Functional Formants: An Overview of Modern Singing Voice Science

Angerhofer, Thomas Erik

University of Colorado Boulder

Author Note

TMUS 8269 Lecture Project for DMA in Voice Performance and Pedagogy

1. Title
2. Singing Voice Science is said to begin with Manuel Garcia and his invention of the laryngoscope in 1854. Much like the instrument used in modern dentistry, the angled mirror allowed him to look down the larynx at the vocal folds for the first time, observing and detailing its behavior during singing. Due to the shape of the vocal tract, particularly the obstruction of view caused by the false vocal folds, and the rapid oscillation of the vocal folds during phonation, observations made by the naked eye with a small and awkwardly placed mirror would make it difficult to gather precise information on the nature of the vocal instrument. Nevertheless, Garcia's efforts have inspired continued scientific inquiry which continues to illuminate our understanding of the anatomy and physiology, acoustics, and psychology of the singing voice.
3. Interest in singing voice science has not always been constant. Many pedagogical texts throughout the 1800's and early 1900's include some anatomy and physiology, but have limited reference to the physical nature of sung sounds. Much of the research regarding the sounds of the human voice during that time was an outgrowth of the invention of the Edison phonograph and the telephone. The phonograph, while primarily used for entertainment, would also allow for capturing and recording sounds for later processing and analysis. At Bell Telephone Labs, the first spectrogram was developed to analyze and better understand the sounds of speech so as to improve the new medium of telephony. These first voice scientists were linguists and / or acousticians, and their work played a major role in our early scientific understanding of the sounds of the human voice.
4. Pierre Delattre was one of these linguists. A professor of French, he was intrigued by the possible applications of the spectrograph for teaching foreign languages. During his time

at The University of Pennsylvania, he made numerous trips to Bell Labs to better his understanding. Described as a “renaissance man,” Delattre was also an amateur musician who was apparently a fine pianist. When he later moved to the University of Colorado at Boulder, he was able to build his own lab that would attract the attention of noted CU voice professor Berton Coffin. Given the similarities between speech and singing, it was not long before the work of Delattre and other acoustic phoneticians would make its way into NATS publications and receive regular citations in a new generation of voice pedagogy texts, including those by Coffin and Vennard.

5. Over the years there have been many notable pedagogues that have contributed to our scientific understanding of the singing voice. Richard Miller’s laboratory at Oberlin was among the first of its kind, and Ingo Titze’s ongoing work through the National Center for Voice and Speech has given rise to the new discipline of Vocology. Newer technologies for gathering data, such as the flexible fiberscope, magnetic resonance imaging, and electroglottography, and software applications for analysis and synthesis, such as VoceVista and Madde, have allowed for the scientific exploration of the singing voice to be simultaneously more rigorous and accessible.
6. While our understanding of acoustics in singing voice science is still evolving, there is much about the physical nature of sung sounds where there is considerable agreement. The next section of this presentation will detail what we have come to understand over the past century with regard to the acoustic nature of sung sounds and how they are produced.
7. The vocal instrument is made up of three distinct parts. The terminology used to describe these parts have changed over the years; once known as the actuator, the oscillator, and

the resonator, many singing voice scientists now refer to them as the power source, the sound source, and the filter.

8. The respiratory system serves as the power source for the vocal instrument. This includes many of the muscular and skeletal structures of the torso and pulmonary system starting as low in the body as the coccyx and as high in the body as the clavicle, with some extensions even beyond those points. One way of thinking of the power source is that it collects and distributes air – the potential energy for the vocal instrument. This breath energy is then transformed to sound energy by the vocal folds.
9. The sound source is housed within the larynx. The vocal folds close, open, and repeat due to a variety of muscular and aerodynamic forces, as well as fluctuations in pressure. The opening and closing of the folds create "puffs" of air, or vibrations that travel through the atmosphere to the ears of the listener. Periodic, or regularly oscillating folds create a sound that has a discernable dominant frequency, which is interpreted as a musical tone. Irregular, or non-periodic oscillations with no discernable dominant frequency are interpreted as noise.
10. The filter or resonator is the area of the vocal tract above the sound source, beginning just above the vocal folds and extending to the lips. The structures of the resonator are highly moveable, allowing for changes in length, shape, and tonus. The cavities of the vocal tract will oscillate when one of the frequencies produced at the sound source is approximate to the natural frequency of the cavity – effectively amplifying or boosting the intensity of that frequency. The vocal tract can also have a dampening effect on the source harmonics when the natural frequencies of its cavities are not approximate to the frequencies being produced at the source.

11. This short video shows acoustic activity in singing. The lines represent the individual frequencies, or source harmonics, that come together to make a complex musical tone. Darker, more blueish colors indicate a lower level of energy, while brighter reds and whites are used to show higher levels of energy in the individual frequencies. This spectrogram is a visual representation of the composite sound as it is produced at the source and modified by the filter / resonator.
12. The sound waves produced by the opening and closing of the vocal folds make up the raw materials for sung musical tones. A sound wave is a pressure wave; in the picture to the left, the lines in section A represents an undisturbed medium – particles of air that are somewhat still. The vocal folds open and close during phonation, releasing pressure built up below in the form of tiny, regular puffs of air. These puffs disturb the particles of the atmosphere, causing them to move closer to one another. This phase of the wave, called compression, can be seen in section B – represented graphically in section C. The particles then rebound back toward their origin in the phase of the wave called rarefaction. The compression phase is higher in pressure, and the rarefaction phase is lower in pressure. The amplitude, which indicates the amount of energy in the wave, is related to how listeners might perceive the "loudness" of the sound. One compression and one rarefaction constitute a cycle, and the number of cycles within a given period of time determine the frequency.
13. Frequency is measured in Hertz, or cycles per second. For every octave, the frequency increases by a factor of 2. For example, the A below middle C vibrates 220 times per second, while the A above middle C vibrates 440 times per second and A5 has 880 cycles

per second. While we perceive octaves as being equal, the number of cycles per second of each successive octave is double that of the previous one.

14. The basilar membrane, inside the cochlea of the inner ear, acts as a frequency spectrum analyzer. It varies in stiffness and will resonate in different places depending on the frequency of the wave. Near the oval window it is stiffer and will resonate with higher frequencies; toward the apex, where it is floppier, it will resonate with lower frequencies. It is thought that because a doubling of frequency causes a response at an equal distance on the basilar membrane, we hear octaves as related. Because of the logarithmic spacing of pitch placement on the membrane, we perceive musical intervals not as differences in frequency, but as ratios between frequencies. A220 to A440 is perceived as the "same" interval as A440 to A880, even though one pair has a difference of 220 Hz and the other a difference of 440 Hz because they share a 2:1 frequency ratio.
15. The frequency of a sound wave is interpreted by the ear and brain as pitch; for most musical tones, however, what we actually hear is many frequencies at once. These frequencies are mathematically related to and superimposed on the fundamental frequency, which we hear as the pitch. The strength and number of the superimposed frequencies being produced at the sound source and how they are resonated or filtered by the instrument are what give the instrument its unique timbre. In the case of the human voice, where the resonator can be reshaped variably tuned, the many timbral possibilities lend to a variety of communicative and expressive colors.
16. When the singing voice is well produced, the vocal folds oscillate in a variety of ways: whole, half, quarters, eighths, etc. Each of these subdivisions produce a wave with a frequency mathematically proportional to the fundamental, or lowest frequency. The first

overtone is 2x the frequency of the fundamental, the second overtone 3x the fundamental, the third overtone 4x, and so on. The most current term for these frequencies is source harmonics, as they follow the harmonic series; however, many authors use the synonyms “overtones” and “partials” interchangeably. [The fundamental frequency (F0) is the same as the first harmonic (H1) or partial. The first overtone is the same as the second harmonic (H2) or partial, and so on.]

