of timbre, where vowels form the sustained portion of musical sounds, their corresponding vowel-formant types containing the bulk of spectral energy and exerting the most powerful shaping forces in the establishment of

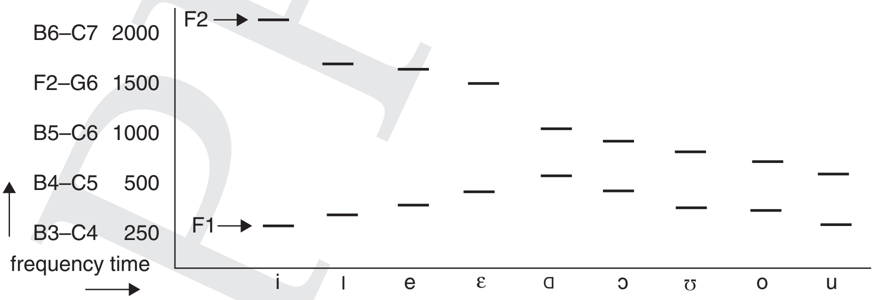


Figure 9.9 Spectrographic model of the ratios of formant frequencies for assorted sample vowels in the IPA.

vocal timbre. While the relationship between vowels, formants, and tone color has long been established in phonological, linguistic, and phonetics research involving speech sounds, the situation is far more complicated when it comes to analyzing the continuous sounds characteristic of music.

If formant frequencies are primarily responsible for defining the vowel and its associated tone color, then the remaining non-formant frequency harmonics further modify these primary vowel colors to produce a variety of available shadings, hues, and intensities. The resonant frequencies of the vocal tract and oral cavity—the formant bandwidths—distinguish between available vowels and thus establish the primary timbre of a vowel in a manner analogous to the concept of hue in contemporary color theory. There will be as many primary timbres as there are distinct vowel sounds as represented by the International Phonetic Alphabet. By extension, the intensity or amplitude of respective timbral elements is loosely analogous to the notion of saturation in color theory, where the amount of light determines the degree of presence similar to that of amplitude being indicative of the degree of presence for spectral features. Finally, the vast diversity of spectral tints and shadings made available by small variances in cardinality, registration, pitch class, content, amplitude, and intensity can be conceptualized as analogous to the notion of value—or lightness—in color theory.[13]

These various forms of timbral correspondence and pitch correspondence are summarized in the following. Timbral equivalencies are predicated on the basis of vowel-formant classification, whereas frequency and pitch-class relationships are construed solely on the basis of octave registration, pitch class, and respective order positions with the harmonic series. Spectral sets which share the same vowel also share the same formant frequencies, the latter of which establishes the primary tone color of the vowel. Such sets will be related to each other in one of four possible ways; and form a hierarchy of correspondence from strongest to weakest:

1. same vowel; same fundamental; same cardinality = identity
2. same vowel; same fundamental; different cardinality = inclusion relation (literal subset)
3. same vowel; different fundamental; same cardinality = related by transposition
4. same vowel; different fundamental; different cardinality = inclusion relation (abstract subset).

Sets may share correspondences defined by pitch relationships yet be unrelated by timbral correspondences due to their use of contrasting vowel colors:

1. different vowel; same fundamental; same cardinality = pitch identity
2. different vowel; same fundamental; different cardinality = pitch-related literal subset
3. different vowel; different fundamental; same cardinality = pitch-related transposition
4. different vowel; different fundamental; different cardinality = not related.

9.11 Measuring vowel formants (F1 and F2)

Phonologists recognize that vowels are defined by the distribution and positioning of their respective formant frequencies. Usually, only the spacing between the first and second formant bandwidths are required in order to achieve vowel disambiguation (Poole 1999). While those harmonics which do align with these formant frequencies appear to be boosted in the spectrographic display, the location and specific frequencies of the formant bandwidths themselves are not. One application available in the PRAAT linguistics software is the ability to select a time segment and then sample the segment in order to reveal both the central formant frequency and the perimeters of the frequency bandwidths for the first and second formants; consideration of additional formants is also possible.

Figure 9.10 demonstrates the distribution of formant frequencies F1 and F2 in phrase A1 of the performance. Observe that the first formant (F1) remains more or less consistent across the lowest horizontal strata of the wide-band spectrogram whereas the frequencies of the fundamental pitches are continually shifting. By using a cursor function one can inventory the approximate formant frequencies defined by the software for F1 and F2 of the spectral segment in phrases A1 and B1. When measuring formants, samples were taken at the midpoint of each spectral segment in order to highlight the vowel as much as possible while minimizing presence of consonants and other phonemes that might appear on either side. Formant frequencies are not consistent nor uniform but in a constant state of fluctuation. While the degree of fluctuation is relatively minimal compared to other aspects of musical sound, the small-scale changes that do occur have the effect of stretching or coloring the distinguishing features of the respective vowel to the extent that, at times, the boundaries which demarcate between respective vowel sounds become obfuscated and uncertain. Pinpointing the exact point where one vowel sound becomes another is no more easily achieved than determining the precise point at which one primary color becomes another in the visual domain.

The vowel timbres generated by the succession of spectral segments and their corresponding formant frequencies represent the primary shaping forces in defining timbre in unaccompanied vocal music. Selecting the accompanying vowel for a given fundamental pitch within the melody is tantamount to coloring the melody; the latter determined by the unique formant design of the corresponding vowel sound. In addition to fluctuations in formant frequencies resulting from subtle modifications of the vocal resonating cavities during performance, the respective octave registration associated with the fundamental pitch also can play a role in vowel ambiguation. When the frequency of the fundamental extends above that of the lower formant frequencies, the timbre becomes thin and weak, and the corresponding vowel sound becomes ambiguous. Typically this happens around the pitch f6, and trained singers will initiate a physical response and engage vibrato to adjust for this resulting in a shift of the formant frequencies in order to increase vocal resonance (Howell 2016).

