
2020

Dynamic Uses of Spectrographic Analysis in Choral Rehearsals and the Voice Studio

Mark McQuade

Valdosta State University, mamcquade@valdosta.edu

Follow this and additional works at: <https://trace.tennessee.edu/jatmi>



Part of the [Music Performance Commons](#), and the [Other Music Commons](#)

Recommended Citation

McQuade, Mark (2020) "Dynamic Uses of Spectrographic Analysis in Choral Rehearsals and the Voice Studio," *Journal of the Association for Technology in Music Instruction*: Vol. 1 : No. 1 , Article 1.

Available at: <https://trace.tennessee.edu/jatmi/vol1/iss1/1>

This article is brought to you freely and openly by Volunteer, Open-access, Library-hosted Journals (VOL Journals), published in partnership with The University of Tennessee (UT) University Libraries. This article has been accepted for inclusion in *Journal of the Association for Technology in Music Instruction* by an authorized editor. For more information, please visit <https://trace.tennessee.edu/jatmi>.



Dynamic Uses of Spectrographic Analysis in Choral Rehearsals and the Voice Studio

Dr. Mark McQuade
Associate Professor of Music
Valdosta State University

Abstract

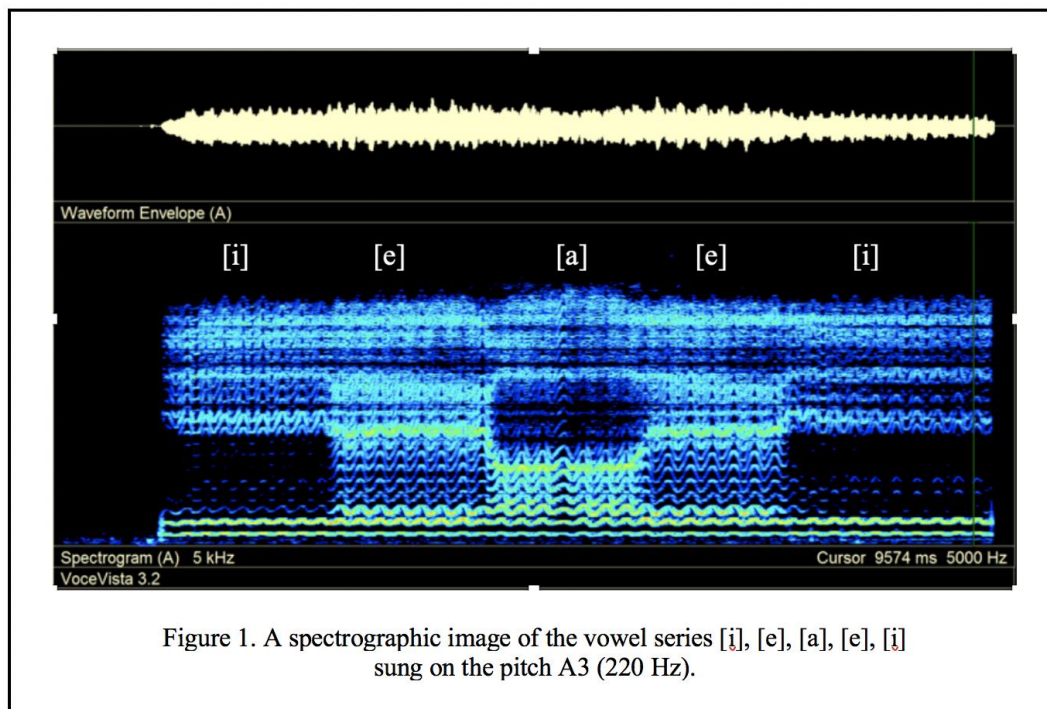
Spectrographic analysis is a powerful tool that allows students to see a visual representation of the acoustic properties of their instruments. For today's technologically savvy students it brings an exciting new element to music making that will allow them to grow musically, artistically, and technically even faster. With a computer, microphone, and screen (or even just a smart phone) teachers can open their students' eyes to see what their ears have always been hearing. By turning an aural/oral art into a visual experience, students will gain a deeper understanding of their instruments and how they can use those instruments efficiently and expressively. The application of spectrographic analysis applies to both choral ensembles and vocal soloists, and while this article will focus on its use with the singing voice, spectrographic analysis also has benefits for instrumentalists as well. To understand the practical applications, we will examine a variety of vocal and musical skills that can be explored through this technology.

What is Spectrographic Analysis?

A spectrogram provides a visual representation of the acoustic properties of the singing voice. After capturing the sung tone through a microphone, the spectrograph generates an image of that tone (the spectrogram), as seen in Figure 1. In essence, the spectrograph serves as a video camera for sound. Using the captured images to study the acoustic and practical aspects of sung tones allows for deeper insight into the art of singing and provides a useful pedagogical tool.

Spectrographic analysis is not new. Acousticians and voice scientists have used it as a research tool for decades, and in fact, it has been used as a diagnostic tool relating to the singing

voice since the 1950s. In his 1958 *NATS Bulletin* article titled, “Vowel Color and Voice Quality - An Acoustic and Articulatory Comparison,” the linguist and experimental phonetician Pierre



Delattre used spectrographic analysis as a measure of “vowel color.” It was not until the 1990s, however, when championed by internationally-renowned voice teacher Richard Miller, that spectrographic analysis started to break into more main-stream voice pedagogy. Furthermore, it was well into the 21st century when spectrographic analysis became common-place in voice studios around the world. Miller’s *The Structure of Singing* (R. Miller, 1986) paved the way for other authors’ works such as *Your Voice: An Inside View* (McCoy, 2004), *Resonance in Singing* (D. Miller, 2008), *Practical Vocal Acoustics* (Bozeman, 2013), and numerous articles in the *Journal of Singing* that all demonstrated the scientific and practical uses of spectrographic analysis as it related to the solo singing voice.

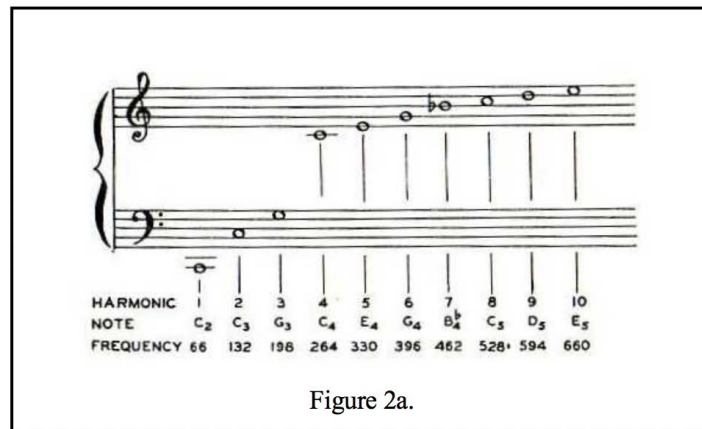
There is a significant body of literature in the voice pedagogy world, both articles and books, that deal with and discuss specific applications of spectrographic analysis for the soloist. One large piece of the vocal art that has been seemingly neglected in the cannon of spectrographic literature is the application of spectrographic analysis to choral singing. With the exception of one feasibility study conducted about chorister perceptions of spectrography

reported in the *International Journal of Research in Choral Singing* (Nix, Mabry, Matthews-Mutwill, 2007) and a dissertation, “Spectrographic Analysis of the Acoustic Properties of Selected Vowels in Choral Sound” (Hunt, 1970), there is a decided lack of literature on the subject of using spectrographic analysis in the choral setting. Therefore, this article explores some of the practical applications that have potential benefit for choral directors and their ensembles, along with solo singers and teachers of voice.

