Spectral Fission in Barbershop Harmony

Why do barbershop chords "ring"? It will be useful to have this sound in our ears as we consider this question, so let's listen for "ringing" in an excerpt by Vocal Spectrum, the 2006 champions of the Barbershop Harmony Society's International Quartet Contest. This is the end of the tag to their "Aladdin Medley" [Audio Example 1]. How have previous listeners explained this powerful sound? Most responses to this question have historically come from barbershop practitioners uninterested in a precise answer, and old issues of their official magazine are particularly telling in this regard. One of its earliest issues tells readers: "One of the joys of barbershopping, unaccompanied, is the fact that true harmony can be attained," although it does not explain why "true harmony" would lead to ringing.[1] The chair of their Barbershop Craft Committee later declared: "We may not have a scientific answer for [why chords "ring"] but we know we do love it and that's enough for us."[2] And an article actually entitled "What makes the barbershop 7th ring?" features a detailed description of just intonation without answering its titular question, instead concluding: "Until [people] feel the thrill of ringing chords, they have not experienced the ultimate in harmony."[3]

Having just presented you with "the ultimate in harmony" courtesy of Vocal Spectrum, I have an answer of my own to provide. In this paper I argue that the distinctive sound of barbershop's "ringing" chords results from our perception of certain upper partials of vocal timbre as discrete pitches, a phenomenon I call spectral fission. I will begin by situating my theory of spectral fission in relation to existing theories of timbre and pitch perception. I will then turn to acoustically deconstructing barbershop's "ringing" chords. Finally, I will combine my perceptual framework with this acoustical perspective through analytical vignettes in which I apply computational models to recordings by two championship barbershop quartets, Vocal Spectrum and Ringmasters. In so doing,

---

1 Reagan 1944, p. 10
2 Ewin 1961, p. 30
3 DePaolis 1954, p. 45

I will demonstrate a methodology that accounts for the relationship between timbre and pitch in this music as a means of explaining our experiences of the chords its singers describe as "locked in."

Timbre is often defined by what it is not, as typified by the widely used ANSI definition: "That attribute of auditory sensation which enables a listener to judge that two nonidentical sounds, similarly presented and having the same loudness and pitch, are dissimilar."[4] Some scholars have even called timbre an "auditory wastebasket" that simply catches what's left after pitch, loudness, duration, and spatialization have all been accounted for.[5] Pitch is a far less slippery target, and its modern neurocognitive definition—the subjective perception of periodicity—positively expresses that pitch is usually perceived due to a particular frequency having a certain amount of acoustical energy.[6] And yet, timbre's reliance on spectral information resembles pitch's reliance on frequency. Stephen McAdams defines perceptual fusion as "the process by which simultaneous things are grouped into single entities."[7] Because upper partials constitute the spectrum of a complex sound, in the case of timbre McAdams specifically calls this spectral fusion.

Previous work on timbre has defined notions of "chimeric sound,"[8] "expanded sound,"[9] and "timbre anomaly by extraction,"[10] all of which describe situations in which ordinary spectral fusion does not take place but without explicitly invoking pitch perception. I propose that under certain acoustical circumstances—including the case of "ringing" barbershop chords—the precise opposite of spectral fusion occurs, meaning certain upper partials are perceived as separate pitches from their fundamentals. Loosely borrowing a term from psychophysics, I call this spectral fission to highlight its close but inverse relationship to the mechanisms of timbre perception.[11] In order for an upper

[4] Quoted in Siedenburg et al. 2019, p. 4
[5] Bregman 1990, p. 92
[6] Langner 2015, p. 14
[7] McAdams 1982, p. 280
[8] Bregman 1990, pp. 459–60
[9] Averill 1999, pp. 44–50
[10] Fales 2002, pp. 66–8
[11] Chalika and Warren 1994

partial frequency to be perceived as a discrete pitch, I posit that two requirements must be fulfilled. First, an upper partial must have enough amplitude so that the auditory system can single it out from other partials in the spectrum. Sufficient amplitude, however, does not always lead to spectral fission because several upper partials in close proximity all having similarly high amplitudes are more likely to be perceived as a timbral feature than as a cluster of pitches. In light of this, my second requirement is that an upper partial must be prominent, clearing a threshold for its amplitude to be processed as a separate pitch by having more than its neighbors, its neighbors' neighbors, and so on. Thresholds for both amplitude and prominence are necessarily relative, varying across sonorities as well as listeners and thus being par for the course as an explanation of subjective pitch perception.