17. When analyzing pitch and frequency, it is important to take into account the role of aural perception. Octaves, for example, have logarithmic spacing with each successive octave having double the number of vibrations as the previous one. Nevertheless, they maintain a 2:1 frequency ratio, and that is why we hear octaves as being the same interval.
18. Source harmonics, on the other hand, are multiples of the fundamental frequency. While there is a constant number of vibrations between them, we hear the intervals as getting closer together as they move farther away from the fundamental.
19. In this video you can see and hear each individual frequency of the complex wave being produced at the source. Using a computerized spectrograph to record the singer, we are then able to use a filter in the software to isolate each of the source harmonics in the spectrogram.
20. Despite the sound having traveled through the resonator / filter in the video, it is noticeable how the power of each successive harmonic mostly decreases. In well produced singing voices, the typical decrease in power throughout the spectrum, also called “roll off,” is about 12 decibels per octave at the sound source (or before having gone through the resonator).

21. Beside the natural, optimal roll off, there are four other factors that influence the strength and number of audible harmonics produced at the sound source: mode of phonation, laryngeal register, intensity, and fundamental frequency.
22. How much air is being supplied by the power source and how the vocal folds either release or resist the flow of air during phonation determines the amount of pressure. Depending on where the singer is in the range, the desired strength and timbre of the tone, and a variety of other considerations, there are many possible positive ratios between flow, release, and resistance. When the power source supplies too little or too much air and the sound source responds by resisting the flow too little or too much, the mode of phonation is said to be out of balance. "Flow phonation" is often used to refer to an optimal state when an appropriate supply of air is met with balanced vocal fold resistance / release.
23. Hyper adductive, or "pressed" phonation occurs when the vocal folds overly resist the supply of air during phonation. This mode of phonation will create a moderately strong fundamental frequency with a moderate roll off in power in the middle harmonics, and more and stronger higher harmonics. The resulting sound is an overly brassy, metallic, or bright timbre.
24. Hypo adductive, or "breathy" phonation occurs when the vocal folds do not resist the supply of air adequately during phonation. This mode of phonation will create a moderate to strong fundamental frequency with a steep roll off in power, leading to fewer and weaker harmonics with an element of air or noise. The resulting sound is airy, fluty, and often noisy.

25. There are two sets of muscles that are primarily responsible for adjustment of pitch. The thyroarytenoids, which make up the body of the vocal folds, shorten and thicken the folds on contraction, thus preparing them for vibrating at lower frequencies. The crycothyroids tilt the thyroid cartilage forward and down upon contraction, thus lengthening and thinning the vocal folds and preparing them for vibrating at higher frequencies.
26. Vibrational Mode 1, also called thyroarytenoid dominant phonation, is typical of the lower, "chestier" part of the range. It exhibits a brusque movement where contact is made from the lower to the upper edge of the vocal folds in each cycle. Because the folds are shorter and thicker, they will release pressure from the bottom up during phonation. This vertical phase difference creates a complex wave with more and stronger high harmonics, resulting in a brassier tone quality.
27. Vibrational Mode 2, also called cricothyroid dominant phonation, is typical of the higher, "headier" part of the range. Because the folds are longer and thinner, they appear to "chop" the air more smoothly and exhibit little to no vertical phase difference during phonation. This creates complex wave with a strong fundamental frequency, but with fewer and weaker high harmonics, resulting in a "flutier" tone quality.
28. Intensity is the physical property of sound that is perceived as "loudness" by the listener. When phonation is more intense, or "louder," the complex wave will have more and stronger harmonics. When phonation is less intense, or "softer," the complex wave will have fewer and weaker harmonics.
29. The final factor influencing the number and strength of the source harmonics is fundamental frequency. The lower the sung pitch, the greater the number of closely spaced harmonics there will be within audible range; that is to say, notes sung lower in

the range and/or by lower voice types produce more sonic material to then pass through the resonator / filter. Notes sung higher in the range and/or by higher voice types are passing less sonic material through the vocal tract. Furthermore, the higher the fundamental frequency in each voice type, the more likely it will be produced in a more CT dominant phonation, resulting in fewer and weaker harmonics. Pictured are harmonic series with H1/F0 at C4 and H1/F0 at C5; notice how there is a cluster of harmonics in the upper octave of the piano when H1=C4. That same cluster would be an octave higher when H1=C5, placing the cluster outside the sensitive range of human hearing. Nevertheless, the ear and brain tend to perceive higher frequencies as being louder than lower ones with the same amplitude. So, for voices producing frequencies over 2000 Hz in the lower end of the harmonic spectrum, the number and power of higher harmonics may be of little importance. We'll revisit this as we explore the topic of resonance.

30. Resonance, as we know it from physics, is the reinforcement or prolongation of sound by reflection from a surface or by the synchronous vibration of a neighboring object. In the case of the human voice, the sounds being produced at the source must pass through the remainder of the vocal tract, including the laryngopharynx, oropharynx, and oral cavity, before clearing the lips and moving on to the atmosphere. The length of the vocal tract, from the sound source to the lips, and the manner in which the tongue can effectively divide the tube into 2 variably shaped containers, plays an extremely important role in modifying the character of the source harmonics.
31. In this video, the washer set in motion could represent a harmonic from the source. The other washers represent the potential resonances of the vocal tract, or how the hollow spaces of the vocal tract might oscillate when they are tuned to one of the frequencies

emitted from the source. On the other hand, just as the washers on the ends of strings of a different length fail to oscillate sympathetically, if the cavities of the vocal tract are not tuned to any of the source harmonics, they will not enhance the source harmonics. By manipulating the shapes of the vocal tract, we can boost certain frequencies while ignoring others – effectively dampening them. In this way, resonance contributes to the timbre of sung sounds, as well as intelligibility of language and efficiency of production.

32. With its highly formable nature, the vocal tract is a special type of resonator. In addition to our ability to manipulate the length of the vocal tract by raising and lowering the larynx and velum or protruding and pulling back the lips, we can also divide the space by moving the bulge in the tongue. Further, the walls of the pharynx can be more or less firm, lending to a more "wet" or "dry" acoustic that will either reflect or dampen some of the waves. Additionally, the uneven texture of the tissue allows the vocal tract to resonate source harmonics that are close in frequency; absolute precision is not required.
33. The vocal tract has been traditionally referred to as a Helmholtz resonator, as it has a resonating chamber and is open at one end. More recently it has been found to have more in common with tube resonance, also called a quarter wave resonator, as it is also open at one end, but is capable of producing standing waves at ¼ the length of the wave, as well as odd multiples ¾, 5/4, 7/4, etc. These frequencies are often referred to as being the "natural" resonances of the vocal tract.
34. If the F0/H1 is 220 Hz (A3), the vocal tract, if shaped well, could resonate at frequencies of 275 Hz (~C#4), 385 Hz (~G4), 495 Hz (~B4/C5), 605 Hz (~D5/D#5), 715 Hz (~F5/F#5), 825 Hz (~G5/G#5), 935 Hz (~A#5/B5), 1045 Hz (~B5/C6), 1155 Hz (~C#6/D6), 1265 Hz (~D#6/E6), 1375 Hz (~E6/F6), 1485 Hz (~F#6/G6), 1595 Hz

(~G6/G#6), 1705 Hz (~G#6/A6), 1815 Hz (~A6/A#6), 1925 Hz (~A#6/B6), 2035 Hz (~B6/C7), 2145 Hz (~C7/C#7), 2255 Hz (~C#7/D7), etc. By the time these frequencies approach the top octave of the piano, they comprise every semitone, as well as many frequencies in between.