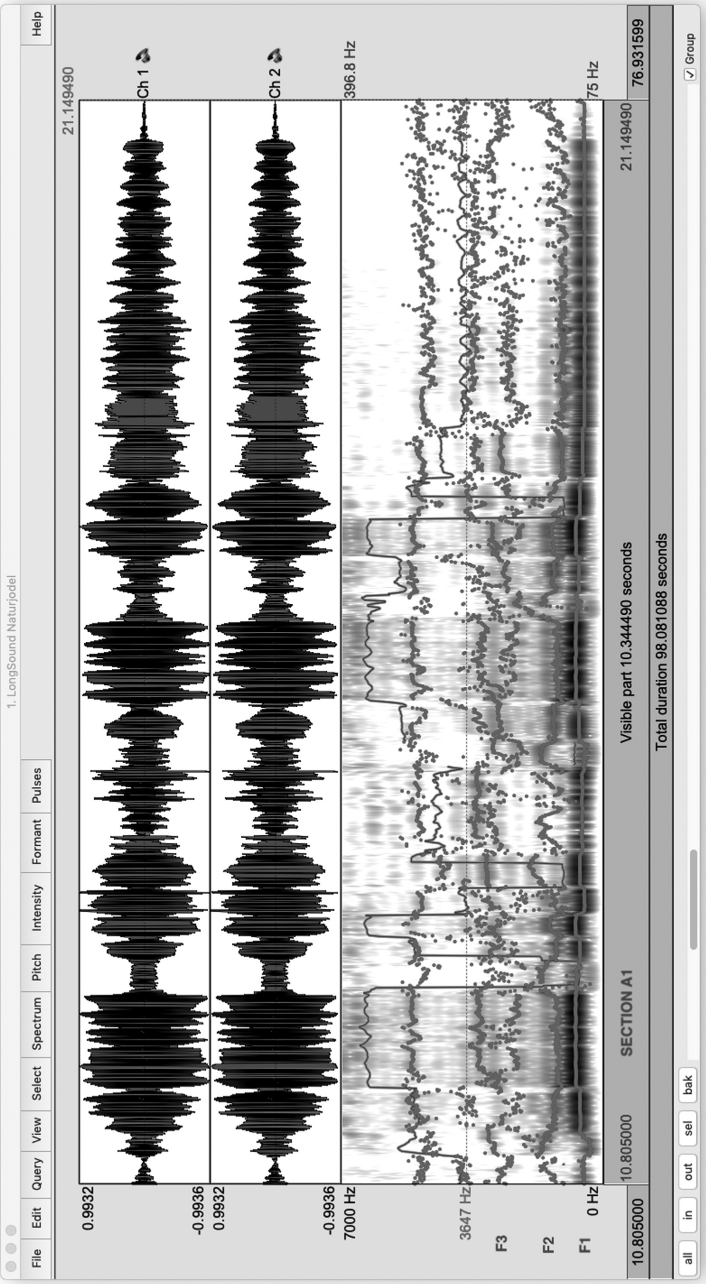


Figure 9.10 Measuring change in order positions of vowel formant frequencies.

9.12 Mapping vocal formants in IPA vowel space

Having disclosed the first and second order position formant frequencies (F1 and F2) associated with each spectral set, the next step involves mapping these formant frequencies onto the corresponding vowel space of the International Phonetic Alphabet, a sample of which is demonstrated in Figure 9.11. The primary advantage the IPA vowel space affords is its ability to compare and contrast vowel colors from any language or dialect—especially those for which no current phonetic models exist.

Using the vowel analysis function available on the VoceVista Video Pro software, the central formant frequencies (F1 and F2) illustrated in the spectrograph (located to the left of Figure 9.11) appear projected directly onto the IPA vowel space (located to the right of the example). The point of intersection between the horizontal and linear axes indicates the respective position of the corresponding vowel sound within the greater IPA vowel space as determined by the frequency and ratio of the first and second order position formant frequencies. There are 29 distinct vowels represented in the IPA vowel space, each distinguished by a unique sonic fingerprint and collectively establishing a diverse spectrum of available vowel colors. Prior to considering further how the pathways traced by the succession of vowels within the IPA space contribute to morphology, let us first examine the variety of relationships—timbre, pitch, frequency, and density—reflected between them.

9.13 Mapping transformations between spectral sets

Figure 9.12 characterizes the organization of the timbral surface of phrases A1 (mm. 1–7) and B1 (mm. 15–21); the former containing 30 spectral sets; the latter, 20. In reading the graph, the following criteria apply: Each node is numbered and represents a distinct spectral set; the harmonics that are included in each set appear as contents within each respective node. Nodes colored black correspond to the "O" vowel; dark gray highlight indicates "U" vowel sets; and light gray designates those sets with an "E" vowel. The transformations that map one node onto another are indicated by transformational arrows and subscripts: T0 means two sets are identical (transposed by "0") and that they have the same vowel and fundamental; the same cardinality; and the same order positions of the harmonic series. If two sets are otherwise identical but differ only in terms of the number of harmonics they contain, they share an inclusion relation as either literal subsets or supersets with the offset value between sets appearing in parenthesis as T0(y).

If two sets share a common vowel but have different fundamental pitches, they are related by transposition shown as Tx, 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 Tx(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, or T0; same vowel, fundamental, and cardinality); salient (same vowel, same fundamental, literal subsets);

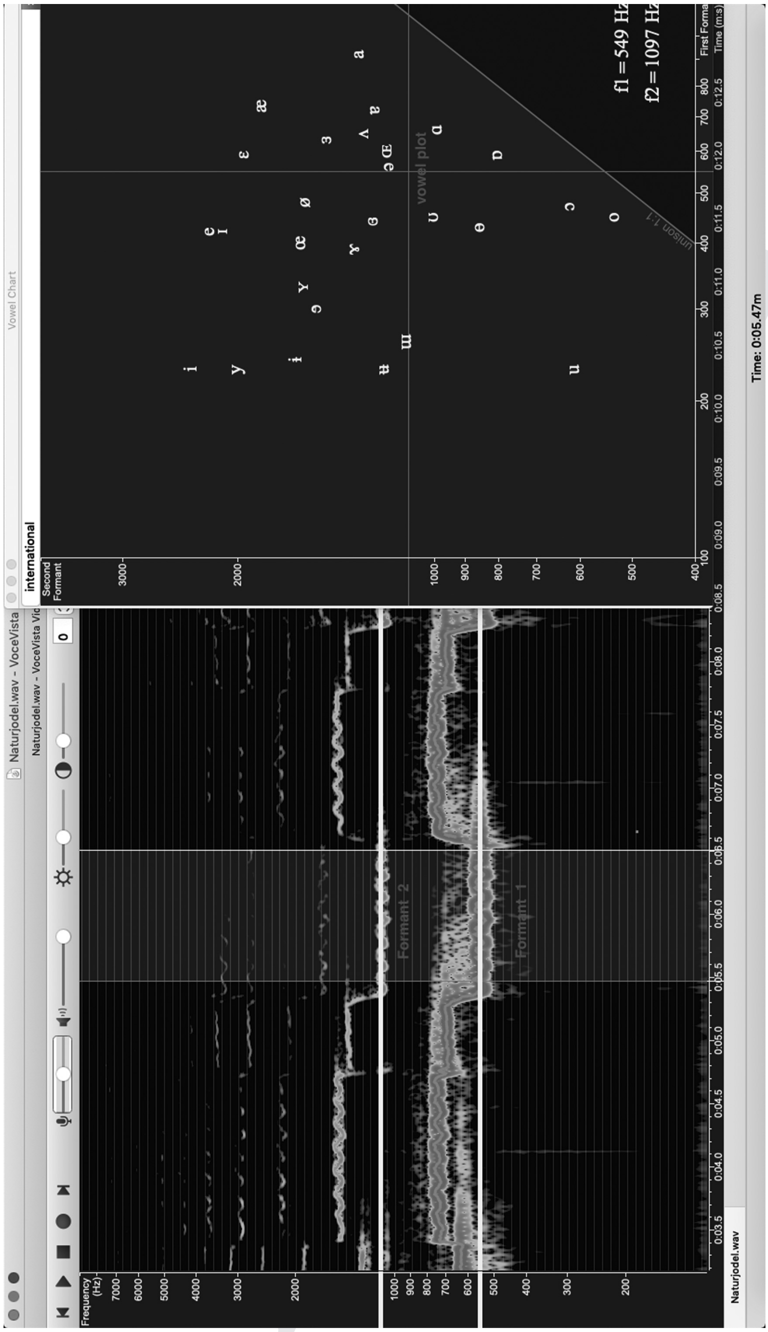


Figure 9.11 Mapping vowel formats in IPA vowel space.