How, then, do "locked in" barbershop chords fulfill these requirements of spectral fission? As performed by the very best barbershop quartets, these chords are defined by maximal blending, minimal use of vibrato, and vertical just intonation.[12] Acoustically speaking, the first two of these characteristics primarily serve to minimize unstylistic "chorus" effects, but vertical just intonation combined with the structure of vocal timbre provides an explanation of certain partials having enough amplitude to be heard as discrete pitches. Vocal production begins with the source spectrum, the raw sonic material resulting from the vibration of vocal folds.[13] The amplitudes of each partial are modified by formants, which vary as a singer changes the shape of their vocal tract to obtain desired vowels and tone colors.[14] Due to the nature of airflow through the vocal folds, the source spectrum of the voice is harmonic: the frequencies of its upper partials are integer multiples of its fundamental frequency.[15] When singing in just intonation and tuning fundamentals in integer relationships, there is therefore overlap between these upper partials. Let's consider the just perfect fifth as an illustration. By

[12] Ternström and Kalin 2007, p. 1796
[13] Sundberg 2019, p. 121
[14] Ibid., pp. 127–9
[15] Ibid., p. 121

definition, its fundamental frequencies are tuned in a 3:2 relationship. If these are realized by sine waves, their combined spectrum is simply two parallel lines [Example 1a]. Due to the harmonicity of the human voice, however, the spectral reality of singing a just perfect fifth involves every 3rd partial of the lower pitch coinciding with every 2nd partial of the higher pitch [Example 1b]. In the following recording, we can hear these overlapping partials when the second voice enters [Audio Example 2].

Acousticians at the Royal Institute of Technology in Stockholm have shown that in the most successful barbershop quartets, singers spread their formants to cause desired "ringing," deliberately tuning at least half to overlapping partials.[16] And a recent corpus study found spectral peaks in the range 1000–3000 Hz consistent with this systematic formant tuning.[17] I accordingly use a vowel-neutral predictive model of vocal timbre to identify frequencies of maximal overlap in the source spectra since additional amplitude is created through formant tuning at such overlaps. My model gives each fundamental 15 harmonic partials with amplitude inversely proportional to frequency, so the 2nd partial has twice the frequency and one-half the amplitude of the fundamental, and so on, accounting for the fact that vocal folds more or less vibrate like a triangle wave [Example 2a].[18] I have written a function that takes the fundamentals of a just barbershop chord and predicts the top four frequencies in the range 1000–3000 Hz that will have the most amplitude, assuming formants tuned to partials are tuned to frequencies of maximal spectral overlap because the quartets in question are exceptionally good [Example 2b]. In this visualization, each color represents the source spectrum of one quartet member, while the boxes denote the frequencies of maximal overlap.

To account for the prominence requirement of spectral fission, I implement an original digital signal processing script. This script first uses the Discrete Fourier Transform to compute the frequency spectrum of a recording, which it splits into thousands of 4.6 ms time slices to strike a

[16] Ternström and Kalin 2007, p. 1801
[17] Drown and Cottingham 2018
[18] Personal communication, Bill Sethares

balance between frequency resolution and temporal accuracy [Example 3a]. In each time slice, it runs a March Madness-style single-elimination tournament between all frequencies in the range 1000–3000 Hz, comparing amplitudes until eight or fewer "winners" remain with more amplitude than their neighbors, their neighbors' neighbors, and so on [Example 3b]. This script tracks which frequencies are "winners" slice by slice, and I consider those which "win" in a majority of the slices to fulfill the relative prominence requirement. In the end, the upper partial frequencies I predict a listener will perceive as pitches through spectral fission are those identified by both my predictive model and by my script, and in the following analyses I notate them as diamond noteheads in double-octave treble clefs [Example 3c].

The final chord of Vocal Spectrum's "Aladdin Medley" represents a typical barbershop close spacing for a major triad [Example 4a]. After I apply my two models to its fundamentals, both agree on three candidate frequencies for spectral fission, corresponding to the pitches E6, G#6, and C#7 [Audio Example 3a]. A more literal notated representation of this "ringing" chord as sung, then, could include these pitches [Example 4b]. I will now play just the last two measures of this tag to present this "locked in" chord in context [Audio Example 3b]. A large number of barbershop tags to up-tempo numbers end with a major triad in this spacing, and its pitches of spectral fission offer an explanation as to why arrangers might consistently favor it. In this spacing, spectral fission essentially "doubles" the fifth and the third that couldn't otherwise be doubled with only four singers while also adding a major seventh that increases the chord's sonic punch. Let's listen one more time to appreciate how these pitches contribute to this tremendous sound [Audio Example 3b].

Barbershop chords in wide spacings are less common than those in close spacings, but they do exist, and an examination of one such chord sheds light on why arrangers' choices of spacing depend most of all on the affect of the song in question. The final chord of Ringmasters' tag to "When You Wish Upon a Star" is a representative example, with the doubling one would expect for a major triad but vast distances between each of its fundamentals [Example 5a]. Applying my two models to this chord, the three consensus candidate frequencies for spectral fission correspond to the pitches D6, F6, and C7 [Example 5b] [Audio Example 4a]. Let's hear this "ringing" chord in context [Audio Example 4b]. Similarly to a close spacing, this wide spacing leads to "doubling" of the third and the fifth, though now a third rather than a sixth apart. The most significant difference is the third pitch of spectral fission, here a mellow major ninth instead of a punchy major seventh. It is therefore unsurprising that wide spacings most commonly end barbershop ballads, where "ringing" is still desired but context dictates a gentler sound. As I play this excerpt again, I invite you to consider how its final chord resonates literally and figuratively with Jiminy Cricket's optimistic message of dreams coming true [Audio Example 4b].