35. Formants are perhaps the most vital resonances of the vocal tract. The most important of these resonances are formed by the way that the bulge in the tongue divides the vocal tract into 2 variable resonance chambers. Not only are these resonances responsible for boosting the frequencies that make vowels identifiable, but they are also responsible for expressive timbres and encouraging efficient phonation.
36. The first formant resonance is tuned in the laryngopharynx and oropharynx, or the space above the vocal folds extending upwards to the velum or soft palate. The natural bulge in the tongue separates this space from the oral cavity. Because it is the lowest resonance, it can lend to depth or fullness in the sound. Its tuning is responsible for acoustic registration, or open and close timbre. If there is more than an octave between the fundamental frequency (F0/H1) and the first formant (F1>H2), the timbre is said to be open. This is typically characterized by a high larynx and low soft palate, shortening the length of the F1 space, and will usually result in an overly bright or “shouty” sound. Maintaining a settled larynx and a raised soft palate will help to produce a close timbre, where the first formant is less than or equal to the second harmonic. This not only contributes to the depth of the sound, but also motivates a more efficient phonation as it provides appropriate supra-glottal pressure due to the rarefaction in its oscillations reflecting back on the source in sync with subsequent vibratory cycles.

37. The second formant resonance is tuned in the oral cavity, from the bulge in the tongue to the lips. Because the frequency of this resonance is higher, it contributes to the "brightness" or clarity of the sound. This timbral perception is often associated with the "front" or "back" dimension of vowels; the more fronted the bulge in the tongue, the higher the frequency of the F2. Depending on the fundamental frequency, it may also help some voices boost the resonance frequencies in the upper octave of the piano where the cavity of the ear tends to resonate, giving the voice a sense of brilliance or ring.
38. This image, borrowed from the National Center for Voice and Speech, represents the traditional concept of vowel resonance. Numerous pedagogy texts over the past 60 years have included a similar graph showing the range of frequencies that determine the intelligibility of each vowel. Coffin states that vowels are timbres, meaning that they are a quality of sound distinct from pitch and intensity. More specifically, vowels have been thought to be a combination of two different resonance frequencies independent from the fundamental frequency, or perceived pitch. Vowels that are said to be "close" have a lower first formant frequency, while vowels that are said to be "open" have a higher first formant frequency. Vowels that are said to be "back" have a lower second formant frequency, while vowels that are said to be "front" have a higher second formant frequency. /i/, for example, is both "close" and "front", as it has a low first formant frequency and a high second formant frequency. By contrast, /u/ is both "close" and "back", as it has a low first formant frequency and a low second formant frequency.
39. The singing voice has been said to have a "special" resonance that helps it to be heard even over the sounds of a large orchestra. This theory goes back at least as far as Vennard in 1949, and has been supported by many notable pedagogues since. This resonance may

be a product of a few factors. First, singing voice scientists, as well as speech and hearing specialists, have noted that the human ear tends to be more sensitive to higher frequencies. Bozeman notes that the human ear, also a quarter wave resonator, has a primary resonance in the 3000 to 4000 Hz range. For reference, this would be in the top octave of the piano or in the octave just above. This resonance within the ear will boost frequencies in this range and cause us to perceive them as louder than lower frequencies with the same amplitude.

40. Nevertheless, the cause of this special resonance remains the source of some controversy. Traditionally speaking, pedagogues have accepted the theory that this resonance is created by the narrowing of the laryngopharynx at the epiglottis, in a space referred to as the “epilarynx.” McCoy states in “Your Voice: the Basics,” published in 2016, that the space between the vocal folds and the tip of the epiglottis must remain in a 6:1 ratio with the surrounding space in the throat. Sundberg further noted that this resonance is also partially due to a partially lowered larynx, and that it does not exist in the upper ranges of the female voice. One assumption of this theory is that voices singing a fundamental frequency above C5 may not be relying on a partially lowered larynx, as they do not exhibit as much of a boost in the resonance frequencies in this part of the spectrum.

41. It is interesting to note that Donald Miller shows how other voice types might not rely on the singer’s formant for carrying power in his book “Resonance in Singing.” Pavarotti’s B4 in “La donna e mobile” has a peak in energy on the third harmonic due to his tuning of the second formant resonance, with less energy in the area of the singer’s formant. Carrerras uses this same strategy, while Kraus and Domingo’s final “pensier” shows less

energy in the third harmonic with greater energy in the area of the singer's formant. Two different resonance tuning strategies, and both with desirable outcomes.

42. It may be useful to revisit the "natural" or tube resonances of the vocal tract for more insight into this special resonance. Here we see the potential quarter wave resonances of A3 (220 Hz) along with the source harmonics at the same pitch. Notice that there is a cluster of potential resonances that coincides with the source harmonics in the upper octave of the piano. Bozeman refers to this as the singer's formant *cluster*, pointing out that there are a number of harmonics being resonated in close proximity, all within the most sensitive range of human hearing.

43. Bozeman also states that maintaining a stable larynx throughout the sung range would keep this resonance cluster in approximately the same part of the spectrum, thus lending to a more consistent vocal timbre. Further, vocal tracts vary in length from person to person, as well as between voice types. Usually, the higher the fach, the shorter the vocal tract length. A longer vocal tract would lower the frequencies of this resonance cluster, while a shorter tract would raise them. This could possibly account for the difference in resonance strategies between Pavarotti/Carrerras and Domingo/Kraus (shorter vocal tracts vs. longer vocal tracts), but also the absence of an audible boost in the SFC for voices that sing higher fundamental frequencies. Simply stated, the SFC in these voices might be higher than the frequency range where our ears are sensitive and/or the SFC might not align with the tight grouping of source harmonics. In any case, all voices will have some harmonic material to resonate in the appropriate part of the spectrum.

44. As researchers have come to better understand the physical properties of the sounds of singing and the roles that each of the individual parts of the instrument play, they have

also developed theories as to how the instrument behaves as a whole. These models become increasingly complex as the understanding of the nature of sung sounds has evolved. This section will outline the major theories, detailing their contributions and possible shortcomings. Further, it will include new research in the field of psycho-acoustics that may have important consequences for models of vowel resonance.

45. A Helmholtz resonator is a container of gas (air in the case of the vocal instrument) with an opening, neck, or port. Named after the German scientist Hermann von Helmholtz (1821-1894), the coupled Helmholtz resonator theory provided an important foundation for understanding the acoustics of the singing voice. Vennard summarizes this theory, stating that the vocal folds "produce a tone with a definite desired pitch, considerable volume, and of complex timbre, having the possibilities of beauty. As this tone passes through the throat and mouth, these cavities encourage those partials which make for power and beauty, and muffle the undesirable ones; or, in the case of the unskillful singer, the opposite occurs (p. 81)." Edward Wheeler Scripture took exception to this position back in 1906, arguing that the vocal folds did not produce overtones. According to him, the puffs of air emitted by the opening and closing of the glottis produced a pitch, and the puffs would then excite the air in the resonators so that they would sound their own frequencies – independent of the sound source.

46. While neither the coupled Helmholtz model nor Scripture's observations are entirely accurate, they provided the framework for what would become Source-Filter theory. Like the Helmholtz model, the source-filter model acknowledges the source harmonics as being the raw materials for potentially resonated sounds. These frequencies may then cause the pharyngeal cavities to oscillate. Intelligible vowels result from the timbres

created by the first and second formant resonances. Further, the source-filter model recognizes the presence of tube, or quarter wave resonances that are related to the source harmonics in odd intervals, but differ in frequency from the source harmonics.