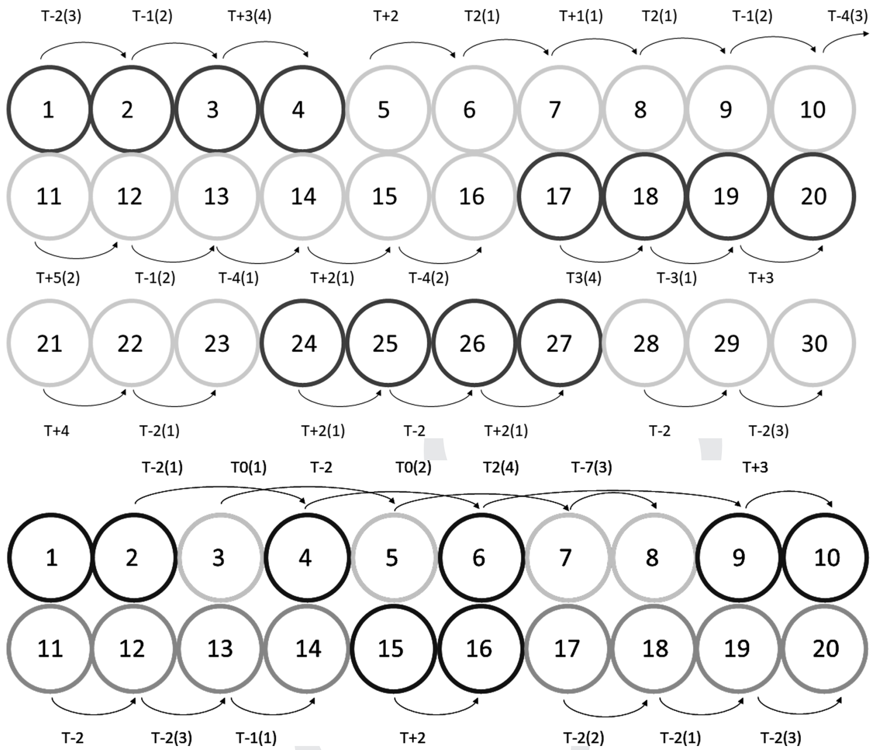


Figure 9.12 Transformations between spectral sets.

less salient (same vowel, different fundamentals, transposed or abstract subsets); to other types of correspondence predicated on the basis of pitch and frequency relationships only. In essence, the same relationships as before, but comparing only contrasting vowels.

Regarding Figure 9.12 we note that subphrase A1a (mm. 1.0–3.5) contains 16 spectral sets; subphrase A1b (mm. 3.5–7.0) contains 14 spectral sets. Moreover, we note that both subphrases begin with four successive instantiations of spectral sets characterized by the O vowel followed by numerous successive repetitions of sets containing the contrasting U vowel. O vowel sets are associated with the low register, chest voice sounds whereas the higher register instantiations of the U vowel correspond to the head voice. The manner in which the different vowel colors are positioned by the performer as well as consideration of how things change when and as they do in relation to everything else leads us to the deduction that the formal components of musical design are distinguished as much by vowel color and other timbral contrasts as by the more conventional pitch and rhythmic relationships. Moreover, whereas the two vowel colors O and U define the primary vocal timbres of the phrase, each of these primary vowel timbres is manifest in multiple shades and intensities as shaped by nuanced contrasts in terms of pitch and octave registration, spectral density, and the relative intensity of constituent harmonics within each discrete set.

Throughout phrases A1 and B1, each subphrase—albeit, with one exception—displays essentially the same distribution of primary vowel colors. In subphrase A1a, the initial portion of the phrase segment, spectral sets 1–4, present the vowel O, whereas the middle and ending of the phrase segment, display sets 5–16, are characterized by the U vowel sound. In subphrase A1b, the O vowel sets in 17–20 are followed by three iterations of U vowel sets in 21–23. The latter half of subphrase A1b then reiterates the same design with sets 24–27 returning to the opening O vowel for an additional four iterations and subsequently being followed by three concluding statements of U vowel sets. Beginnings are characterized by lower frequency sounds and colored by O vowels whereas the middle and ending portions of the phrase are colored by various shades of the U vowel.

Sometimes the differences, or in this case, distances between things, can become as important as the things themselves; the numerous instances of stepwise adjacencies implicit in the melodic design of the performance coupled with the tactical alternation of vowel colors make this performance an excellent example. Throughout section A1, it is the transformation T2 that occurs most frequently, with 11 occurrences out of a total of 24 with T1, T3, T4 each occurring four times and a single instance of T5.

This organization of the timbral surface is coordinated with transformations in pitch and rhythmic design resulting in the bipartite formal shape of phrase A1. The tripartite distribution of subphrases in phrase B1 are also articulated and defined through a combination of timbral shaping coordinated with other parameters of musical design. Similar to both subphrases A1a and A1b, subphrases B1b and B1c present timbral designs already seen. Each starts with an O vowel for several repetitions before moving to a longer succession of U vowel spectral sets. Indeed, with one exception, each subphrase in both phrases A and B employs the same succession of vowel colors beginning with O vowel color followed by longer succession of U vowels. Subphrase B1a is the exception: it features a consistent alternation of vowel colors O and E distributed within a stepwise linear octave descent spanning F4 down to F3. The introduction of the new vowel color E provides a stark timbral contrast to the binary vowel color scheme in play up to this point. Interestingly, the acute vowel E appears positioned on the low pedal pitch F3 whereas the grave vowel O is positioned in the higher register and coincides with the linear descending line F4 to F3.

9.14 Mapping spectral morphologies

We now shift our focus from consideration of the objects themselves to the manner in which they change over time. In our 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, vowel modulations in IPA space, and the intensity fluctuations of individual harmonics between successive spectral sets. We understand that there are numerous other contributing factors involved in our perception and experience of spectral morphology, but we distinguish these four analytical parameters in particular as important constituents that are universal to sonic design and morphology in all styles and traditions of unaccompanied vocal music and therefore provide a foundation on which more sophisticated future studies can build.