Barbershop tags overwhelmingly tend to end on major triads or dominant seventh chords. Nonetheless, barbershop's non-dominant seventh chords are perhaps the most compelling to the ear. The acoustical basis of these chords' different flavor of "ringing" is the complexity of their intervallic content in vertical just intonation. The fundamental frequencies of a just major triad in the closest possible spacing, for instance, can be expressed as 4:5:6, and a just dominant seventh chord can similarly be expressed as 4:5:6:7. A minor seventh chord in such a spacing, on the other hand, has fundamentals in a 10:12:15:18 relationship as typically sung by barbershop quartets, while a half-diminished seventh chord's fundamentals are in 45:66:110:165 relationship. In any spacing, these relationships mean there are fewer overlapping partials, and thus fewer opportunities for singers to tune formants accordingly. In an excerpt from Ringmasters' tag to "She's Out of My Life," the last chord we'll hear is one such half-diminished seventh, for which my models agree on two candidate frequencies of spectral fission, approximately the pitches G#6 and D#7 [Example 6] [Audio Example 5a]. Here's what this moment sounds like in context [Audio Example 5b]. To my ear, the "doubling" of the seventh and the addition of a hard-hitting eleventh to this chord contribute to its hair-raising effect, and though it is not sustained as long as the final chord will be, it more than makes up for its shorter length through sheer viscerality. Let's listen one more time to this extraordinary chord [Audio Example 5b].

Three analytical vignettes obviously cannot represent every "ringing" chord in barbershop music, but they nonetheless demonstrate a type of understanding made possible by treating timbre's upper partial frequencies as possible generators of pitch perception. These analyses also highlight the importance of considering chord spacing in addition to intervallic content in the context of timbre-generated pitch perception. At this point, it bears repeating that my timbral model and digital signal processing script are predictive and descriptive, respectively. Frequencies they agree on correspond to pitches I believe listeners will perceive through spectral fission, but the influence of listener as well as mode of listening on this perception should not be understated. Behavioral experiments are undoubtedly an eventual next step for this theory. In addition, one could easily apply these or similar computational methods to other vocal repertoires, for instance using the Hilliard Ensemble's excellent recordings of Renaissance polyphony as another corpus of a cappella singing in vertical just intonation. Another inviting possibility would be to consider voice-leading between the pitches of spectral fission and its implications for barbershop arranging. And on the technical end, my vowel-neutral model of vocal timbre could be replaced with several vowel-specific models if one had access to several very good barbershop quartets and the opportunity to record and analyze vowels as they sing them both individually and together.

I would like to close with a reflection on the role of notation in this line of inquiry. One could uncharitably conclude that the direct outcome of my exploration of spectral fission in barbershop has been replacing what could be found in published arrangements with representations that more literally reflect what we hear when singers hit these "locked in" chords. To briefly channel the spirit of Harry Partch, far be it from my intention to replace one tyranny of notation with another. I hope instead that my “transcriptions” including pitches of spectral fission have served as a reminder that while notation can be an invaluable resource, it is the smallest piece in a much larger perceptual and acoustical puzzle when we choose to directly analyze relationships between timbre and pitch. The meticulous performance practice of the best barbershop quartets serves as a convenient point of entry for such analysis, but barring sine wave ensembles suddenly taking the world by storm, timbre is by no means exclusive to barbershop. I believe that exploring the boundaries between timbre and pitch is a choice worth making any time our ears remind us how much there is to appreciate and understand by directly considering music as sound.

Bibliography

Averill, Gage. 1999. "Bell Tones and Ringing Chords: Sense and Sensation in Barbershop Harmony." *The World of Music* 41(1): 37–51.

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

Chalika, Magdalene, and Richard Warren. 1994. "Spectral Fissioning in Phonemic Transformations." *Perception & Psychophysics* 55(2): 218–226.

DePaolis, Paul F. 1954. "What Makes the Barbershop 7th Ring?" *The Harmonzier* 14(2): 43–45.

Drown, Colin, and James Cottingham. 2018. "The Barbershop Sound: The Characteristic Timbre of the Male Barbershop Quartet." Paper presented to the Acoustical Society of America, Minneapolis, MN.

Ewin, Jim. 1961. "Barbershop Harmony — A Natural Science." *The Harmonizer* 21(2): 19, 30.

Fales, Cornelia. 2002. "The Paradox of Timbre." *Ethnomusicology* 46(1): 56–95.

Langner, Gerald. 2015. *The Neural Code of Pitch and Harmony*, ed. Christina Benson. Cambridge University Press.

McAdams, Stephen. 1982. "Spectral Fusion and the Creation of Auditory Images." In *Music, Mind and Brain: The Neuropsychology of Music*, ed. Manfred Clynes, 279–298. New York: Plenum Press.

Reagan, Maurice E. 1944. "Mechanics of Barber Shop Harmony." *The Harmonizer* 4(1): 10.

Sethares, William. 2005. *Tuning, Timbre, Spectrum, Scale*. London: Springer-Verlag.