47. Non-linear source-filter theory came to prominence in the mid-2000's by notable researcher Ingo Titze at the National Center for Voice and Speech. Building on the source-filter model, this revision recognizes the potential for interactions between the source and filter. An appropriately tuned resonator can utilize the power of rarefaction, the reflective portion of the sound wave, to decrease subsequent effort at the sound source. In other words, the energy reflected back on the source from a well tuned resonator can improve the efficiency of phonation and the power of the source harmonics; thus advancing sustainability during prolonged activity and at higher intensity levels.
48. While the linear model assumes that the power source, sound source, and filter / resonator operate independently, the non-linear model shows how the different parts of the instrument can influence each other's behavior. Christian Herbst has noted that these interactions would make isolating the behavior of any singular part of the instrument problematic, as each part is subject to the forces of the other parts. As an example, imbalances in flow or resistance might actually be remedied by resonance tuning, rather than adjusting the air flow or muscular activity associated with adduction.
49. While we are becoming increasingly aware of strategies that improve efficiency and sustainability, it is important to note that these qualities are not always desirable when making certain expressive and stylistic choices. Singers often select strategies that do not take advantage of efficient interactions to meet communicative demands, especially when

the function of projection is managed by external amplification. Here, Pavarotti expertly demonstrates open vs. close timbre while attempting to explain the benefits of the more efficient interactions resulting from the latter.

50. Ian Howell's investigations of the intersectionality of acoustics, or the physical nature of sound, and psychoacoustics, the human perception of sound, seems to have some interesting implications that both further our understandings and challenge long held beliefs. As discussed earlier in the section on the Singer's Formant Cluster, it may not be the case that the larynx rises in voices singing fundamental frequencies that approach the top of the staff. The perceived absence of this cluster could simply be that it naturally lies above the top pitches of the piano and outside the range where the quarter wave resonance of the ear would give those frequencies a boost of energy. This could be either due to the high fundamental frequency and the spacing of its harmonics, the tube length of the singer, or both. It wouldn't necessarily mean that the singer has a high laryngeal position; while the tuning of the SFC would be somewhat constant, frequencies sung lower in the range would have at least one harmonic that could be resonated near the upper octave of the piano, making the SFC irrelevant.
51. Howell extends this logic to vowel formants and intelligibility. The long-accepted practice of vowel modification or substitution prescribes altering vowels, or changing the formant tuning, to encourage more efficient source-filter interactions. Howell states, "essentially, above the treble staff, vowel clarity disappears entirely not because of vowel substitutions or modifications, but because the simplicity of the listener's percept is too distant from the timbral complexity of speech (p. 4)." In other words, it is not necessarily the adjustments of the vocal tract that cause vowels to diminish in intelligibility in the

higher parts of the range. Instead, it is because these frequencies are so far removed from the typical range of speech intonation.

52. Further, Howell's idea of *absolute spectral tone color* challenges our traditional acceptance of vowels requiring two resonances in order to be understood. Vennard first noted the spectral peaks shared by different vowels; "…when one sings Ay [e], he is really singing Oh [o] plus a high partial which is not heard in the Oh [o]; and when one sings Ee [i], he is really singing Oo [u], plus a still more ringing overtone."6 Rather than focusing on their objectively measurable frequencies, he calls attention to their shared *tone color*. Howell's research takes this idea further, noting that the listener might perceive vowels by the *tone color*, which he defines as a combination of "the harmonics of the voice" and "other elements of timbre idiomatic to the voice (such as the attack, decay, and release) (p. 24)." Therefore, a sung harmonic has it's own *absolute spectral tone color* that might allow it to be identifiable as a vowel in and of itself. This is why we can sometimes hear the [u] in an [i] vowel; the resonance of the F1 is sounding [u] while the resonance of the F2 is sounding [i]. Bozeman refers to this as the "back vowel" and the "front vowel." When combined with the non-linear source filter model, *absolute spectral tone color* provides a framework for efficiency and intelligibility. The F1 resonance is responsible for warmth, roundness, depth, efficiency, and for /u o/ primary vowel identity, while the F2 resonance is responsible for clarity, degree of brightness, efficiency, and for /ɑ a e i/ primary vowel identity.

53. For nearly 200 years, the scientific exploration of the singing voice has yielded important and useful information about the physical nature of sung sounds. As the tools for collecting such data have become more precise and accessible, more accurate models

have been developed to describe the behaviors of the singing voice. The teaching of singing is at least as complex as the act of singing itself, requiring at least some functional knowledge of anatomy, physiology, acoustics, and psychology, among other things. Just as individual singers have a personalized approach to their art, so, too, do teachers of singing. One value of singing voice science is that it should help to clarify our understanding of the vocal instrument. To that end I offer the following observations or suggestions.

54. First, there are numerous challenges regarding the terminology used in singing voice science. In some cases, we see different terms used interchangeably, such as source harmonics, partials, and overtones. In other cases, we see remnants of an earlier acoustic understanding or model. Hypo- and hyper- functional phonation, terms made famous in McKinney's *Diagnosis and Correction of Vocal Faults*, are examples of this. While it does indicate under or over energized singing, phonation is influenced by breath flow, vocal fold resistance / adduction, and the supra-glottal forces of resonance. There now exists more accurate language to describe these phenomena with greater exactness.
55. Yet another case we see is terminology borrowed from acoustic phonetics, such as open, closed, front and back in reference to vowels. These words were originally used to describe tongue position; closed vowels have a high bulge of the tongue, closer to the roof of the mouth, while open vowels have the tongue seated lower in the mouth. Back vowels have the bulge of the tongue farther back in the mouth, while front vowels move the front further toward the front of the mouth. While these positions may accurately describe tongue position in speech, they might not accurately describe how vowels are formed during singing. In addition, singing voice science uses open and close to refer to

F1 tuning; $F1 > H2$ = open timbre, $F1 \leq H2$ = close timbre. Authors like Bozeman try to reconcile this by using front/back perception, but perhaps there is a way to express this with greater clarity.

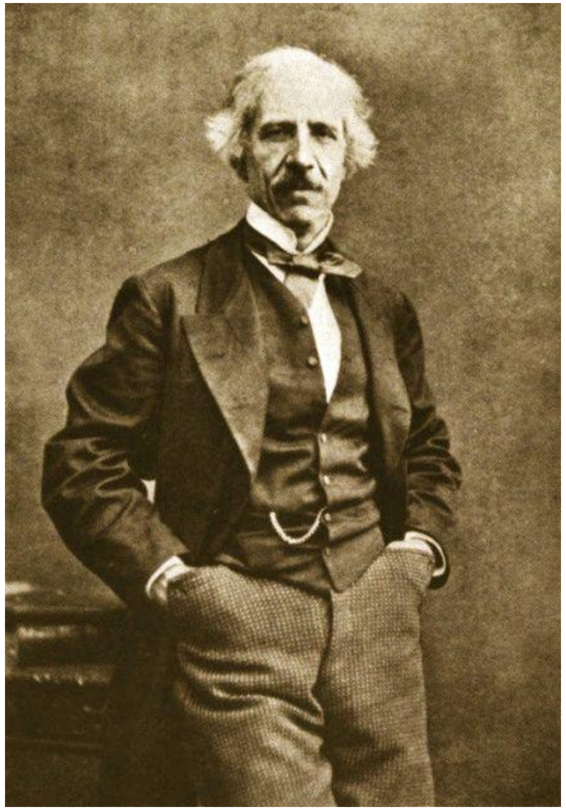
56. Just as singing voice science has borrowed language, so too has it borrowed research. While this helped to propel the science of singing, it has also been the cause of some dissonance. While singing and speech share the same basic structures, they differ in at least three important ways.
57. Projection: In conversational speech, the proximity and size of the audience is minimal. Intensity levels, therefore, can remain relatively low while still being intelligible. In classical singing, the distance and size of the audience is far greater. Balanced source harmonics with greater intensity levels than speech are required so that the resonator can play a more active role in helping sounds travel through the atmosphere to the listener.
58. Intonation: In conversational speech, variation of pitch is rather limited – except in moments of heightened emotion. Intonation is typically limited to less than a perfect 4^{th} up or down from the habitual or mean pitch, resulting in a functional pitch range of less than an octave. Professional classical singers are required to have a functional pitch range of over two octaves in order to meet the demands of many compositions, and singers often shape or enunciate sounds differently than in speech in order to maintain more favorable source filter interactions – especially in the extremes of the range.
59. Duration: With the exception of auctioneers and patter songs, duration varies greatly between speech and singing. It is frequently the case that the sung line is delivered at a slower tempo than in conversational speech, requiring greater specificity from the singer in defining vowel sounds and timing the change from sound to sound. In other words, in