The analytical tools and strategies developed herein have their origins in the brilliant, innovative work of Robert Cogan; in particular, his adaption of Roman Jakobson's phonological theories developed for modeling the constituents of speech sound to the analysis of tone color in music. Central to Cogan's early explorations of spectral morphology is his table of 13 binary oppositions adapted from Jackobson's structuralist approach. The primary characteristic that distinguishes our approach to morphology as distinct from Cogan's table of binary oppositions is that in our methodology the different features of sonic design—spectral density, compass, and loudness—are conceptualized not simply as binary oppositions but rather as a spectrum of incremental gradations or values, which can be measured and evaluated and thus provide a more detailed characterization of how various timbral elements interact on the timbral surface and contribute to spectral morphology.

Moreover, the various oppositions themselves are not equal in terms of perceptual weight and register. Some of the features involved in the binary pairings play a much stronger, more formative role, and impact our perceptions in more meaningful ways than others. The premise of this type of structuralist's feature vector analysis is that by viewing an object from multiple vantage points, each emphasizing a different feature or opposition, one could begin to make sense of an object's essential structure. Each feature reveals a single cross section. When these individual sections are multiplied and combined, they reveal the composite design. The more features or oppositions employed, the more analytical perspectives afforded. This results in the underlying structure becoming increasingly articulate. Our initial consideration, spectral density (i.e., cardinality), indicates the respective weight of a spectral set contingent upon the number of elements contained in the set. Increases in density represent progressive or increasing intensity fluctuation values, while the reverse represents decreasing, or recessive intensity fluctuations. Mapping successive changes in set density across the performance reveals the organization of an important consideration regarding spectral morphology.

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. In unaccompanied vocal music there is a direct correlation between the number of sounding harmonics (density) and the corresponding spectral compass of the set, given that the spectral sets usually increase or decrease membership in alignment with the respective order position of harmonics within the harmonic series. A spectral set with a cardinality of five will typically consist of the first five order positions of the corresponding harmonic series; a seven note set will use the order positions 1–7 of the corresponding series and so forth.

Only rarely are one or more harmonics missing due to masking or other interference, and an incomplete set results. Whereas complete sets sound full, incomplete sets sound increasingly hollow as the number of missing elements increase. Figure 9.13a demonstrates the correlation of spectral density and spectral compass values as they occur in section A1 (segments 1–30) and B1 (segments 31–50). Figure 9.13b illustrates the relationship between vowel types on cardinality.

Vowel modulations in IPA vowel space—one of the most important shaping forces regarding vocal timbre—results from the successive modulations of vowels and their defining formant frequencies across the performance. Videos 9.1a and 9.1b demonstrate how the individual vowel-formant plots demonstrated previously can be combined to reveal distinct pathways within the IPA vowel space for sections A1 and B1 respectively (**<URL HERE for associated eResources>**). Moreover, we can partition the IPA vowel space into distinct sectors corresponding to the respective degree of intensity associated with each corresponding vowel sound: acute—such as the high intensity vowels "I" or "e"; neutral—vowels such as "a"; or grave—low intensity vowels such as "o" or "u." Each of these three categories includes multiple gradations depending on the exact positioning of the various vowel sounds within the greater IPA vowel space.

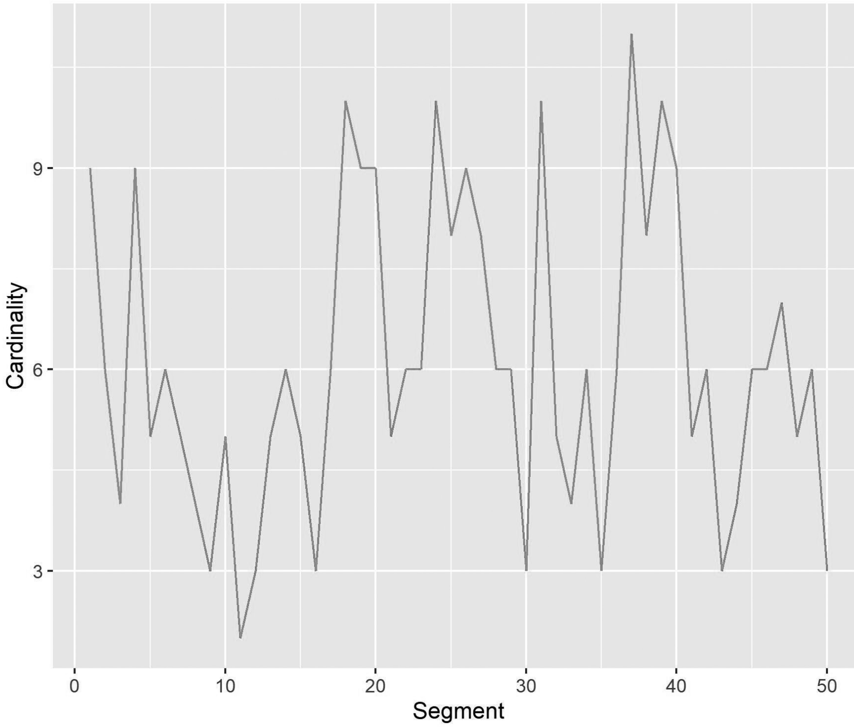
9.14.1 Intensity fluctuation (dBFS)

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, organizes these into sets, and develops a variety of available correspondences between them predicated on the basis of implicit vowel color in conjunction with other considerations involving pitch, frequency, respective order position within the harmonic series, and set cardinality.

We will now construct a new perspective, one that attempts to understand the relative intensity of these sonic details and how they change over time. 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.

Now we consider how the respective intensity of each harmonic in each spectral set changes over time—a phenomenon that we consider to be descriptive of spectral morphology. This dynamic perspective expands the manner in which we conceptualize spectral sets. In addition to now being able to distinguish and compare sets in terms of the particular harmonics contained therein, we can now map out the relative intensity each member harmonic contributes to its corresponding spectral set and view their changes over the course of the performance.