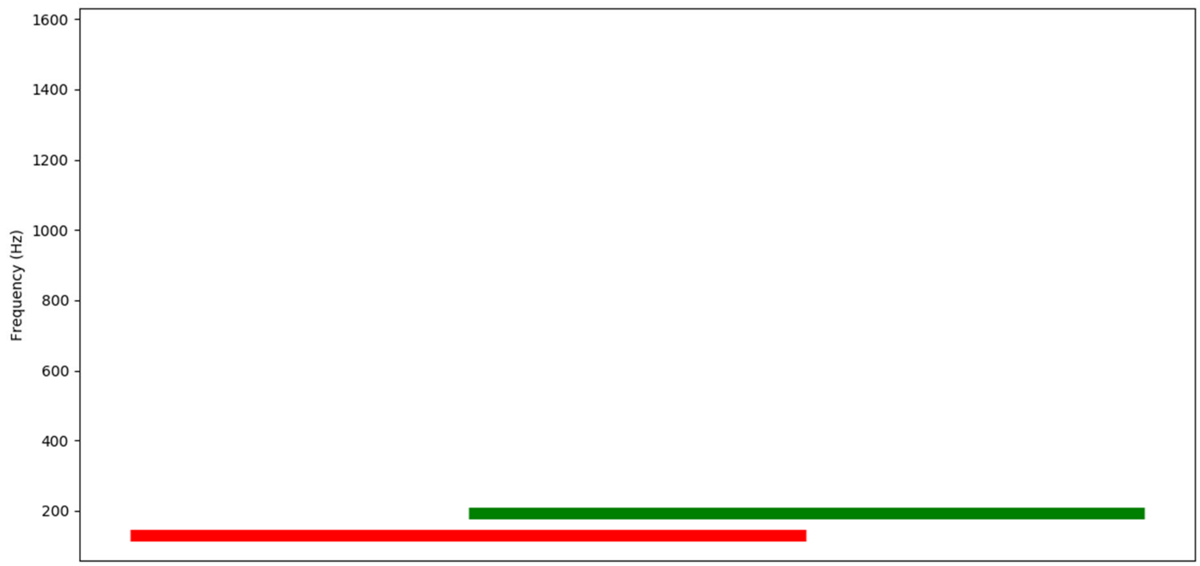
Siedenburg, Kai, and Stephen McAdams. 2017. "Four Distinctions for the Auditory 'Wastebasket' of Timbre." *Frontiers in Psychology* 8:1747.

Siedenburg, Kai, et al. 2019. "The Present, Past, and Future of Timbre Research." In *Timbre: Acoustics, Perception, and Cognition (Springer Handbook of Auditory Research 69)*, ed. Kai Siedenburg et al., 1–19. Cham, Switzerland: Springer Nature.

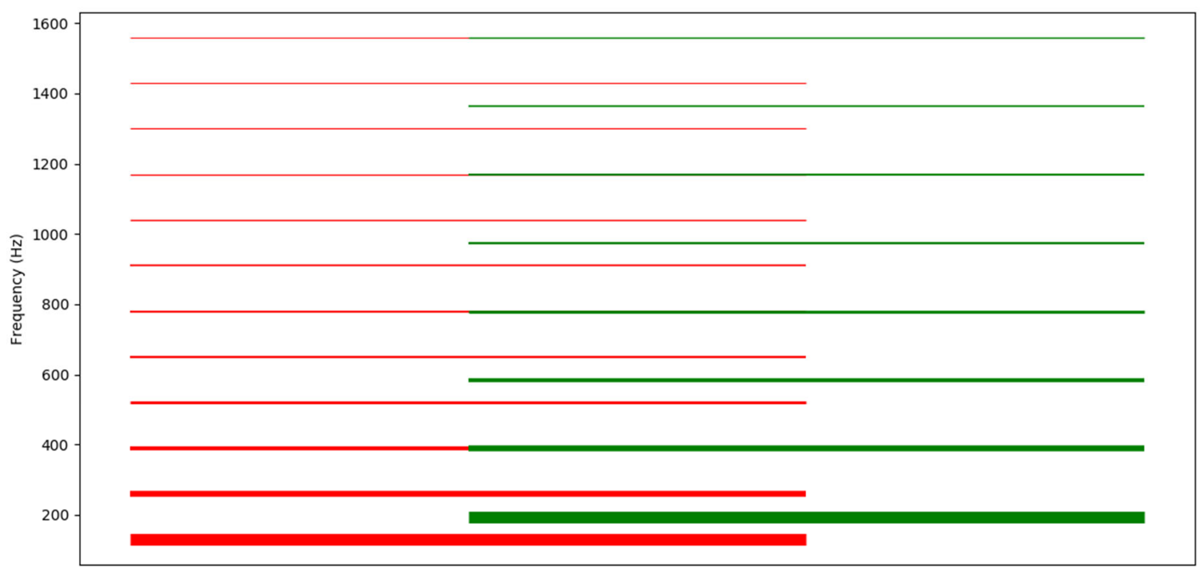
Sundberg, Johan. 2019. "The Singing Voice." In *The Oxford Handbook of Voice Perception*, ed. Sascha Frühholz and Pascal Belin, 117–142. Oxford University Press.