Spectrogram technology allows one to take an art form that is both *oral* and *aural* and make it a visual experience. For today’s tech-savvy students who are always “plugged in,” this technology can bring a stimulating new element to music making that allows them to grow musically, artistically, and vocally even faster. It also provides quantitative real-time feedback for teachers and students, which is invaluable as they try to grow as an ensemble and as individual artists. (As one gains familiarity with its functionality and interpretation, this technology can also be used as a qualitative reference for certain vocal maneuvers.)

Before considering spectrographic images and what they mean, it is important to first establish or be reminded of some acoustic fundamentals as they relate to the singing voice. When the vocal folds are set into vibration, they emit a complex tone which is made up of a fundamental frequency and its overtones, also known as harmonics or partials. These overtones are multiple integers of the fundamental frequency. For example, the pitch A440, which is A4 or the A above middle C on the piano, has a fundamental frequency of 440 Hz. When that pitch is sung, the frequency 440 is present, but so are 880, 1320, 1760, 2200 and so on, as you continue to add 440 Hz to each previous overtone frequency.

Figures 2a and 2b display the distribution of overtones in the modern Western equal tempered tuning system. The first chart displays the overtone series for the pitch C2, while the second lists the overtone series for the pitch G2 (*frequencies in figure 2b are rounded to the nearest hundred*).



Musical Harmonic Series	
1600 Hz	G
1500 Hz	F#
1400 Hz	F
1300 Hz	E ^b /E
1200 Hz	D
1100 Hz	C#
1000 Hz	B
900 Hz	A
800 Hz	G
700 Hz	F
600 Hz	D
500 Hz	B
400 Hz	G
300 Hz	D
200 Hz	G
100 Hz	G

Figure 2b.

The sequence of overtones is relationally the same for both pitches. The series starts with the fundamental frequency. Then the first overtone is an octave higher than the fundamental frequency. Overtone 2 is an octave plus a perfect fifth higher. Overtone 3 is two octaves higher than the fundamental. Then the next is two octaves plus a major third higher, followed by two octaves plus a perfect fifth higher, then two octaves plus a minor seventh, and so on. All this is to show that when one looks at a spectrographic image of a sung tone, many lines will be visible scrolling across the screen, and these lines correspond to the frequencies and overtones that the singers' voices are producing.

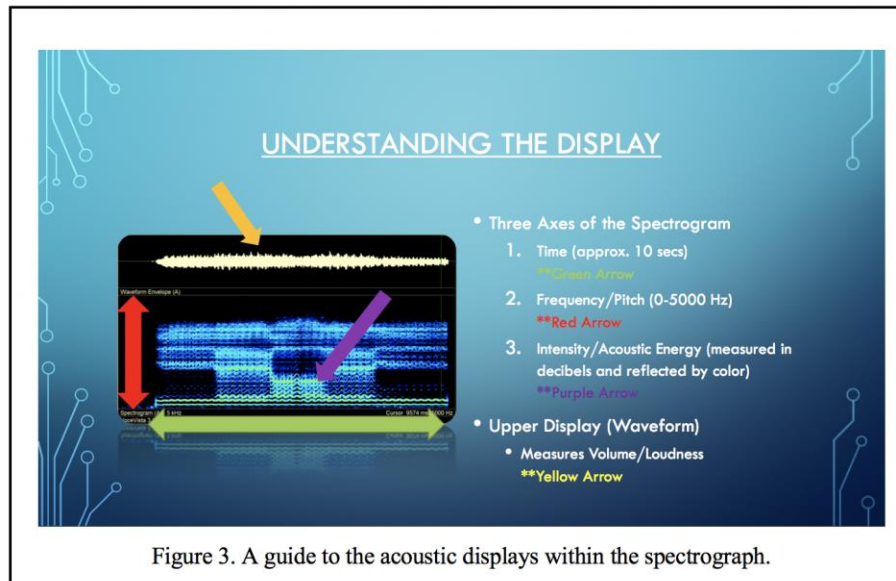
As a point of reference, it was previously mentioned that the voice emits a complex tone, but not all sound sources produce such a complex tone. A tuning fork, for example, emits a pure tone – one single frequency with no overtones.

In addition to the idea that the human vocal folds produce a complex tone, one must also understand that the vocal instrument does not stop at the laryngeal level. The vocally produced tone is filtered through the vocal tract (which is comprised of the larynx, the pharynx, the oral cavity or mouth, and sometimes the nasal cavity) and the sounds that are perceived as pitch and timbre result. The vocal tract is the resonator for the voice, and thus, any alterations in the resonator shapes, created by the singer, will result in variations in spectral images. As certain harmonics are filtered (or not filtered) by the resonator shapes, the harmonics appearing on the spectrogram will vary in intensity.

Understanding the Display

With these acoustic ideas in mind, it is time to turn to the spectrographic display. Figure 3 provides a guide to the acoustic information that may be viewed when used with choirs and voice students. The first thing to recognize is that most spectrographs have three axes: 1) Vertical, 2) Horizontal, and 3) Color. The horizontal axis represents time and the images scroll from right to left on the screen. Depending on the specific spectrogram one uses, the amount of time displayed will vary and some can even be manually adjusted. In the Voce Vista¹ software that was used for this study, the images last approximately eight seconds before they scroll off into cyberspace. The vertical axis represents frequency, with the lower frequencies starting at the bottom of the display and ascending to higher frequencies at the top of the display. The image in Figure 3 shows a range of approximately 0 to 5000 Hz. Much like the time settings, some spectrograms allow the user to adjust the viewable frequency range to fit their specific needs.

¹ The software used in this study was Voce Vista Pro 3.2, c. 2002. The most recent version of this software is called Voce Vista Video and may be purchased or licensed at <http://www.vocevista.com/purchase-vocevista-video-for-mac-and-pc/>.



The color axis represents acoustic energy or intensity (as measured in decibels). Weaker acoustic signals are indicated by darker/warmer colors such as purples and blues, where stronger or more intense acoustic signals appear brighter in color, such as yellow or red.

The final part of the display in Figure 3, is actually not part of the spectrogram itself, but is a supplementary tool that works in conjunction with the spectrogram and that many users find quite valuable and easy to understand. The upper window provides a *waveform* display that measures loudness or volume. It is quite simple in that the louder one sings, the wider the waveform grows. And inversely, the softer one sings, the slimmer the waveform becomes. This can be a fun way to explore dynamics visually with vocal ensembles.