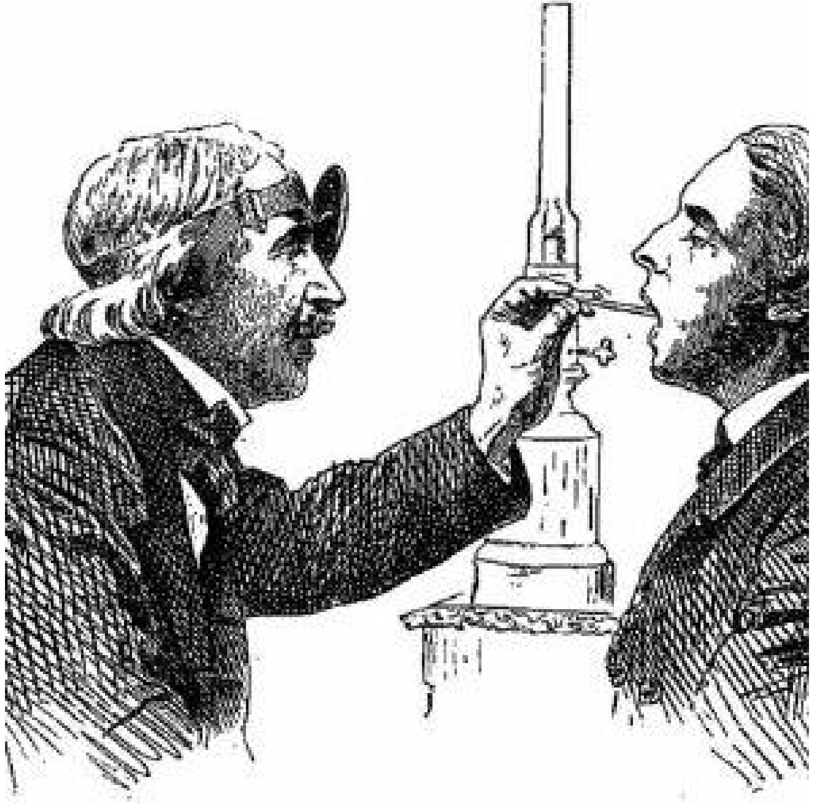
order to have the sounds of sung text have the same communicative effect as spoken language, they must be enunciated differently.

60. Given these fundamental differences, it would stand to reason that even if communicating with a speech-like quality was desirable in singing, it would have to be done differently in order to have that effect. As such, the physical nature of the sung sounds and the behavior of the instrument would also differ from those of speech. With many of the building blocks firmly in place from a century of research, clarifying the terminology of singing voice science and continued study of the nature of sung sounds and their perception *as a discipline separate from speech science* will benefit teachers and students of singing alike. The surge in research activity over the past 20 years, led by Titze, Miller, Herbst, Bozeman, Howells, and others, has connected vital pieces of the puzzle, yielding more accurate models for comprehending not only the physical nature of sung sounds, but also the behaviors of the instrument resulting from the interaction of its parts. Ongoing work in the field of singing voice science can help to refine our understanding of both the product and the process, as well as improve the methods of delivery to facilitate skill appropriation and execution, improving functionality in our practice as teachers and singers. (Masonic Interpretation Igne Natura Renovatur Integra meaning "Through fire, Nature is reborn whole" or "By fire Nature is renewed whole", symbolizing Humankind's spiritual regeneration by the sacred fire of truth and love.)

Functional Formants

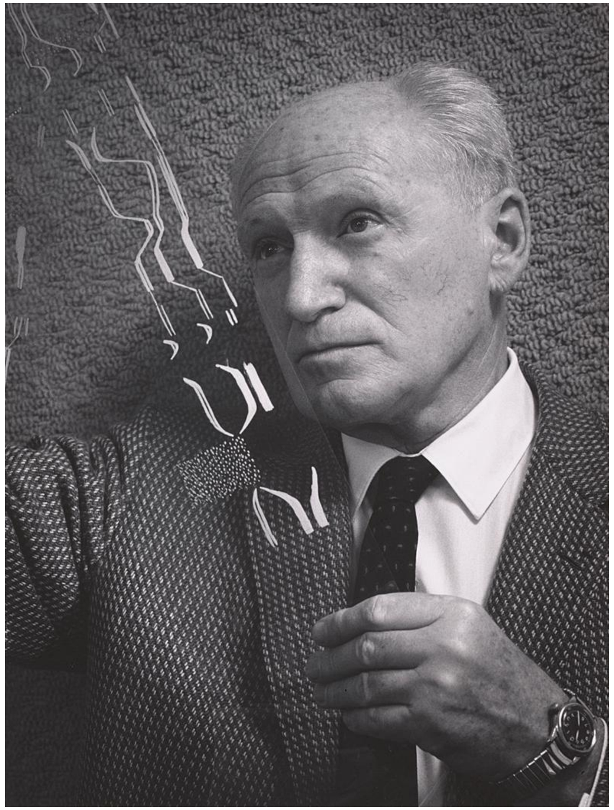
An Overview of Modern Singing Voice Science