In unaccompanied vocal music, the loudest, most powerful harmonics are typically those positioned in the lower order positions of the harmonic series, just above the fundamental pitch. In general, these harmonics tend to occur within the 1–4 hz frequency range. Because this part of the spectrum is intensified in our perception, pitches in this range seem louder in our ears than their actual acoustic measurements would suggest. Moreover, the formant frequency bandwidths F1



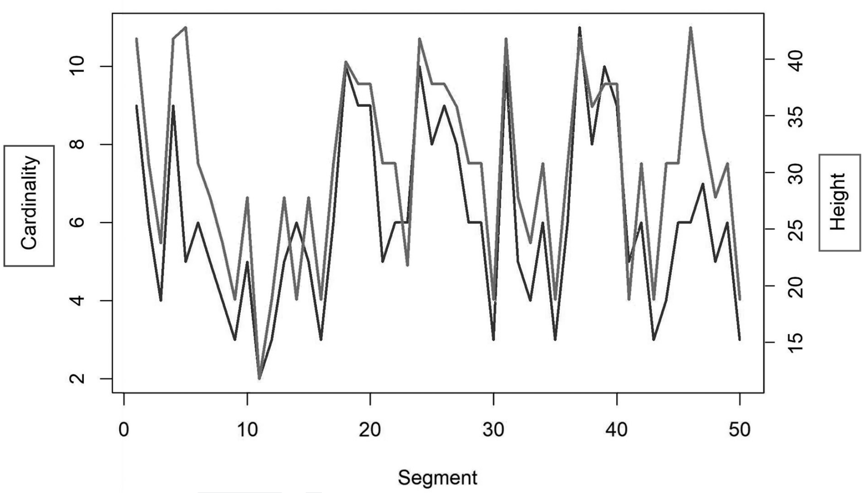


Figure 9.13 Contour graphs showing correlation between spectral density and height.

and F2, which also occur in this frequency range, have an intensifying effect on these lower-order harmonics.

Figures 9.14a and 9.14b display the harmonics from order positions one-through-three of the harmonic series and illustrate the fluctuations in their respective intensity values over time as measured in dBFS values. Note that in phrase A1 (segments 1–30) fluctuations in intensity levels for harmonics H1–H3 are coordinated, and produce a succession of two available combinations of harmonics arranged in relation to their respective dBFS intensity values. One configuration (A) presents the first harmonic (H1) as the loudest, with the third harmonic, H3, being the weakest. The second harmonic is generally in between the two extreme values, as is evident in sets 5–16, or until the end of subphrase A1a.

The second available configuration (B) places H2 as the loudest harmonic, with H1 and H3 positioned together at the opposing extreme. The two available configurations are inversely related: the A configuration of harmonics include sets 5–16; 20–24; and 28–30 with B arrangement occurring in sets 16–20; and 24–28. Phrase B1 presents similar relationships amongst its lower order harmonics and their respective intensity level fluctuations. Segments 1–9 are characterized by the B configuration with a particularly prominent second harmonic. The remainder of the phrase displays the arrangement A with the equally prominent first harmonic now replacing the second harmonic of the previous arrangement. When the O vowel is in play as in segments 1–4; 17–20; and 23–27, the first harmonic is most prominent and the third harmonic recedes. When the U vowel is in play the second harmonic emerges as prominent and the first harmonic recedes. Similar relationships may be construed for phrase B1 as well, where changes in melodic tessitura and accompanying vowel sounds create an exchange of roles between the first and third harmonics similar to that of phrase A1.

If it is the lower order position harmonics and the formants that intensify and define the basic vowel color, it is the mid-to-upper tier harmonics that primarily determine the specific shade and hue a particular vowel color assumes. In general, the more upper end harmonics, the more rich and brilliant the sound, but there are numerous factors that mitigate this finding, foremost among them being the respective octave register of the fundamental pitch. Lower fundamentals leave room for more high harmonics sounding within the frequency range of human perception. The higher the fundamental pitch, the less its harmonics will be within the boundaries of our perception. Figures 9.15a and 9.15b demonstrate the number of occurrences for harmonics occupying order positions 1–10 of the harmonic series in phrase A1. Instantiations appear as individual dots appearing: (a) within their respective vertical strip in the main body of the graph, and (b) intersecting with the corresponding dBFS values located in the vertical scale on the left.

A complementary perspective is afforded by the interactive Webpages 9.1a and 9.1b (**<URL HERE for associated eResources>**). These webpages demonstrate the same collection of harmonics shown in Figure 9.14, but here displayed according to their actual linear distribution and indicating the changes in relative intensity for each harmonic in each spectral set in phrases A1 and B1 respectively.

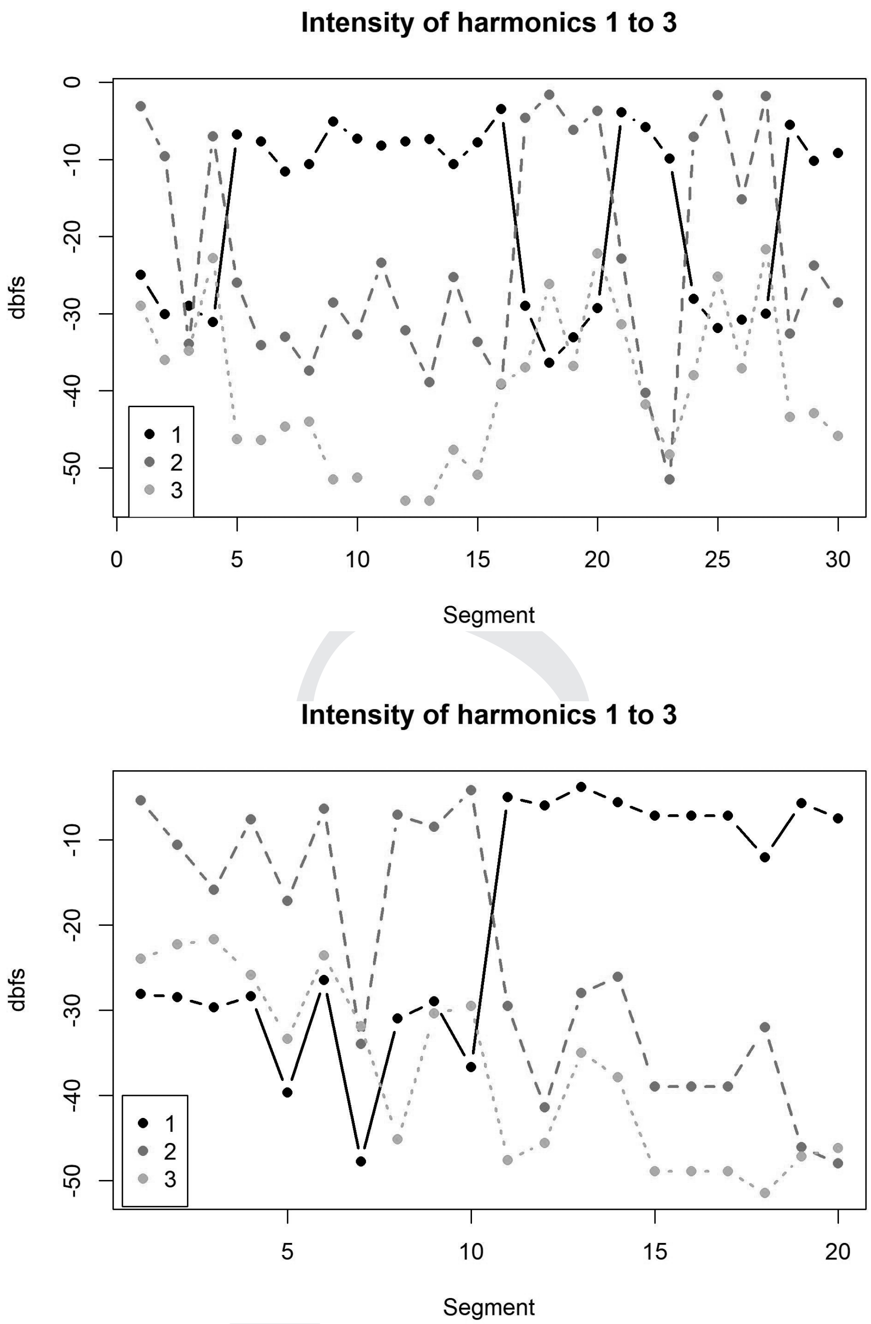


Figure 9.14 Harmonic intensity order positions 1–3.