Ternström, Sten, and Gustaf Kalin. 2007. "Formant Frequency Adjustment in Barbershop Quartet Singing." In *Proceedings of the 19th International Congress on Acoustics (ICA 2007)*, 1796–1801. Red Hook, NY: Curran Associates.

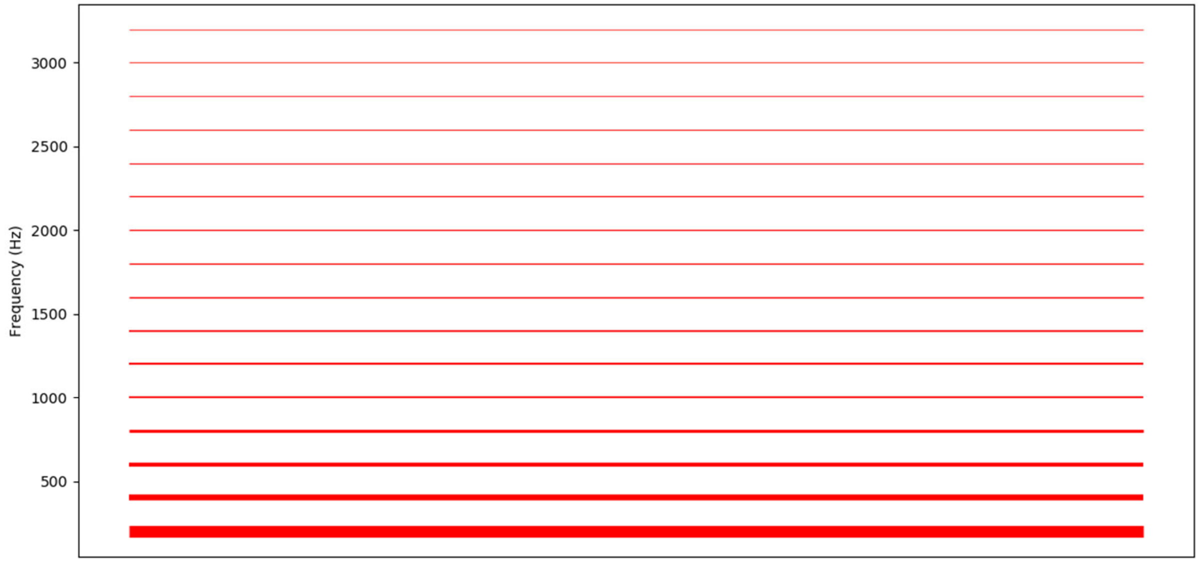
Example 1a. Fundamental frequencies of a just perfect fifth, where red denotes the lower pitch (130 Hz) and green denotes the upper pitch (195 Hz) and 195 Hz : 130 Hz = 3 : 2



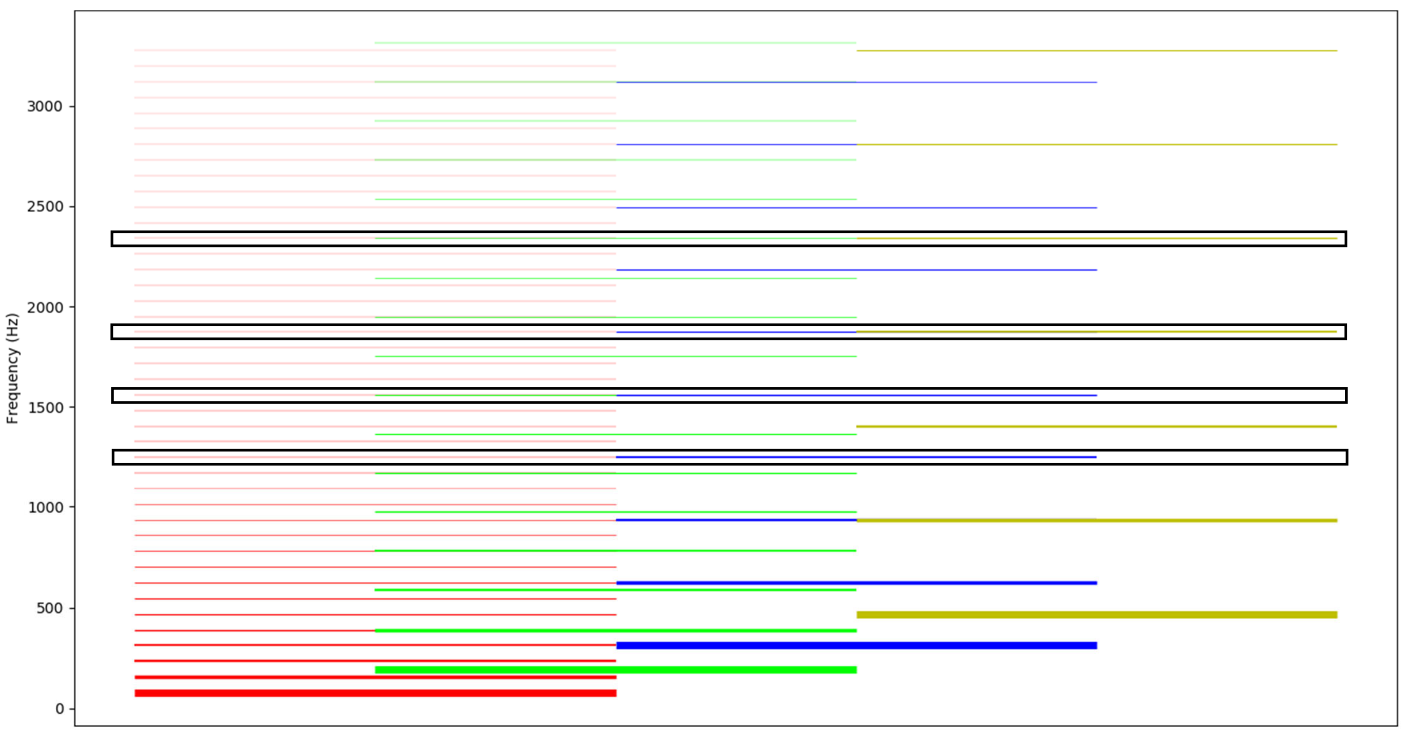
Example 1b. Overlapping upper partials of a sung just perfect fifth, where red denotes the spectrum of the lower pitch and green denotes the spectrum of the upper pitch