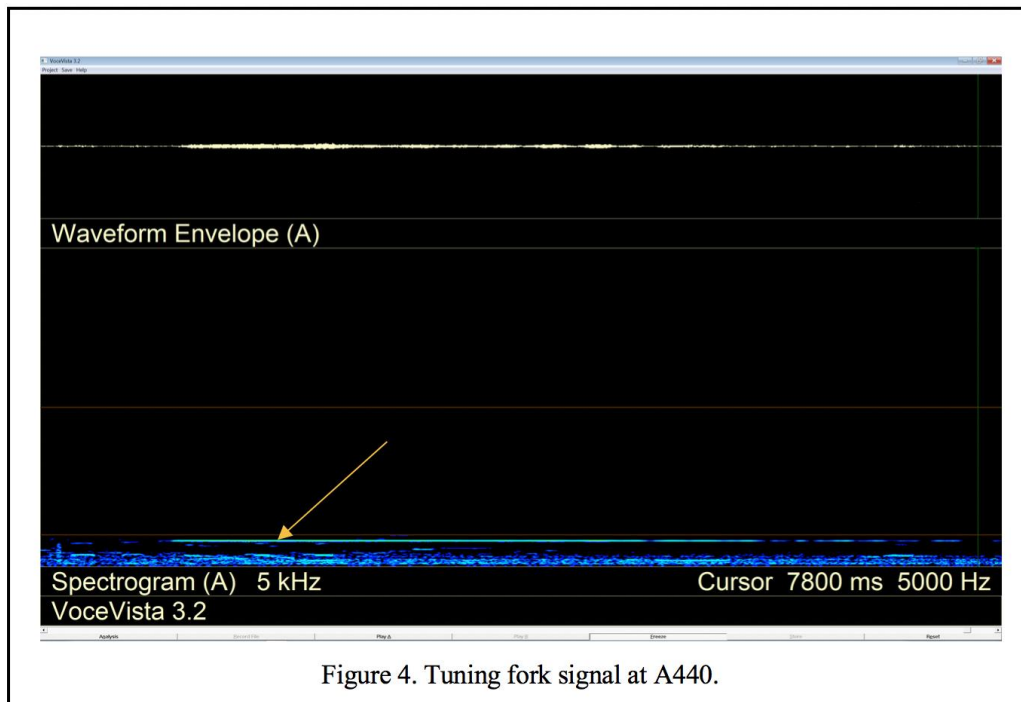
Seeing Our Voices

Voce Vista is a robust spectrographic analysis tool that provides very clear images. The latest version of the software, Voce Vista Video, can be purchased and downloaded from the Voce Vista website² for either Mac and PC. (This is an exciting new option, as previous versions of the software were only compatible with PC operating systems.) When speaking or singing commences, the software creates a spectrogram that registers phonation through microphone input and transforms those sound waves into graphic images. Because the

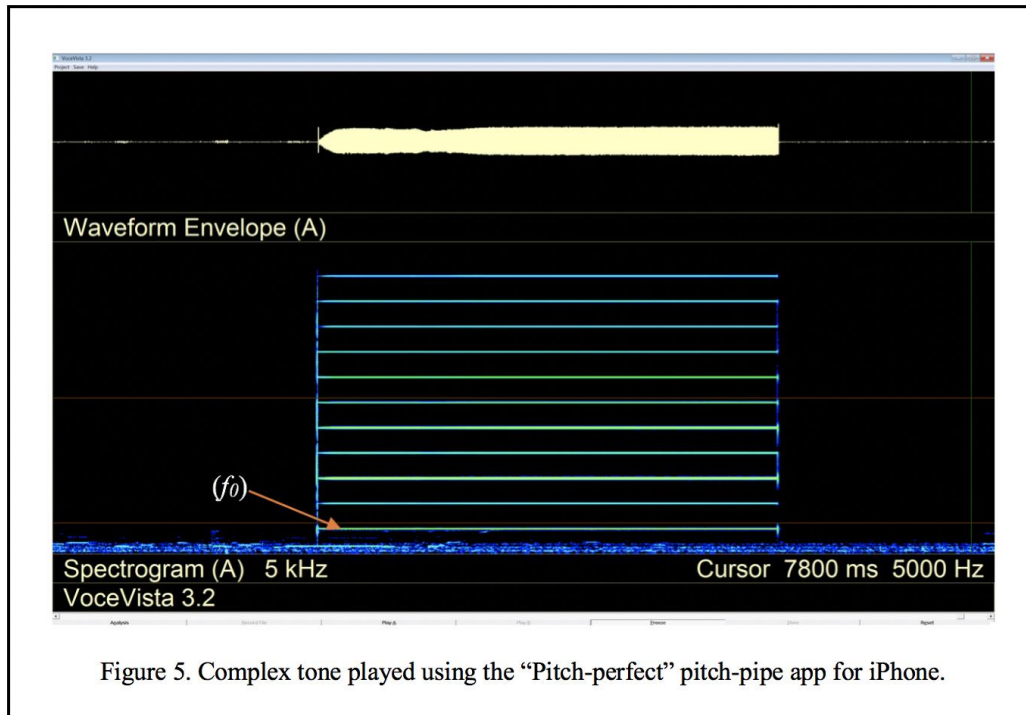
² <http://www.vocevista.com/purchase-vocevista-video-for-mac-and-pc/>

spectrogram will register any audible sound, it is best to sing in a space with as little background/ambient noise as possible in order to keep the spectral image as pure as possible for the singer(s).

Figure 4 demonstrates the image that occurs when a tuning fork, tuned to standard concert pitch, A440, is played and the pure tone emerges. A single thin green frequency line is present at the 440 Hz level. (The thicker/fuzzy-looking section below the line is ambient room noise being picked up by the microphone.) Figure 5 uses the same pitch but builds upon this single frequency by displaying a very basic complex tone. The image was captured from sound played using the “Pitch Perfect”³ pitch-pipe app on an iPhone. This screen clearly conveys the overtone structure as displayed spectrographically.



³ DepollSoft, 2018.



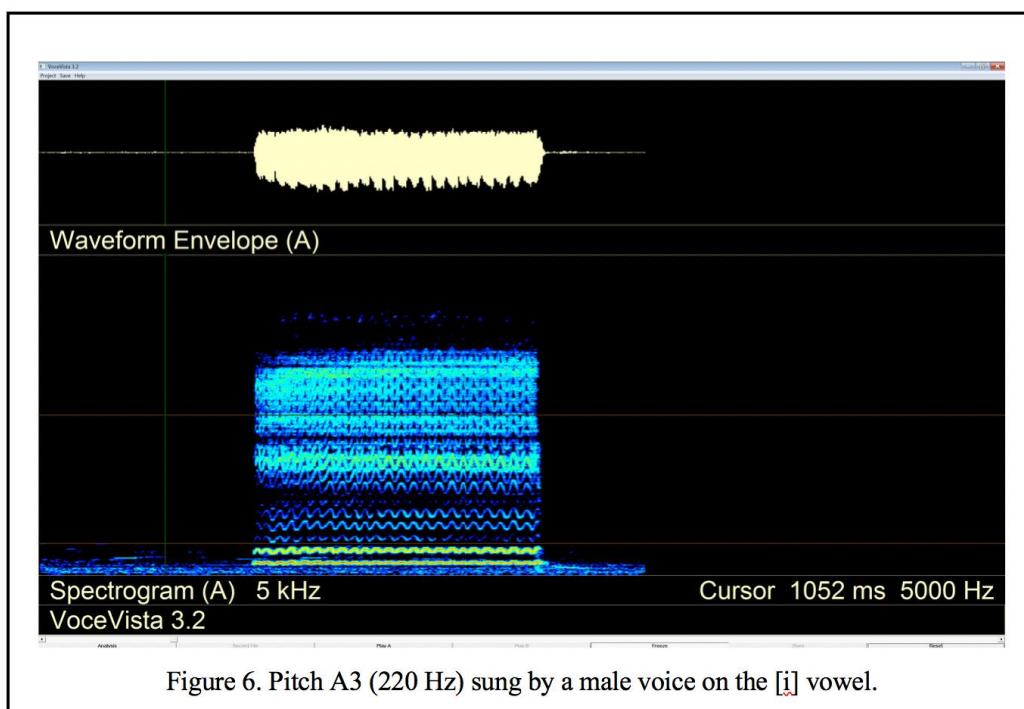
The fundamental frequency (f_0) in Figure 5 is indicated by the lowest line on the display (f_0). Just as with the tuning fork, the pitch A4, with a frequency of 440 Hz, was played. Above the fundamental frequency one notices the overtones appearing every 440 Hz. This “bare-bones” tone has no fluctuation in pitch, the harmonics are quite evenly spaced, and they exhibit a rather uniform intensity with only minor variations in color.