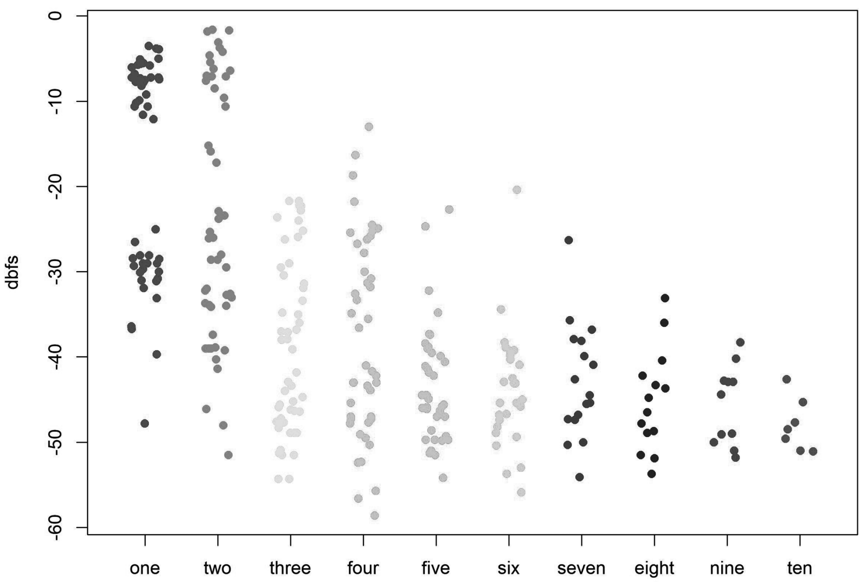


Figure 9.14 Continued

Due to their being interactive, the user can select any point within the graphic configuration to obtain additional information regarding pitch class, frequency, intensity, and intonation.

9.15 Conclusion

In this chapter, we have demonstrated a broad assortment of analytical tools and strategies useful for characterizing the organization of timbral surfaces, along with other select features of sonic design including pitch drift, intonation, vibrato, vowel types, and various considerations involved in the evaluation of spectral morphology. The analytical technique of disclosing spectral sets provides an efficient means not only with which to characterize the precise organization of any given spectral segment, but also as a means to compile a descriptive inventory of the available sonic contexts within a given performance—and by extension, a given style or genre. Instead of relying on adjectival descriptors to characterize the timbre of complex musical sounds, we can characterize the exact contents of a sonic event, thus not only defining the constituent elements of which it is comprised but also indicating with technical precision the exact power or intensity of each discrete spectral element within a given set as well as numerous other nuanced details regarding sonic design.

By isolating the lower order position formant frequencies for each vowel sound on the spectrograph we then projected these onto the corresponding IPA vowel space in order to show how the succession of vowel sounds in *Höch Obe*

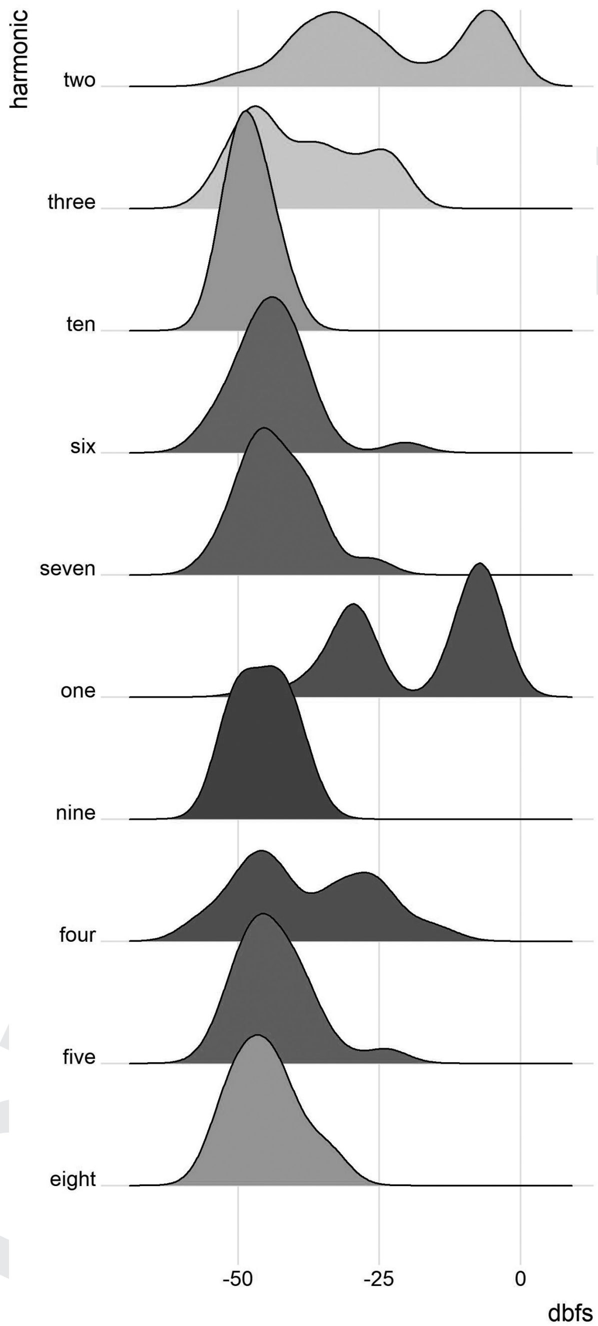


Figure 9.15 Harmonic intensity order positions 1–10.

can be visualized and shown to be tracing discrete transformational pathways within the greater spectrum of available vowel colors of the IPA space. Finally, we have demonstrated that in unaccompanied Alpine yodel, the timbre of the vowel selected by the performer is primarily contingent upon the formant frequencies which disambiguate the vowel and define its unique timbral color. The remaining considerations such as set density, spectral compass, and variances in fundamental pitch tweak this primary color to achieve instantiations of the enormous variety of timbral shades and intensities that are assumed by the primary vowel tone color.