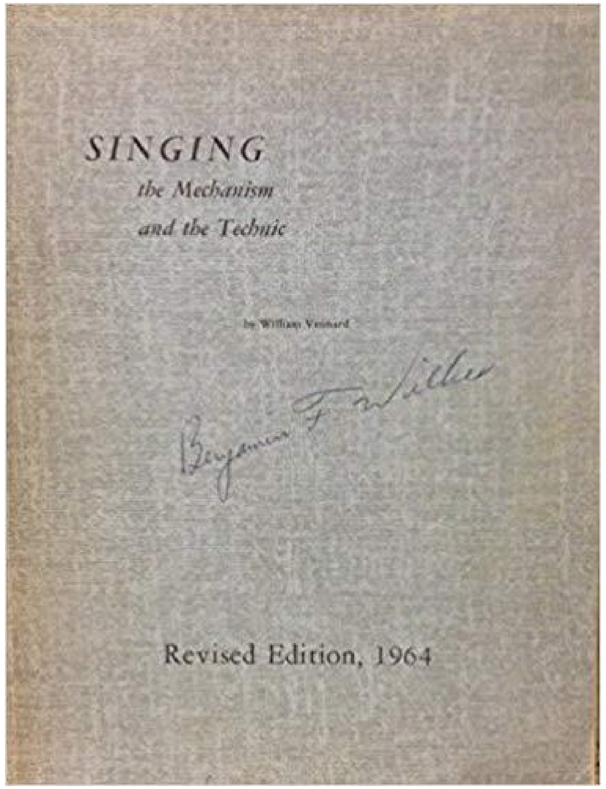


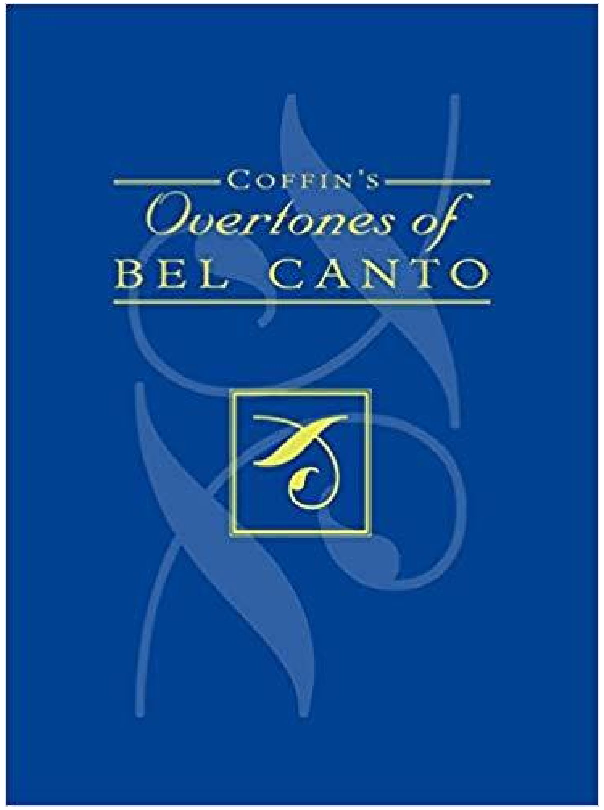


Manuel Garcia, 1805-1906
The “Father” of Voice Science

Tools for Recording and Analysis: Edison Phonograph and Bell Telephone

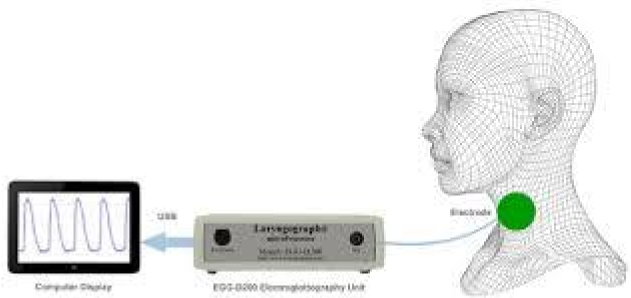


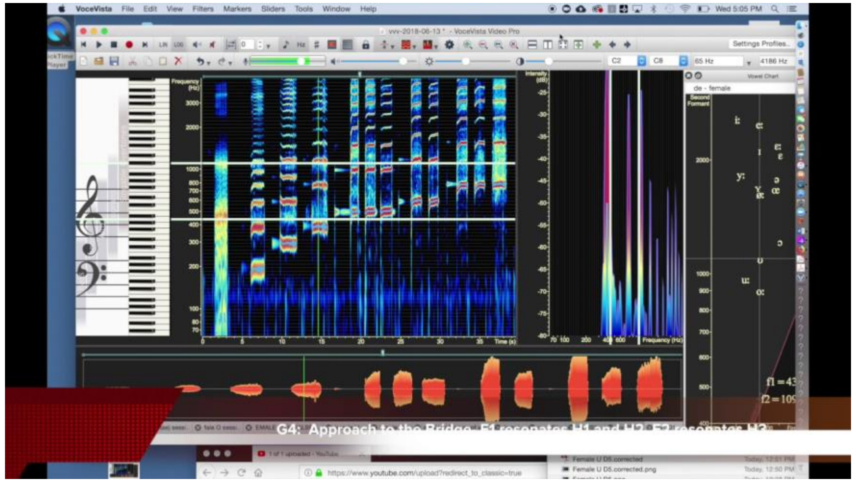


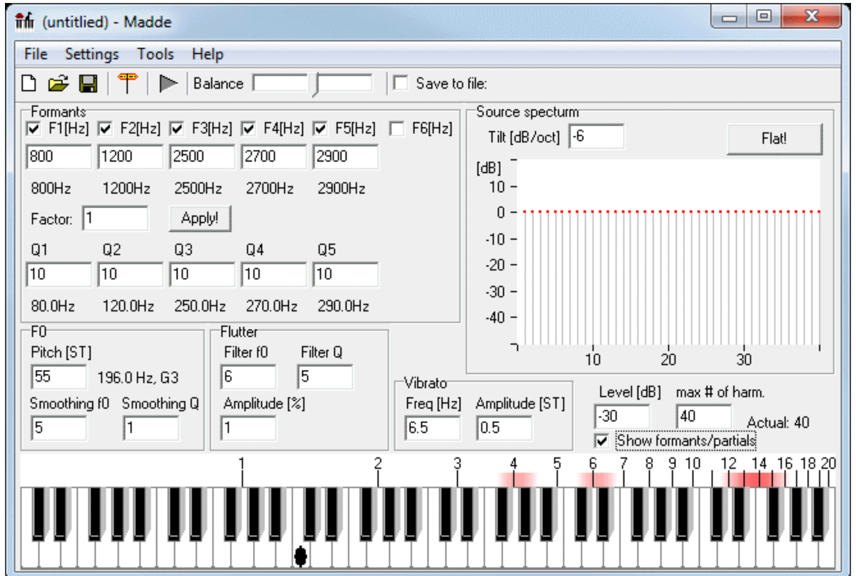


From Linguistics and Acoustic Phonetics to Vocal Pedagogy