Figure 6 is a sharp contrast to Figure 5 and displays a complex tone sung with an [i] vowel by a male voice on the pitch A3 – the A below middle C, which has a frequency of 220 Hz.

Why does the sung phonation look so much different than the pitch-pipe tone? The first difference to note is that of pitch. Because the sung tone was an octave below the pitch-pipe tone, there are twice as many harmonics/overtones present in the sung tone. Therefore, more lines are present in the image. The singer also used a vibrant tone; he was singing with vibrato. Allowing vibrato to happen creates a fluctuation in pitch above and below the fundamental frequency, and that pitch fluctuation results in the “bumpy” or “squiggly-looking” lines that are

seen throughout the spectrum. Another difference between the pitch-pipe tone and the sung phonation are the variations in color (or acoustic energy) within the harmonics. Where the pitch-pipe tone had a rather uniform energy level among its harmonics, there are obvious areas within the sung tone that have significantly greater levels of acoustic energy than others.

There is a great deal more that could be seen and discussed in Figure 6, but the basic acoustic knowledge and understanding of the spectrographic display that was just outlined will serve as the point of departure for the application of this technology to choirs and soloists.



Practical Applications

Anytime something new is introduced to an ensemble or student, it is a good idea to make it fun. Students will buy into it if they think it is going to be enjoyable. Spectrographic analysis could be introduced with a simple warm-up exercise such as a siren, which will give the students the opportunity to observe their ascent and decent through the range in a somewhat whimsical way. Figure 7 demonstrates a series of sirens that systematically rise in pitch. The

sirens were sung on an [u] vowel by a male using a falsetto cheer sound that one might hear at a sporting event to recognize an outstanding play.

The spectrogram in Figure 7 graphically depicts the rising and falling action of the sirens. For those students who have difficulty or may not hear if they are actually raising the pitch higher, or for those who feel like they are not able to sing higher, this tool gives them a way to see what they are doing and provides another avenue to achieve their goal. If singers can watch the pitch ascend and descend on a spectrogram while they are physically producing the tone, and then connect the feelings they experience with the images they see, it gives them something much more tangible than simply an aural sound they are hoping to create. Much like the old

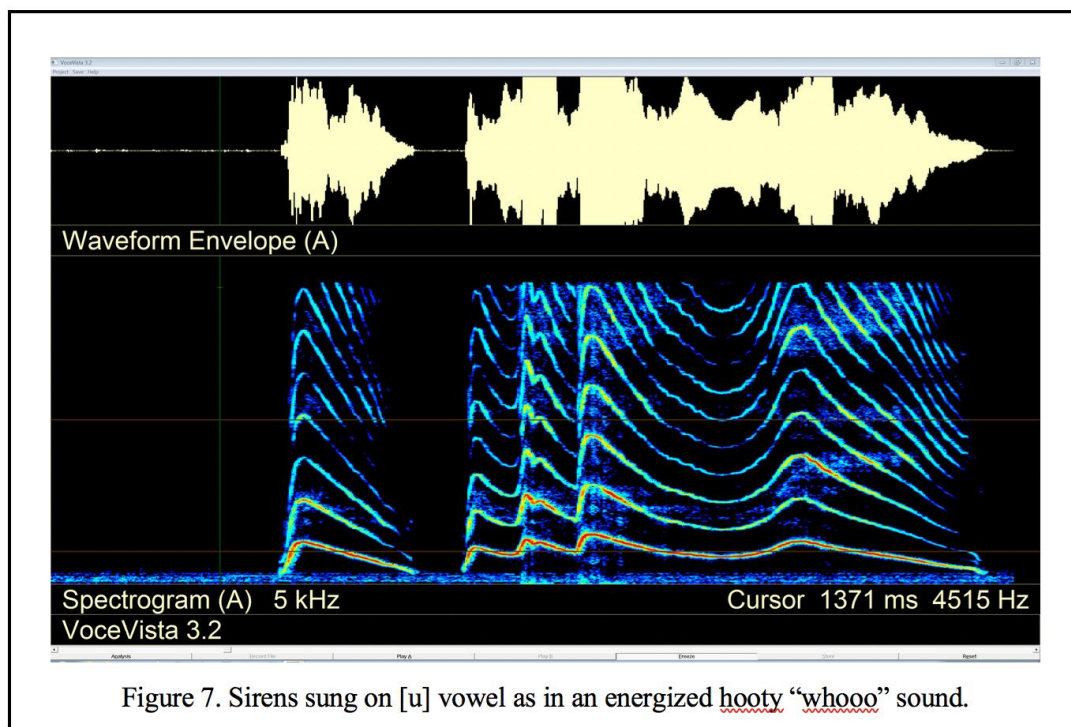
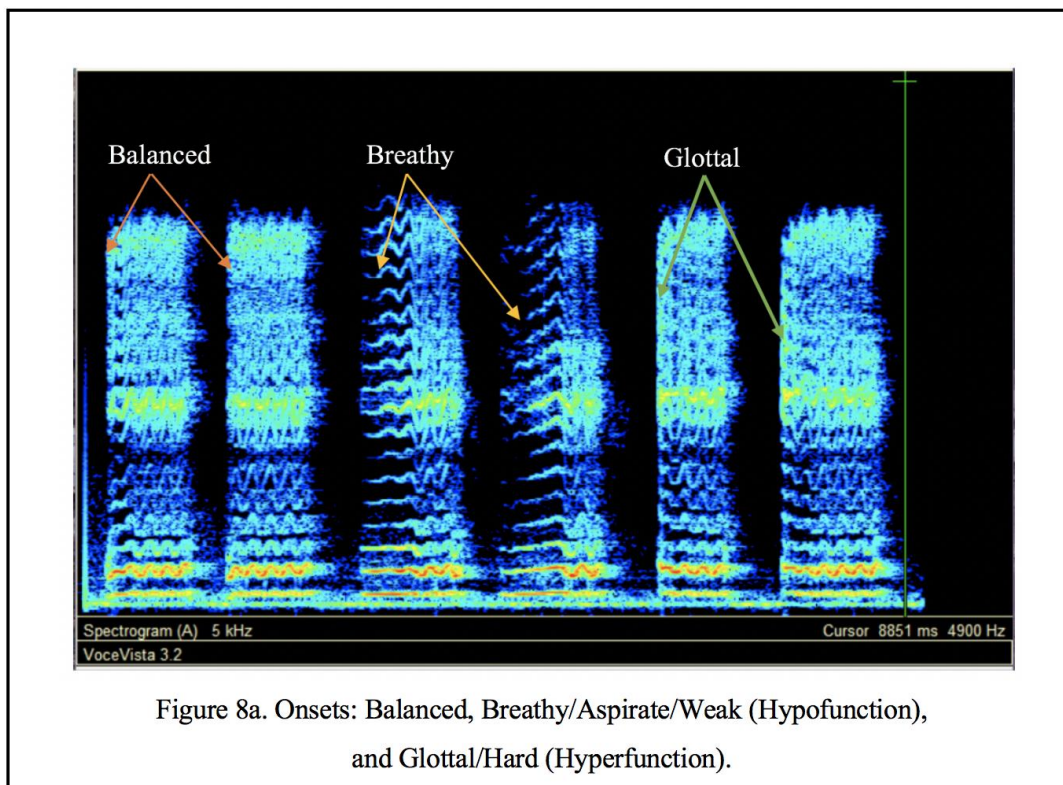