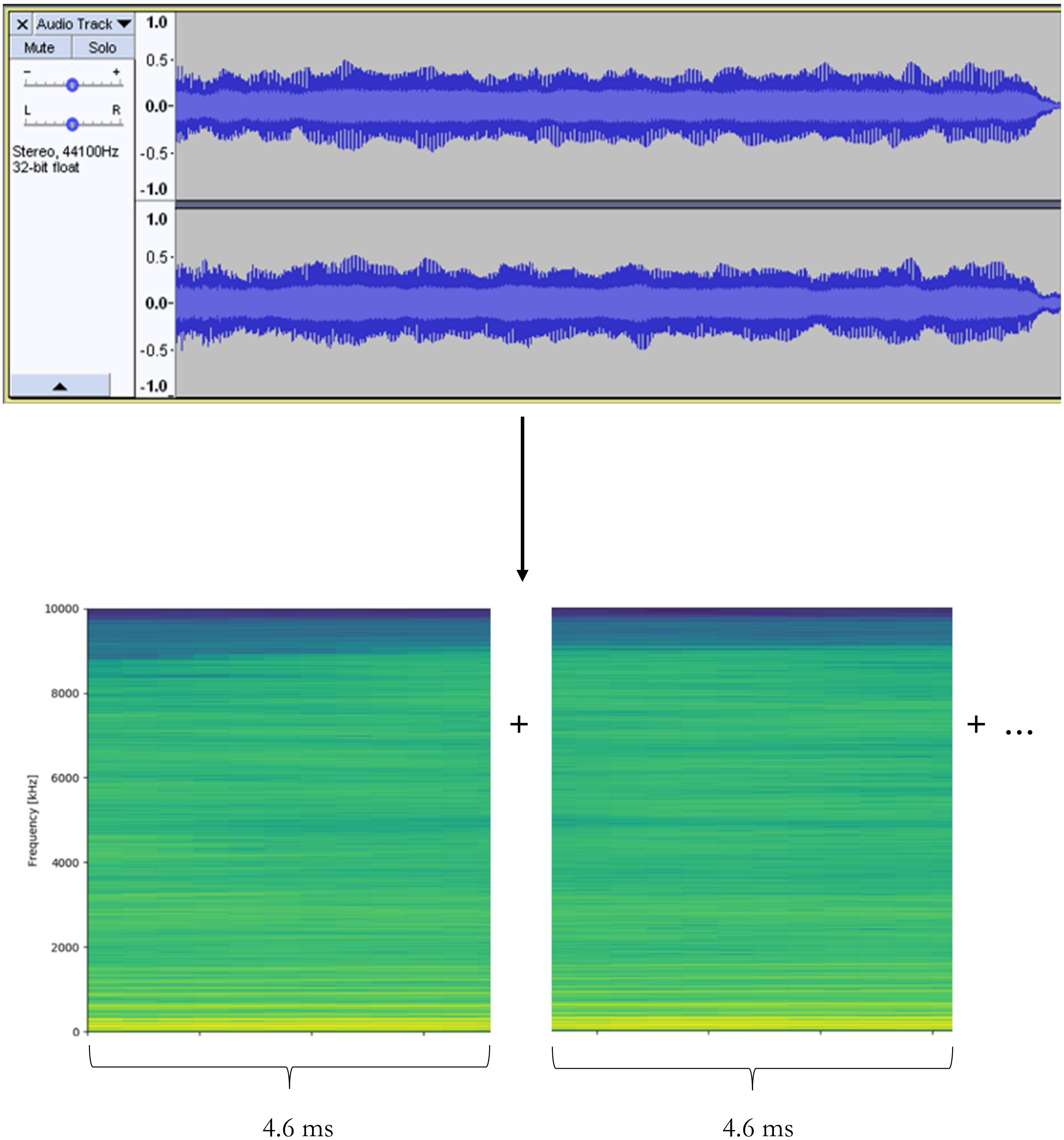
Example 2a. My vowel-neutral predictive model of vocal timbre applied to a fundamental of 200 Hz, where lines of decreasing thickness represent harmonic upper partials (400 Hz, 600 Hz, etc.) of decreasing amplitude (1/2 of the fundamental, 1/3 of the fundamental, etc.)