New Technology – New Information

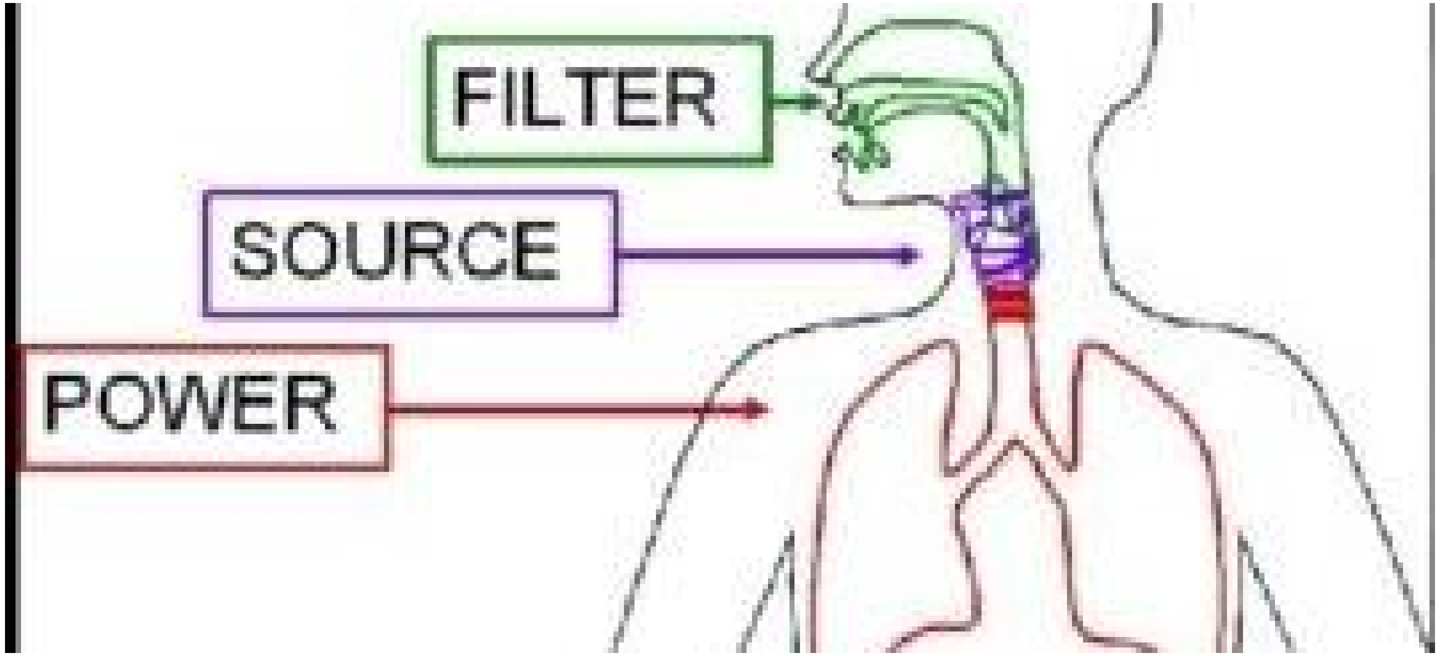






Singing Voice Science

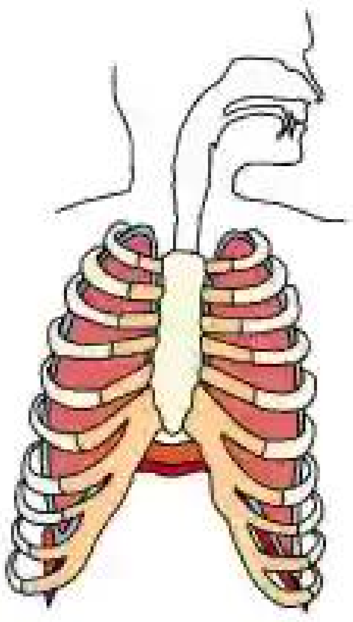
Sound Production and Acoustics

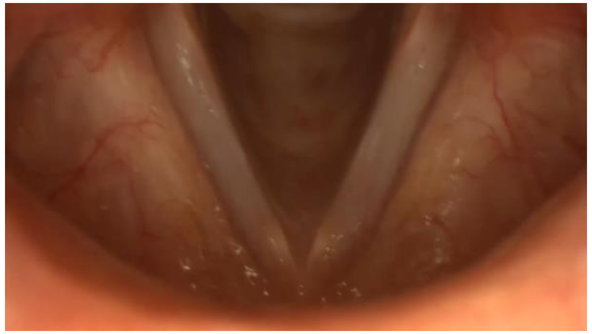


3 Parts of the Vocal Intrument

power source, sound source, and filter or resonator

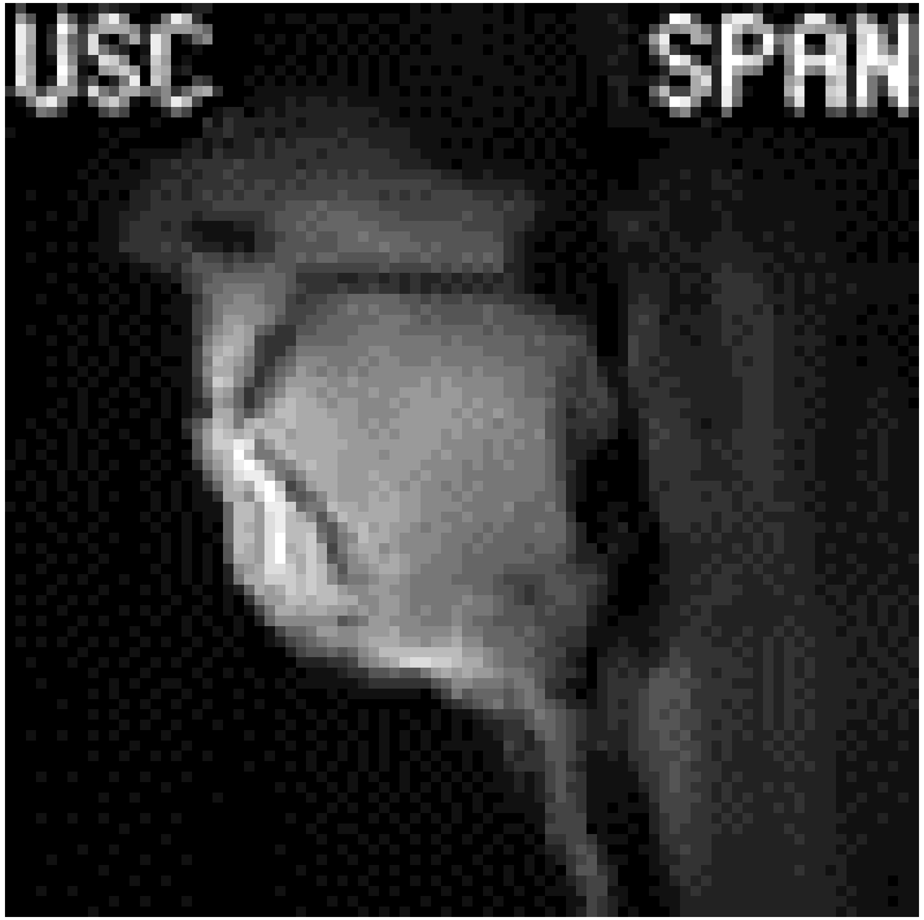
Power Source – Structures of Inhalation and Exhalation