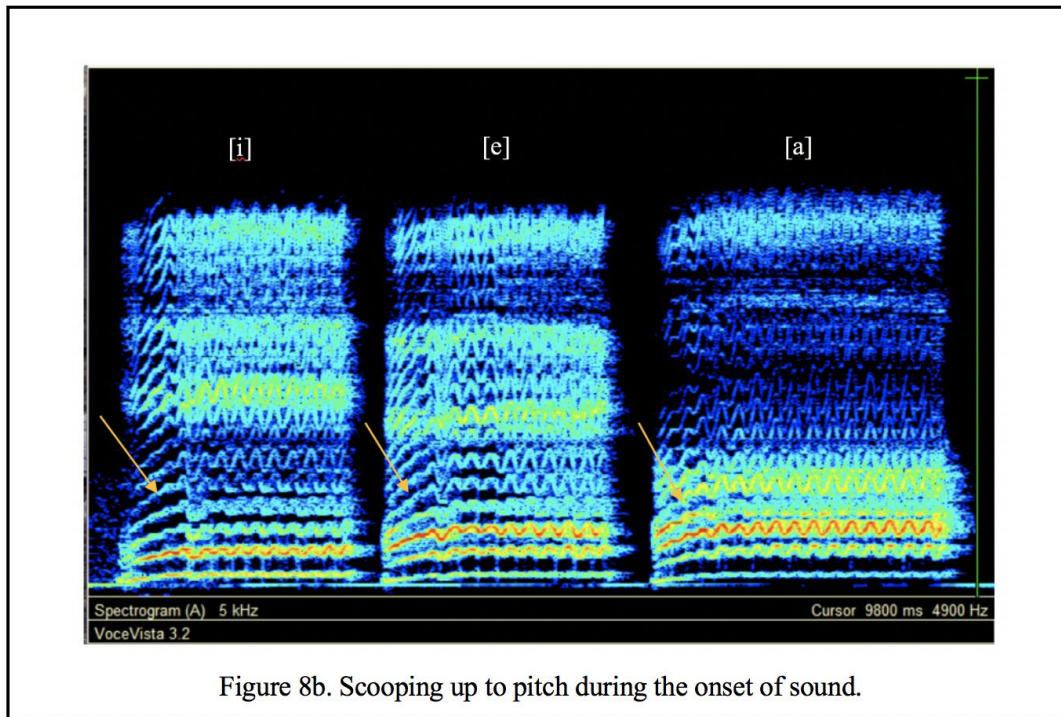
image of throwing darts at a dart board while in the dark, seeing a targeted sound (or vocal maneuver) on a screen often makes it more real for students than simply aiming for a sound they are supposed to imagine and bring into existence. Once singers have had some fun seeing what their voices look like and how they can “paint with sound,” it is time to get down to work.

Every sung tone must have a beginning and an end. In the singing world, these are called onsets and releases. The quality and uniformity of both onsets and releases are vital components of artistic singing. Every choral conductor in history has had to, at one time or another, instruct

their ensemble to “watch” in order for all to enter as one voice with one clear onset of sound. Also, how many voice teachers have ever had to ask a student to start in the center of the pitch, rather than scooping up to it? Most likely, all! Figure 8a illustrates how various types of onsets and releases might look to singers who are watching a spectrographic image scroll by during their rehearsal or lesson, and Figure 8b displays what would be seen if singers scooped up to a pitch.

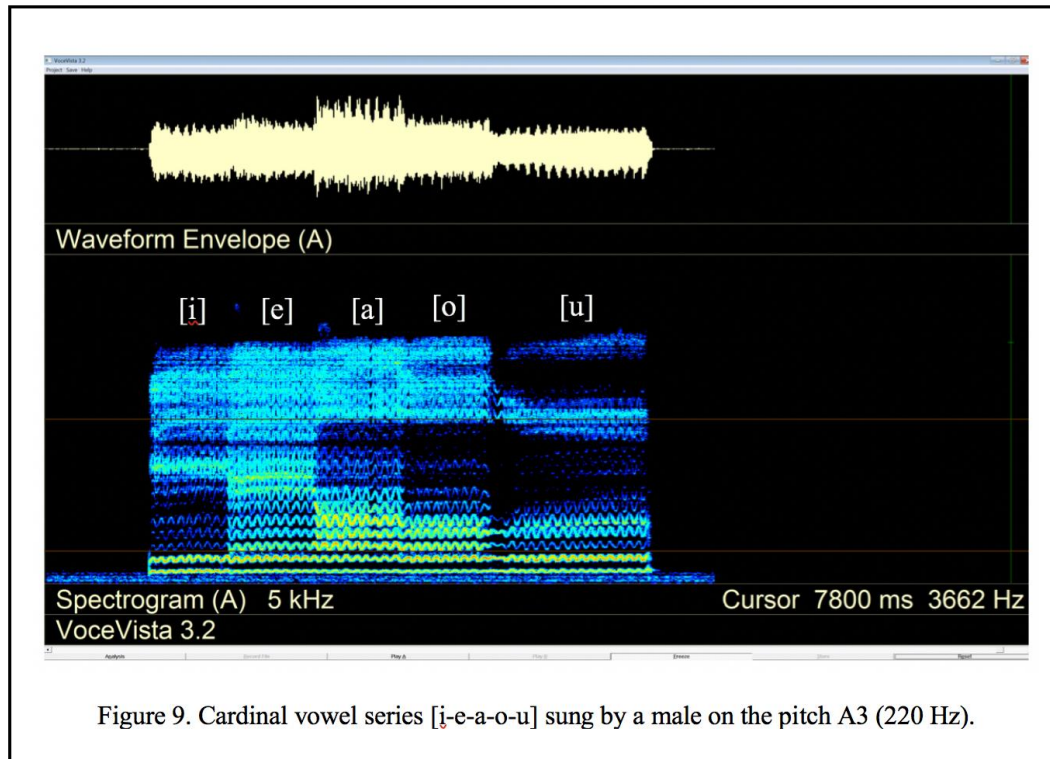
In Figure 8a, three different types of onsets are visible. Each onset is sung twice using the [i] vowel on the pitch A3. The first is a balanced onset that displays a clear beginning with an immediacy of sound and a clean release at the end of the phonation. The second onset is breathy, as noted by the lack of energy and the way in which the sound “creeps” into a fully vibrant tone. The final onset is a glottal or hard attack. This is typically distinguishable by the solid line that forms between the harmonics down the left side of the phonation and usually starts with a downward motion of the vibrato cycle as a result of the vocal folds being pressed together and then blown apart with air.