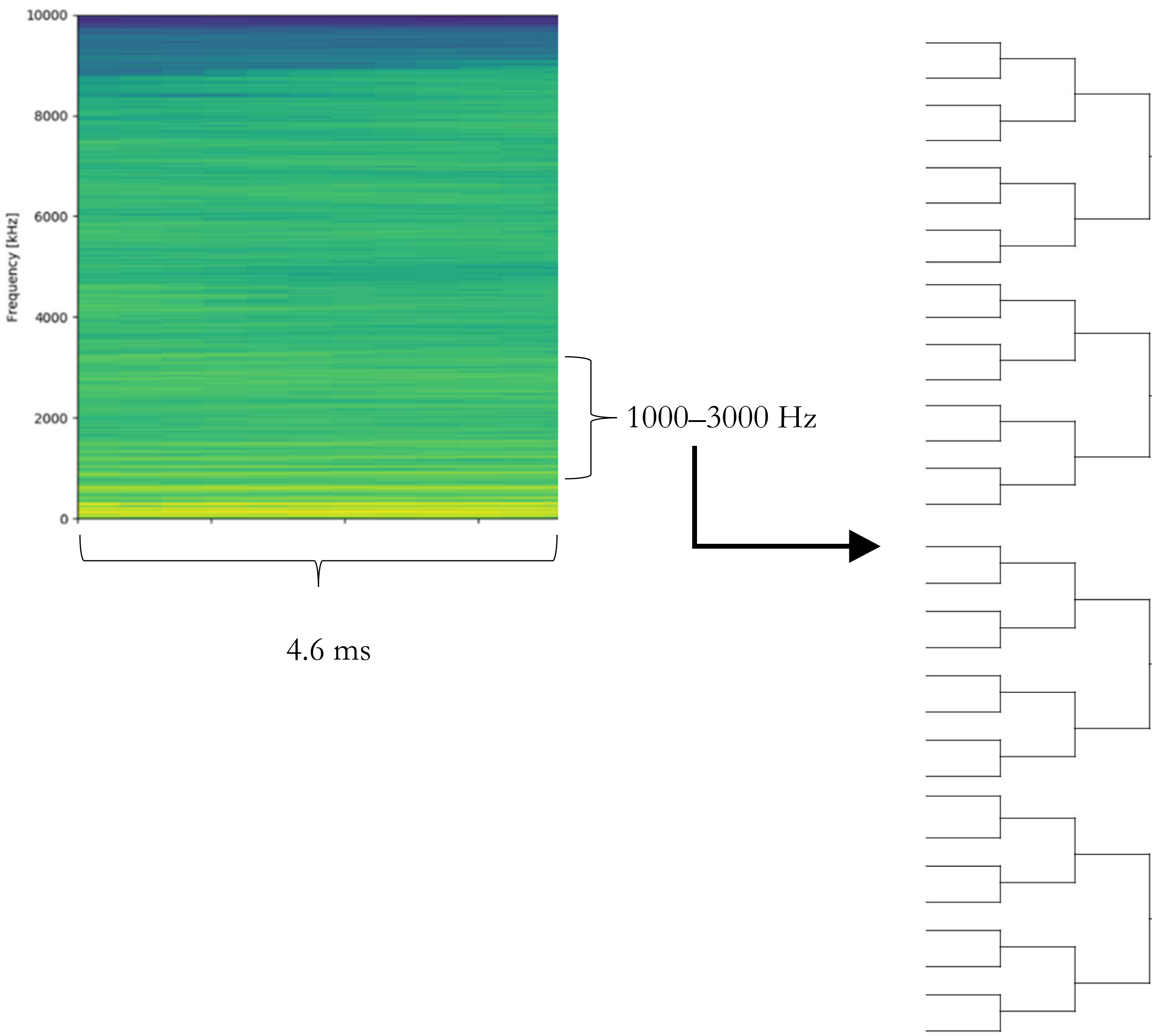
Example 2b. A visualization of my model predicting frequencies with the most amplitude