Notes

1 This definition is incomplete insofar as yodels that are exclusively sung in the chest voice do exist, notably in the Appenzell region of northeastern Switzerland.
2 The first term designates a folk song with a yodeling refrain, which was first introduced in the 1810s by Ferdinand Huber (Kammermann et al. 2016: 22). "Natural yodel" designates a melody entirely performed on wordless yodel syllables. "Natural yodel" and yodel songs are widely regarded as the vocal forms that represent Swiss traditional music (Ammann et al. 2019: 13). In various other Alpine regions, where yodeling is practiced (e.g., Tyrol), both forms are equally designated as yodel ("Jodler").
3 Rymann (1933–2008) became famous for his interpretation of the folk song "Dr Schacher Seppli" (Am Acher 2002). He received international acknowledgement, and his obituary appeared in *The New York Times*.
4 The reason for this effect lies in the linear model of the degree of E, which is based on only seven notes (compared to 17 to 29 notes in all other cases) and is distorted by the very low intonation of the first note. Therefore, this degree was omitted in Figure 9.4.
5 Definition of timbre by the Acoustical Society of America: https://asastandards.org/Terms/timbre/.
6 The fundamental is the first harmonic; the second harmonic is the first overtone.
7 See www.voicescienceworks.org/harmonics-vs-formants.html.
8 The ear is not equally sensitive to all frequencies, particularly in the low and high frequency ranges. The response to frequencies over the entire audio range has been charted, originally by Fletcher and Munson in 1933, with later revisions by other authors, as a set of curves showing the sound pressure levels of pure tones that are perceived as being equally loud. The curves are plotted for each 10 dB rise in level with the reference tone being at 1 kHz—also called loudness level contours and the Fletcher-Munson curves. The curves are lowest in the range from 1 to 5 kHz, with a dip at 4 kHz, indicating that the ear is most sensitive to frequencies in this range. The intensity level of higher or lower tones must be raised substantially in order to create the same impression of loudness. The phons scale was devised to express this subjective impression of loudness, since the decibel scale alone refers to actual sound pressure or sound intensity levels. See www.sfu.ca/sonic-studio-webdav/handbook/Equal_Loudness_Contours.html.
9 In terms of a specific frequency, the critical band is the smallest band of frequencies adjacent that activate the same part of the basilar membrane. In a complex tone, the critical bandwidth corresponds to the smallest frequency difference between two

partials such that each can still be heard separately. See www.sfu.ca/sonic-studio-webdav/handbook/Critical_Band.html.

10 Salience verification protocols attempt to provide some type of phenomenological accountability to justify and support our analytical insights and assertions given the current deficit of perceptual models on which to base our study. Numerous factors impact the perceptual acuity of a sonic feature including considerations regarding recording and playback technology, the psycho-physical constraints of the listener, the "angle" of listening employed, familiarity with the sound sample, etc. In the end, the analytical data generated through this analytical process reflects our listening experience but we assert that most attentive listeners with repeated hearings would also be able to identify the same if not more sonic features.

11 See www.voicescienceworks.org/harmonics-vs-formants.html.

12 See www.u.arizona.edu/~ohalad/Phonetics/notes/Formants%20Spectrograms%20and%20Vowels.PDF.

13 https://thevirtualinstructor.com/Color.html.

Bibliography

Ambrazevičius, Rytis. 2014. "Performance of Musical Scale in the Traditional Vocal Homophony: Lithuanian Examples." *Muzikologija* 17: 45–68. https://doi.org/10.2298/MUZ1417045A.

Am Acher, Paul. 2002. *Ruedi Rymann und der Schacherseppli: Jodler, Komponist, Dichter, Schwinger und Wildhüter: Eine Biographie*. Bettlach: Stubete-Verlagshaus.

Ammann, Raymond, Andrea Kammermann, and Yannick Wey. 2019. *Alpenstimmung. Musikalische Beziehung zwischen Alphornmusik und Jodel – Fakt oder Ideologie?* Zürich: Chronos. www.chronos-verlag.ch/public-download/2381.

Aruffo, Christopher, Robert Goldstone, and David Earn. 2014. "Absolute Judgement of Musical Interval Width." *Music Perception: An Interdisciplinary Journal* 32 (2): 186–200. https://doi.org/10.1525/mp.2014.32.2.186.

Ashby, Michael. 2005. *Introducing Phonetic Science*. Cambridge: Cambridge University Press.

Babbitt, Milton. 1965. "The Use of Computers in Musicological Research." *Perspectives of New Music* 3 (2): 74–83. https://doi.org/10.2307/832505.

Bachmann-Geiser, Brigitte (ed.). 2010. Die JodelArten der Schweiz, CD, Oberhofen am Thunersee: Zytglogge.

Clark, John, and Colin Yallop. 1995. *An Introduction to Phonetics and Phonology*. Cambridge: Basil Blackwell Ltd. https://doi.org/10.1017/S0025100311000351.

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.

———. 1999. *The Sounds of Song*. Cambridge MA: Publication Contact.

Fales, Cornelia. 2008. "The Implicitness of Timbre: Attitudes Towards Timbre in Barundian and Western Art Musics." In *Spectral World Musics: Proceedings of the Istanbul Spectral Music Conference*, edited by Robert Reigle and Paul Whitehead, 63–74. Istanbul: Pan Yaymcilik.

———. 2002. "The Paradox of Timbre." *Ethnomusicology* 46 (1): 56–95. https://doi.org/10.2307/852808.

Fink-Mennel, Evelyn. 2007. *Johlar und Juz. Registerwechselnder Gesang im Bregenzerwald (mit Tonbeispielen 1937–1997)*. Graz and Feldkirch: Neugebauer.

Goldstone, Robert, and Andrew Hendrickson. 2010. "Categorical Perception." *Cognitive Science* 1: 69–78. https://doi.org/10.1002/wcs.26.

Graf, Walter. 1965. "Naturwissenschaftliche Gedanken über das Jodeln. Die phonetische Bedeutung der Jodelsilben." *Schriften des Vereins zur Verbreitung naturwissenschaftlicher Kenntnisse in Wien* 105: 2–25.

Hänggi, Eugen. 2011. *Йодель в истории музыкальной культуры Швейцарии*. Doctoral Dissertation, St. Petersburg Conservatory.

Heidemann, Kate. 2016. "A System for Describing Vocal Timbre in Popular Song." *Music Theory Online* 22 (1). https://mtosmt.org/issues/mto.16.22.1/mto.16.22.1.heidemann.html.

Howard, David. 2007. "Intonation Drift in A Capella Soprano, Alto, Tenor, Bass Quartet Singing with Key Modulation." *Journal of Voice* 21 (3): 300–315. https://doi.org/10.1016/j.jvoice.2005.12.005.

Howell, Ian. 2016. Parsing the Spectral Envelope. Doctoral Dissertation, New England Conservatory. https://pdfs.semanticscholar.org/1008/defb70e0c283c4450ee0e3c2061841e89eb9.pdf.