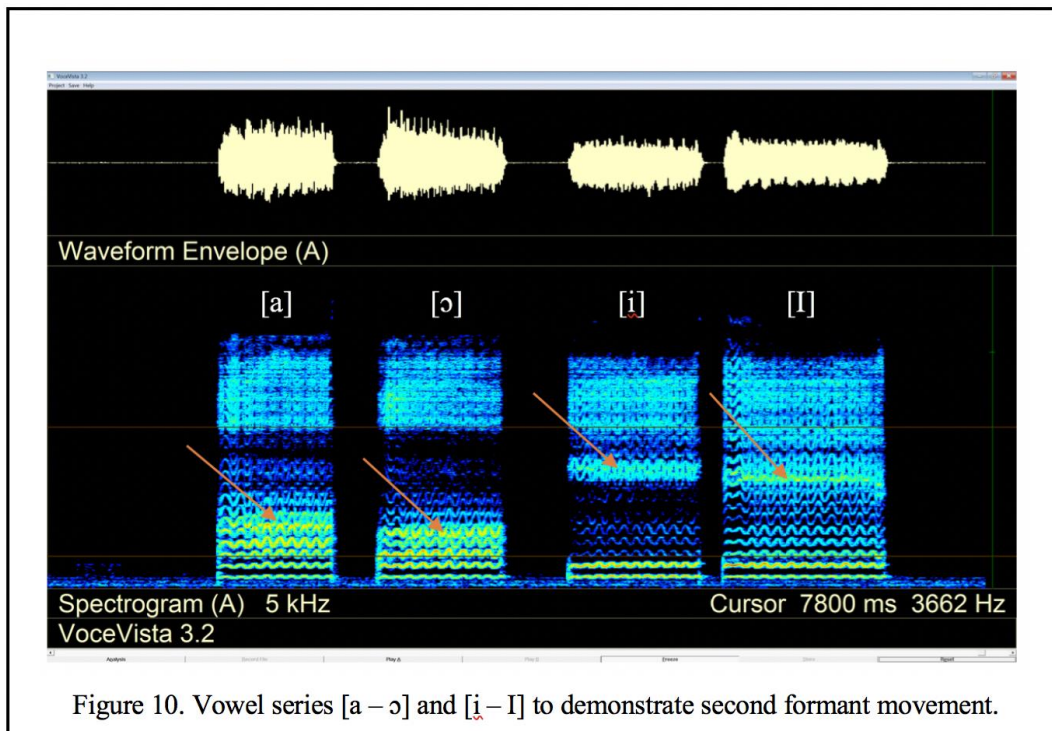
While scooping and sliding between notes may be an acceptable gesture in certain genres of vocal music, in most choral and classical vocal music settings, it is usually less desirable. The importance of singers starting in the center of each pitch cannot be overstated. The spectrographic images in Figure 8b vividly depict what happens when singers approach the pitch A3 from below. The yellow arrows indicate the upward sliding action of phonations all aiming for the same pitch on three different vowels – [i], [e], and [a].

Another task that takes up copious amounts of time in choral rehearsals and in the voice studio is finding appropriate vowel shapes and then making those shapes consistent. In the acoustic fundamentals section above, a conscious choice was made not to delve into vocal formants (which are resonance potentials of the vocal tract) because they are part of a larger resonance discussion that lies outside the parameters of this article. However, there is one formant that deserves mentioning here as it plays an important role in vowel tracking.



In the center of the spectrographic image presented in Figure 9, there appears to be a staircase descending from left to right. This downward descent results from the singer's changing tongue position inside the vocal tract as determined by the chosen vowel series.

The basic rule of thumb is that the more a singer closes the vowel (or the more lateral the vowel becomes) the higher the second formant will rise, and the more a singer opens the vowel (makes the mouth taller or the lips rounder), the lower the second formant will go. Figure 10 demonstrates the subtle differences of second formant position that may be seen when slight changes are made to a single vowel, such as the contrasting pairs of [a – ɔ] and [i – I].



By singing a bright [a] vowel (as in the English word *father*) and its more rounded pair, the [ɔ] vowel (as in the English word *thought*), one can see the affect that the lip rounding required for moving from [a] to [ɔ] has on the spectrographic image. This movement results in the second formant moving to a slightly lower position. The same downward motion is a noticeable response to the mouth opening (i.e. a more vertical position when the jaw lowers) required to move from the closed [i] vowel (as in the English word *see*) to its more open partner, [I] (as in the English word *his*). When students make these different shapes and they see the images change before their eyes, they will start to more consciously hear and feel the differences while they are producing the different sounds. This process then supports the student's development of their aural conception to actively choose the various sounds that a conductor/director/teacher desires in order to accomplish their goals of timbre, color, and balance.

Once students achieve vowel shapes that produce a desired timbre, it is important to capitalize on those shapes and work systematically to reinforce and retain them for vowel consistency. Students must develop recall strategies so that they can find a shape and timbre at any moment. Figure 11 illustrates how spectrographic analysis can help reinforce vowel

consistency. A repeated vowel sequence ([i-a-i-a-i-a-i]) was first sung maintaining a uniform position for all of the corresponding vowels—each of the [i] vowels matched and each of the [a] vowels matched. In the second sequence of Figure 11, certain vowels were intentionally changed to imitate what might happen when consistent vowel choices are neglected. The pattern started identically to the first sequence with [i] and [a], but as can often happen with young singers, those vowels morphed into [ɪ] and [ɔ] before reverting back to [i], [a], and [i] to finish. Diversions such as these would clearly show the singers the evolution of their vowels and then could be addressed by the conductor/teacher. The students and conductors could work together to bring the vowels back to consistent uniformity.

Another prominent adversary to vowel consistency is the diphthong. Diphthongs pose a problem because, if not given specific consideration, they will happen too soon, last too long, or even be added into words where they should not be. Because diphthongs are created by combining two different vowels, the change from one vowel to the next can easily be tracked on a spectrogram. A diphthong produces a rather striking image. Commonly problematic diphthongs are [eɪ] as in the English word *they*, [aɪ] as in the English word *my*, and [ɔɪ] as in the English word *boy*. The transition and duration of these dastardly duos are easily discerned spectrographically.