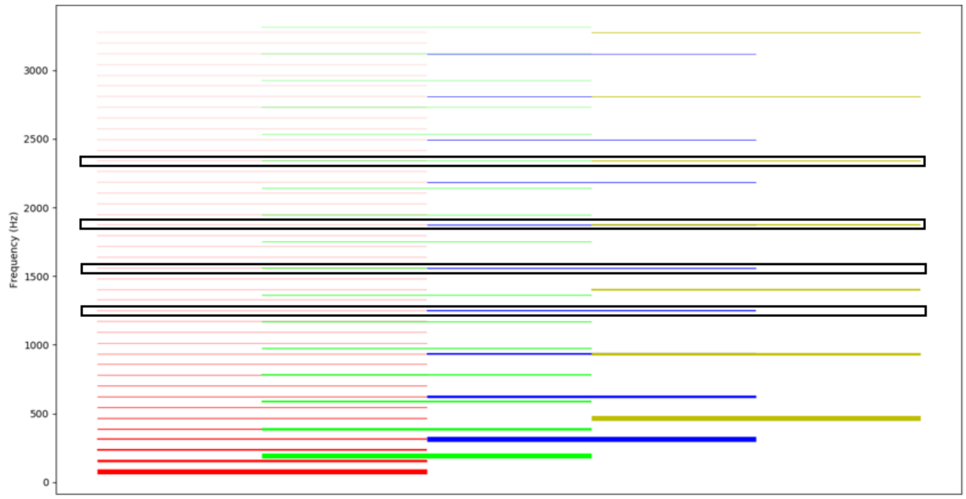
Example 3a. Using the Discrete Fourier Transformation to compute the spectrum of a recording



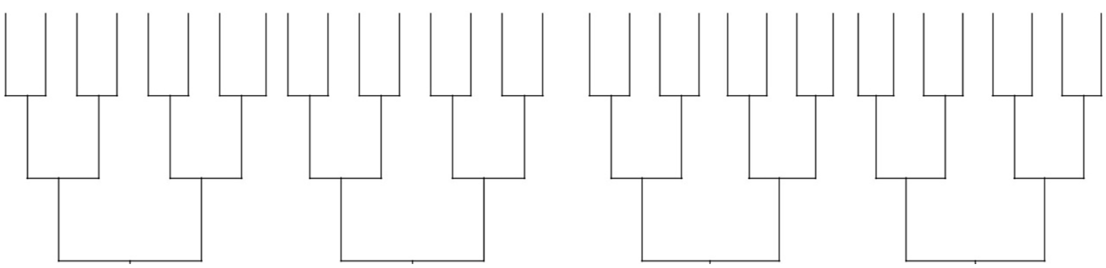
Example 3b. Identifying spectral prominence using a rivalry model



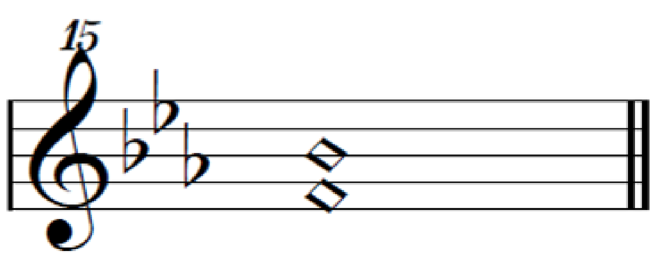
Example 3c. Predicted pitches of spectral fission have frequencies identified by both models



+



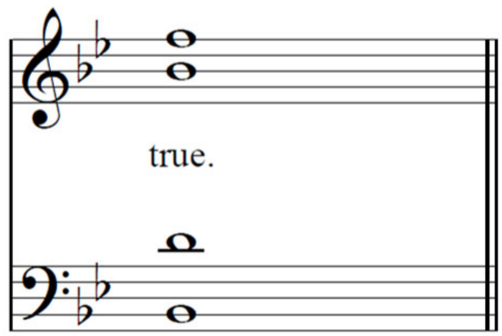
=



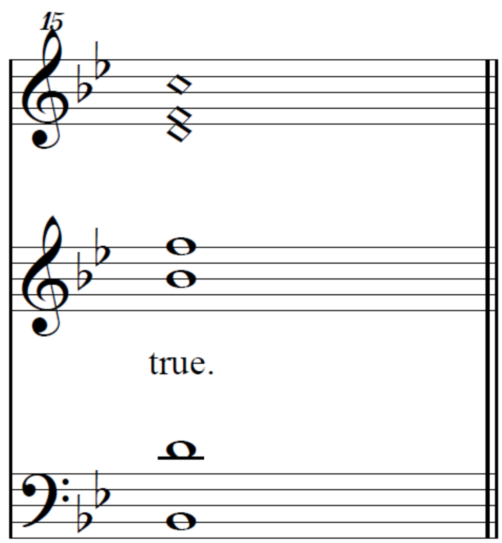
Example 4a. The final chord of Vocal Spectrum's tag to their "Aladdin Medley"

Example 4b. The final chord of Vocal Spectrum's tag to their "Aladdin Medley," including my models' predicted pitches of spectral fission

Example 5a. The final chord of Ringmasters' tag to "When You Wish Upon a Star"



Example 5b. The final chord of Ringmasters' tag to "When You Wish Upon a Star," including my models' predicted pitches of spectral fission



Example 6. The final chord from the penultimate phrase of Ringmasters' tag to "She's Out of My Life," including my models' predicted pitches of spectral fission