Iverson, Paul, and Carol Krumhansl. 1993. "Isolating the Dynamic Attributes of Musical Timbre." *Journal of the Acoustical Society of America* 94 (5): 2595–2603. https://doi.org/10.1121/1.407371.

Jakobson, Roman, Gunnar Fant, and Morris Halle. 1967. *Preliminaries to Speech Analysis*, 7th edition. Cambridge MA: MIT Press.

Jakobson, Roman, and Linda Waugh. 2002. *The Sound Shape of Language*, 3rd edition. New York: Mouton de Gruyter. https://doi.org/10.1515/9783110889451.

Kammermann, Andrea, Yannick Wey, and Raymond Ammann. 2016. "Ferdinand Fürchtegott Huber, Initiator der musikalischen Beziehung zwischen Alphorn und Jodel." Schweizer Jahrbuch für Musikwissenschaft – Annales Suisses de Musicologie – Annuario Svizzero di Musicologia 36: 11–37.

Kharuto, Alexander. 2005. "About the Analysis of Vibrato Properties in Academic Vocal Performance." *XVI Session of the Russian Acoustical Society*, 296–299.

Krumhansl, Carol. 1989. "Why Is Timbre So Hard to Understand?" In *Structure and Perception of Electroacoustic Sound and Music*, edited by Sören Nielzén and Olle Olsson, 43–53. Amsterdam: Elsevier. http://music.psych.cornell.edu/articles/timbre/Why_Is_Musical_Timbre_so%20hard_to_understand.pdf.

Latartara, John. 2004. "Multidimensional Musical Space in Hildegard's 'O rubor sanguinis': Tetrachord, Language Sound, and Meaning." *Indiana Theory Review* 25: 39–74. www.jstor.org/stable/24045281.

———. 2008a. "Machaut's Monophonic Virelai *Tuit mi penser*: Intersections of Language Sound, Pitch Space, Performance, and Meaning." *Journal of Musicological Research* 27 (3): 226–253. https://doi.org/10.1080/01411890802212592.

———. 2008b. "Pedagogic Applications of Fourier Analysis and Spectrographs in the Music Theory Classroom." *Journal of Music Theory Pedagogy* 22: 61–90.

Lerdahl, Fred. 1987. "Timbral Hierarchies." *Contemporary Music Review* 2: 135–160. https://doi.org/10.1080/07494468708567056.

Leuthold, Heinrich. 1981. *Der Naturjodel in der Schweiz*. Altdorf: Robert Fellmann-Liederverlag.

Luchsinger, Richard, and Gottfried Arnold: 1949. "Die Jodelstimme." In *Lehrbuch der Stimm- und Sprachheilkunde*, edited by Richard Luchsinger and Gottfried Arnold, 70–72. Vienna: Springer.

Malloch, Stephen. 2000. "Timbre and Technology: An Analytical Partnership." *Contemporary Music Review* 19 (2): 155–172.

Martin, Silke. 2012. "Alpenloopings in HEIMATKLÄNGE. Jodeln als Globalisierungsbewegung zwischen Tradition und Experiment." *Kieler Beiträge zur Filmmusikforschung* 8: 287–315. http://geb.uni-giessen.de/geb/volltexte/2014/10697/pdf/Samples11_martin.pdf.
Mauch, Matthias, Klaus Frieler, and Simon Dixon. 2014. "Intonation in Unaccompanied Singing: Accuracy, Drift, and a Model of Reference Pitch Memory." *Journal of the Acoustical Society of America* 136 (1): 401–411. https://doi.org/10.1121/1.4881915.
McAdams, Stephen. 1999. "Perspectives on the Contribution of Timbre to Musical Structure." *Computer Music Journal* 23 (3): 85–102. https://doi.org/10.1162/014892699559797.
Odden, David Arnold. 2005. *Introducing Phonology*. Cambridge: Cambridge University Press.
Poole, Stuart. 1999. *An Introduction to Linguistics*. New York: St. Martin's Press.
Prame, Eric. 1998. "Vibrato Extent and Intonation in Professional Western Lyric Singing." *The Journal of the Acoustical Society of America* 102: 616–621. https://doi.org/10.1121/1.419735.
Räss, Nadja, and Franziska Wigger. 2010. *Jodel: Theorie & Praxis*. Altdorf: Mülirad.
Rymann, Ruedi. 2010. "Hech Obe." In *Traditionelle Schweizer Musik / Traditional Swiss Music*, edited by Brigitte Bachmann-Geiser, CD 2, Track 11. Basel: Zytglogge.
Seaton, Richard, David Sharp, and Dennis Pim. 2014. "Pitch Drift in A Capella Choral Singing: The Outcomes from an International Survey." *Proceedings of the Institute of Acoustics* 36 (3): 12–319. http://oro.open.ac.uk/42111/3/rkseaton.pdf.
Sethares, William. 2005. *Tuning, Timbre, Spectrum, Scale*, 2nd edition. London: Springer. https://doi.org/10.1007/b138848.
Sichardt, Wolfgang. 1939. *Der Alpenländische Jodler und der Ursprung des Jodelns*. Berlin: Bernhard Hahnefeld.
Slawson, Wayne. 1987. "Sound-Color Dynamics." *Perspectives of New Music* 25(1–2): 156–181. www.jstor.org/stable/833097.
———. 1985. *Sound Color*. Berkeley: University of California Press.
———. 1981. "The Color of Sound: A Theoretical Study in Musical Timbre." *Music Theory Spectrum* 3: 132–141.
Vennard, William. 1967. *Singing: The Mechanism and the Technic*. New York: Carl Fischer.
Wey, Yannick. 2019. *Transkription wortloser Gesänge. Technik und Rückwirkungen der Verschriftlichung des Jodelns und verwandter Gesänge im deutschsprachigen Alpenraum*. Innsbruck: Innsbruck University Press. https://doi.org/10.15203/3187-81-8.
———. 2020. "Transformations of Tonality: A Longitudinal Study of Yodeling in the Muotatal Valley, Central Switzerland." *Analytical Approaches to World Music* 8/1: 144–163. www.aawmjournal.com/articles/2020a/Wey_AAWM_Vol_8_1.html.
Wiggins, Geraint. 2009. "Computer Representation of Music in the Research Environment." In *Modern Methods for Musicology: Prospects, Proposals, and Realities*, edited by Tim Crawford and Lorna Gibson, 7–22. Farnham: Ashgate. https://doi.org/10.4324/9781315595894.

Computer Software Applications

Audacity (Version 2.3.3) [computer software]. 2020. Available from www.audacityteam.org.
Boersma, Paul and David Weenink. 2020. PRAAT [computer software] (Version 6.0.46). Amsterdam: University of Amsterdam. Available from 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. Available from 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. Available from www.sygyt.com/en/.

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