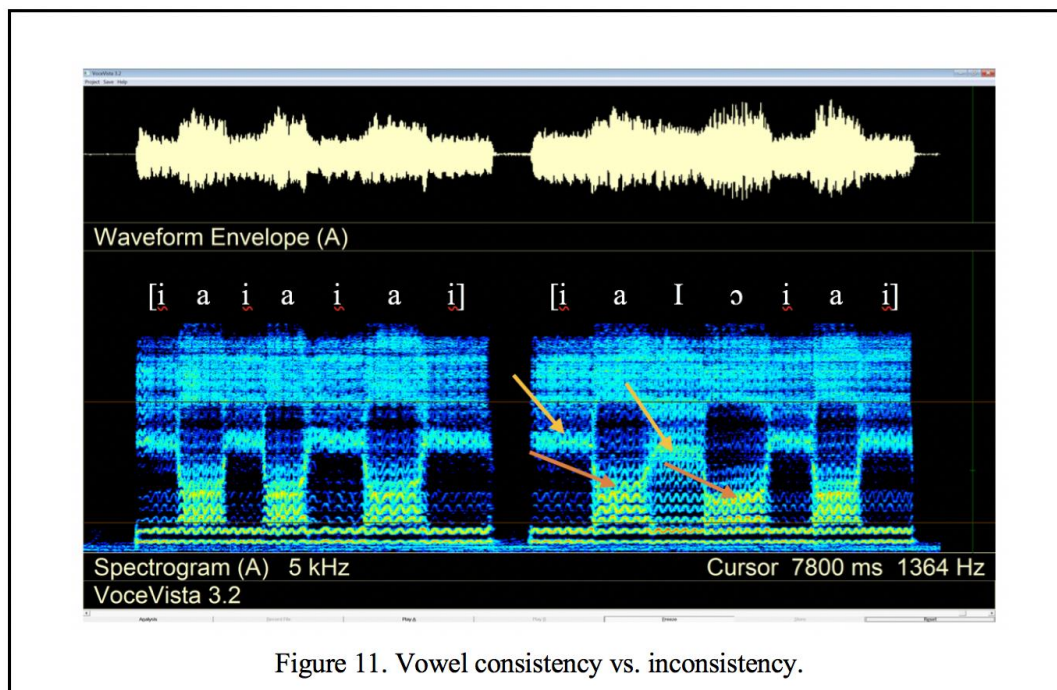
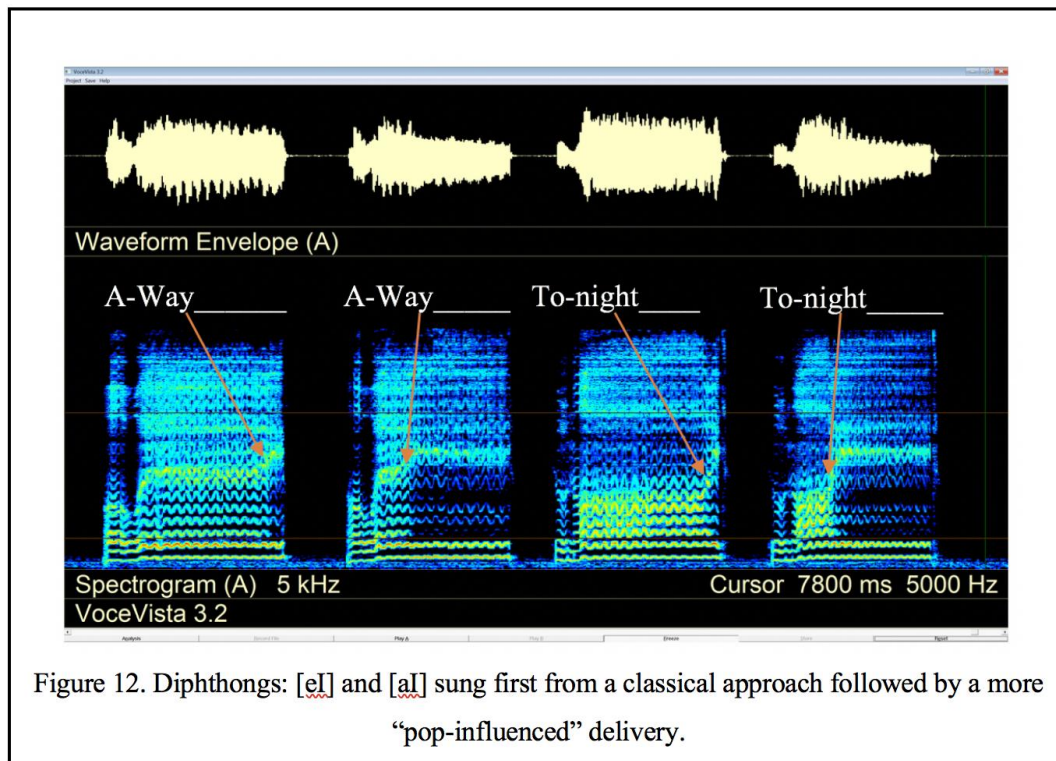


Figure 11. Vowel consistency vs. inconsistency.

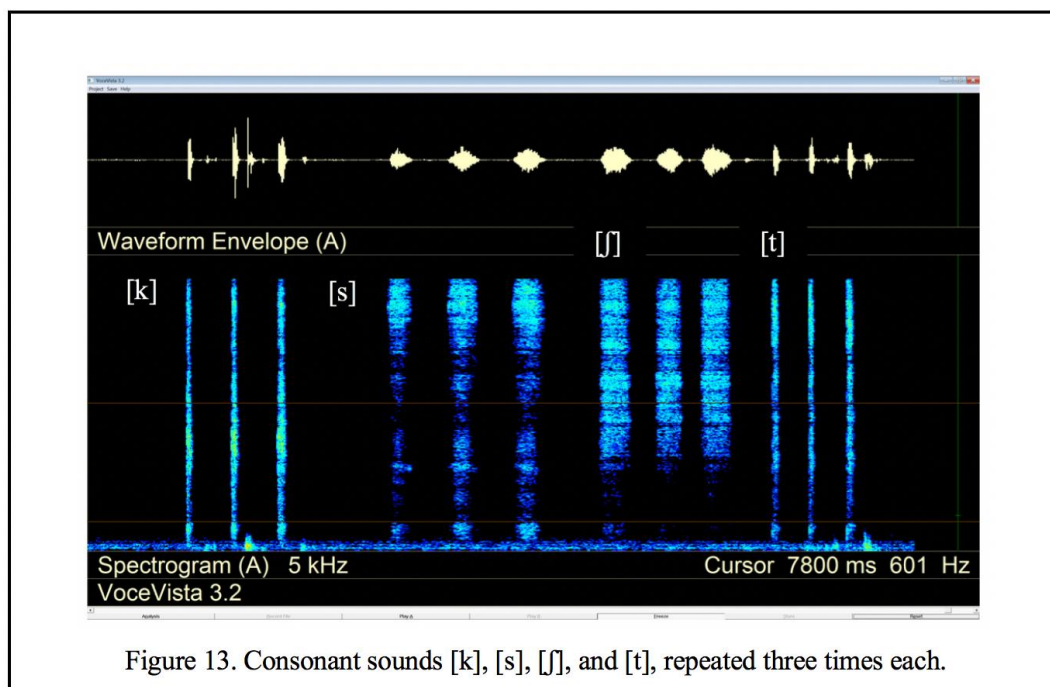


Two problematic diphthongs are considered in Figure 12 by singing the words, *Away* and *Tonight*, both of which are found in countless song settings across genres. Each example begins in the generally accepted delivery for Western art music, where singers stay on the first vowel of the diphthong for the majority of the pitch and then change to the second vowel late. (The orange arrows indicate where the change of vowel takes place within the diphthong. One can see the transitional “slide” that is created between the two vowels – an upward slide in both cases – as the acoustic energy changes based on tongue position and vocal tract shape.) The second performance of each example in Figure 12 was delivered in the more commercial or “pop” style where singers move to the second vowel of a diphthong much sooner. Contrasting the two phonations of the same word graphically reveals the delay or immediacy of diphthong transitions. The technology makes it simple for singers to practice precision within these transitions.

Vowels are paramount in the creation of beautiful tone and evenness of the vocal line, but consonants become king for artistic expression and the intelligibility of text. How many times have conductors asked ensembles for more [k] at the beginning of words like *Kyrie* and *King*, or

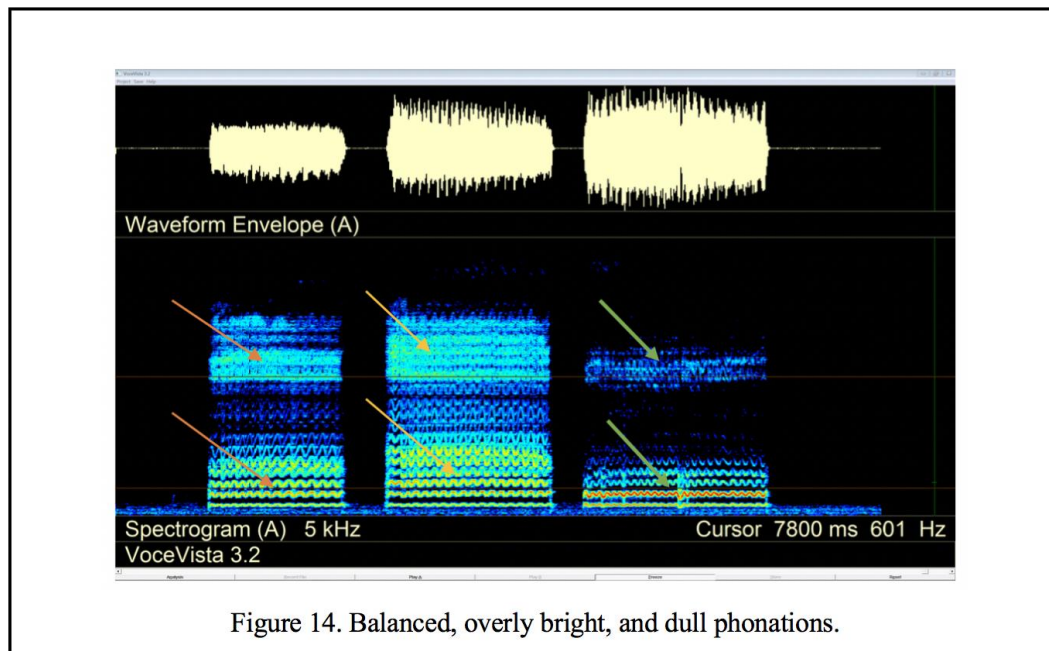
for a crisper, more unified [t] at the end of words such as *Out* and *Night*? Laziness in articulator energy or weakness of consonant sounds can render a song incomprehensible or cause an audience to miss a deeper meaning. Spectrographic analysis provides a sharp graphic image for the exactitude and length of voiced and unvoiced consonants, which are clearly identifiable from the vowel sounds of a sung phonation.

Figure 13 displays the spectrographic images generated by the consonant sounds [k], [s], [ʃ], and [t]. Because some consonant sounds, such as [k] and [t], have only a very short duration, each sound is articulated three times which allows one to gain a clearer picture of the image created by the consonant sound. Other sounds, such as [s] and [ʃ], are able to be sustained for longer durations. Conductors and teachers can use all of these visual cues to work on the desired uniformity, clarity, and length of consonants within ensembles or with soloists.



One last item that deserves mentioning for which spectrographic analysis is most useful is resonance balancing. It may be more helpful in solo singing than ensemble singing, but it can still be useful for ensembles – especially in sectionals. The upper overtones of a sound (in the range of 2500-3500 Hz) provide the brightness (*chiaro*), and the lower overtones of a sound provide the richness (*oscuro*). If one of these areas has significantly more acoustic energy than

the other, the tone can often sound out of balance; the sound may be either too tinny/edgy or too hollow/dull. Three different types of resonance balances are visible in Figure 14. Each phonation was sung by a male using the [o] vowel on the pitch A3 (220 Hz). The first phonation demonstrates a rather even balance of acoustic energy between the upper and lower overtones as indicated by the orange arrows. The second phonation reveals significant increases, and actually an over intensification, in the 2500-3500 Hz range. In fact, the entire phonation grew in acoustic intensity, as indicated by the yellow arrows, and would be an example of an overly bright or “tiny” sound caused by a very lateral mouth posture and a narrowed vocal tract. The final phonation in Figure 14 looks decidedly different than either of the first two. As can be seen with the green arrows, there is a decisive lack of acoustic energy in the upper overtones but a significant increase in acoustic energy of the first and second harmonics. This is an example of an overly covered/dull phonation caused by lip rounding and exaggerated tense vocal tract shape.



Spectrogram Software & Apps

Spectrogram technology is more accessible now than ever before, so there is no reason not to experiment with it. The software that was used for this article was Voce Vista Pro. The

same company now offers Voce Vista Video and Voce Vista Video Pro, which have even more features and higher resolution than Voce Vista Pro. Full specifications and functionality for these products may be found at <http://www.vocevista.com/purchase-vocevista-video-for-mac-and-pc/>. Voce Vista software has been used by singing voice professionals for nearly two decades, and it continues to evolve. There are also several spectrogram apps, both free and for purchase, that can be downloaded for portable devices. Like most things, one gets what one pays for, so quality can vary a great deal. However, there are two free apps that would offer good functionality for basic use. These are “SpectrumView”⁴ by Oxford Wave Research Ltd., and “Spectro Real”⁵ by WEVOSYS Medical Technology. Both of these apps have pros and cons but are worth exploring.

In summary, there are a myriad of ways in which spectrographic analysis and a waveform display may be used to help improve ensemble and solo singing. Conductors and singers only need to experiment fearlessly. Below is a list of exercises discussed in this article that can be enhanced by this technology:

- For Fun – with things like sirens in warm ups.
- Establishing clean onsets and releases
- Working to avoid scooping or sliding (unless they are desired)
- Creating appropriate vowel shapes
- Maintaining vowel consistency
- Strength and precision of consonants
- Presence and/or absence of vibrato
- Achieving desired resonance balance
- Developing dynamic control and contrast

⁴ Oxford Wave Research Ltd., 2018.

⁵ WEVOSYS Medical Technology GmbH, 2014.

Music and technology are both powerful tools, and, when the two work together in harmony, an incredibly productive and rewarding experience can result.

References

- Bozeman, K. (2013). *Practical Vocal Acoustics: Pedagogic Applications for Teachers and Singers*. Hillsdale, NY: Pendragon Press.
- Dellatre, P. (1958). Vowel Color and Voice Quality - An Acoustic and Articulatory Comparison. *NATS Bulletin*, October, volume 15, issue 1, 4-7.
- McCoy, S. (2004). *Your Voice: An Inside View*. Columbus, OH: Inside View Press.
- Miller, D. (2008). *Resonance in Singing: Voice Building through Acoustic Feedback*. Gahanna, OH: Inside View Press.
- Miller, R. (1986). *The Structure of Singing: System and Art in Vocal Technique*. New York, NY: Schirmer Books.
- Nix, J., Mabry, G., Matthews-Muttwill, A. (2007). Chorister Perceptions of Real-Time Displays of Spectra in the Choral Rehearsal: A Feasibility Study. *International Journal of Research in Choral Singing* (3) 1, 1-15.

Software & Apps

- Pitch Perfect, Pitch Pipe App: DepollSoft, 2018.
- Spectro Real, Spectrogram App: WEVOSYS Medical Technology GmbH, 2014.
- Spectrum View, Spectrogram App: Oxford Wave Research Ltd., 2018.
- Voce Vista Spectrogram Software: <http://www.vocevista.com/>