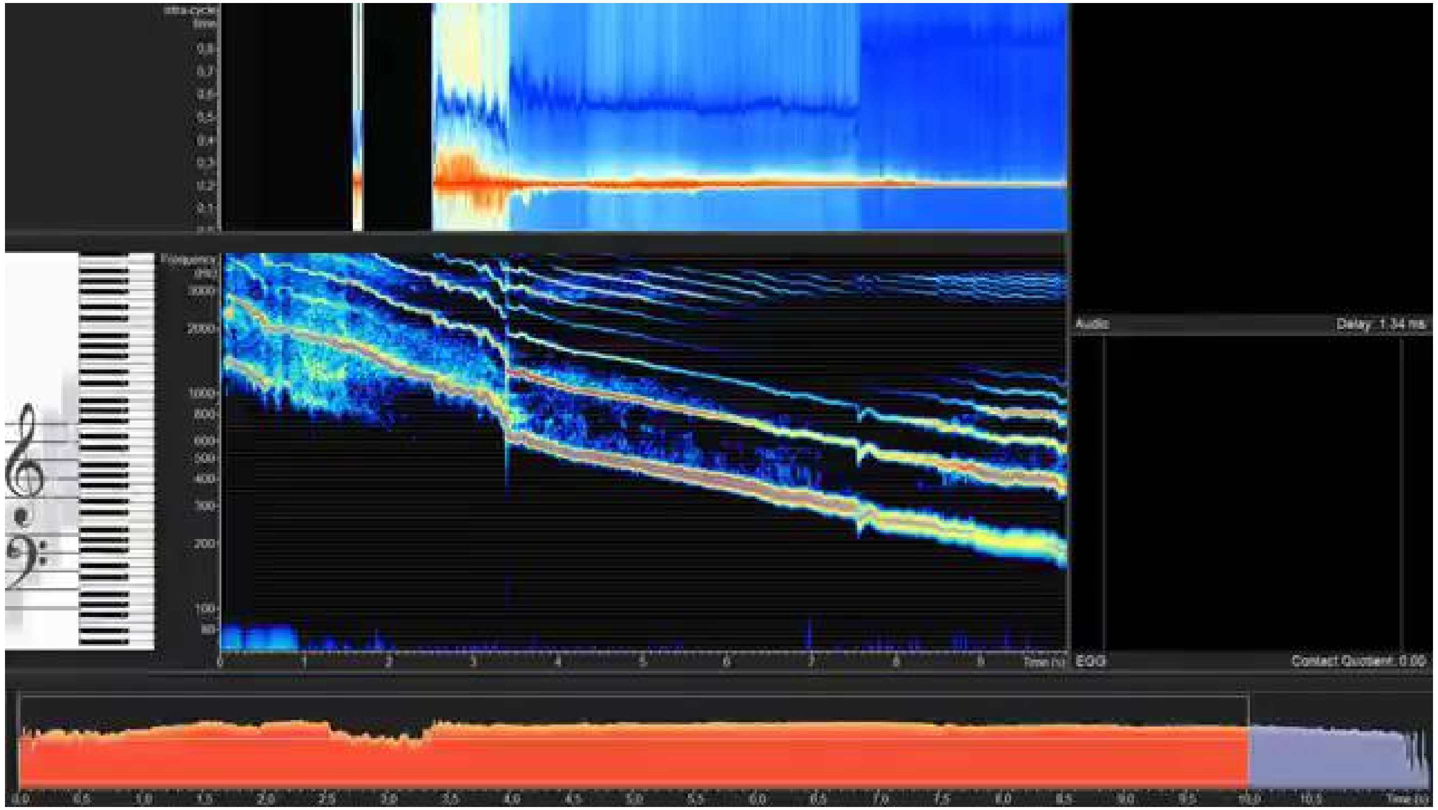




Sound Source – Musical Tone and Noise

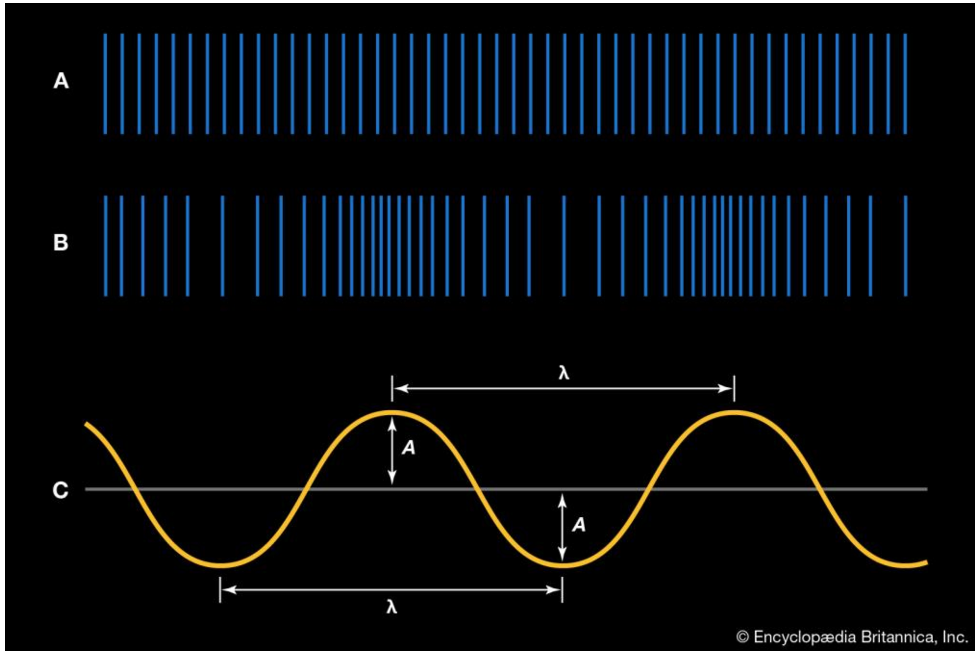
Resonator / Filter

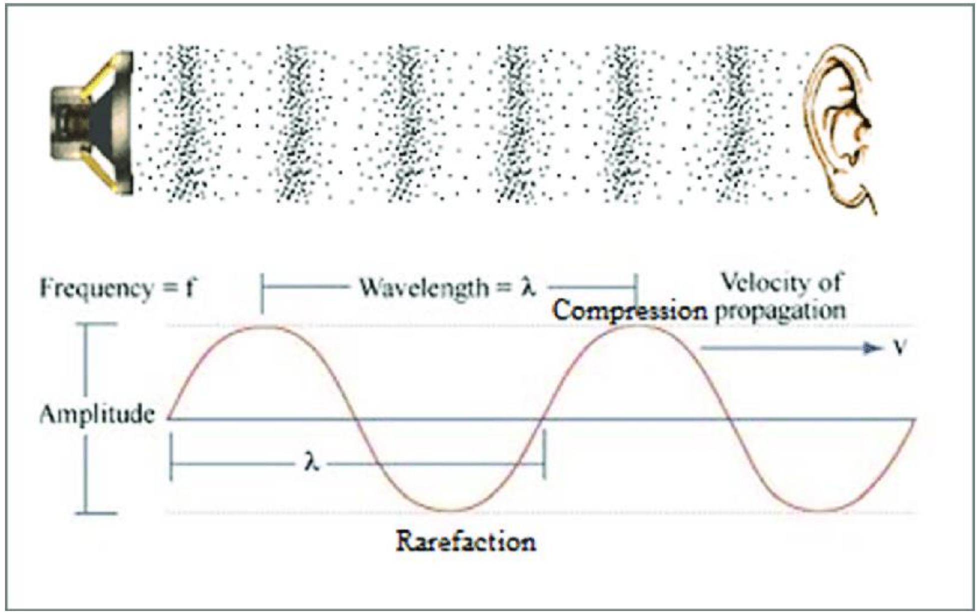




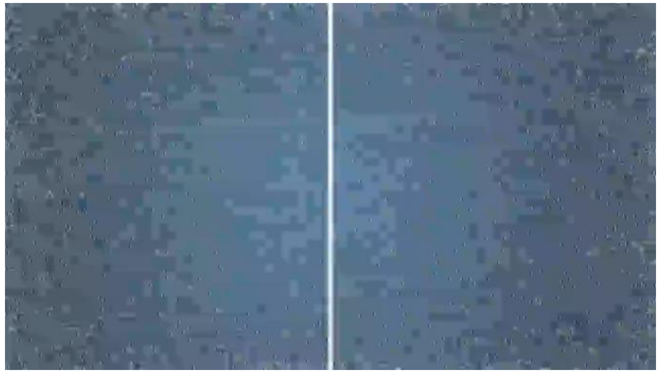
Singing Voice Acoustics

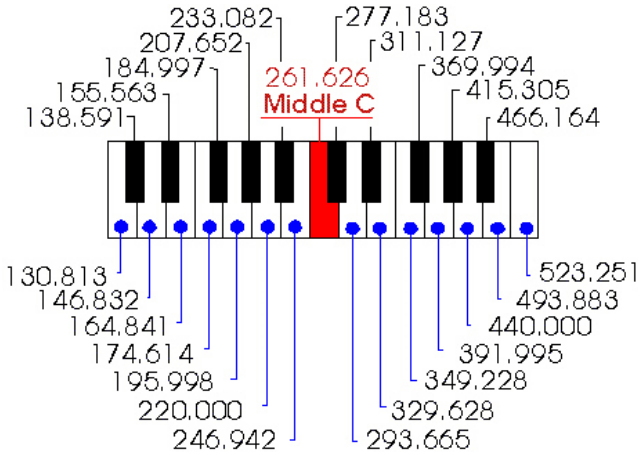
The Basics





Sound Waves – the "raw materials"

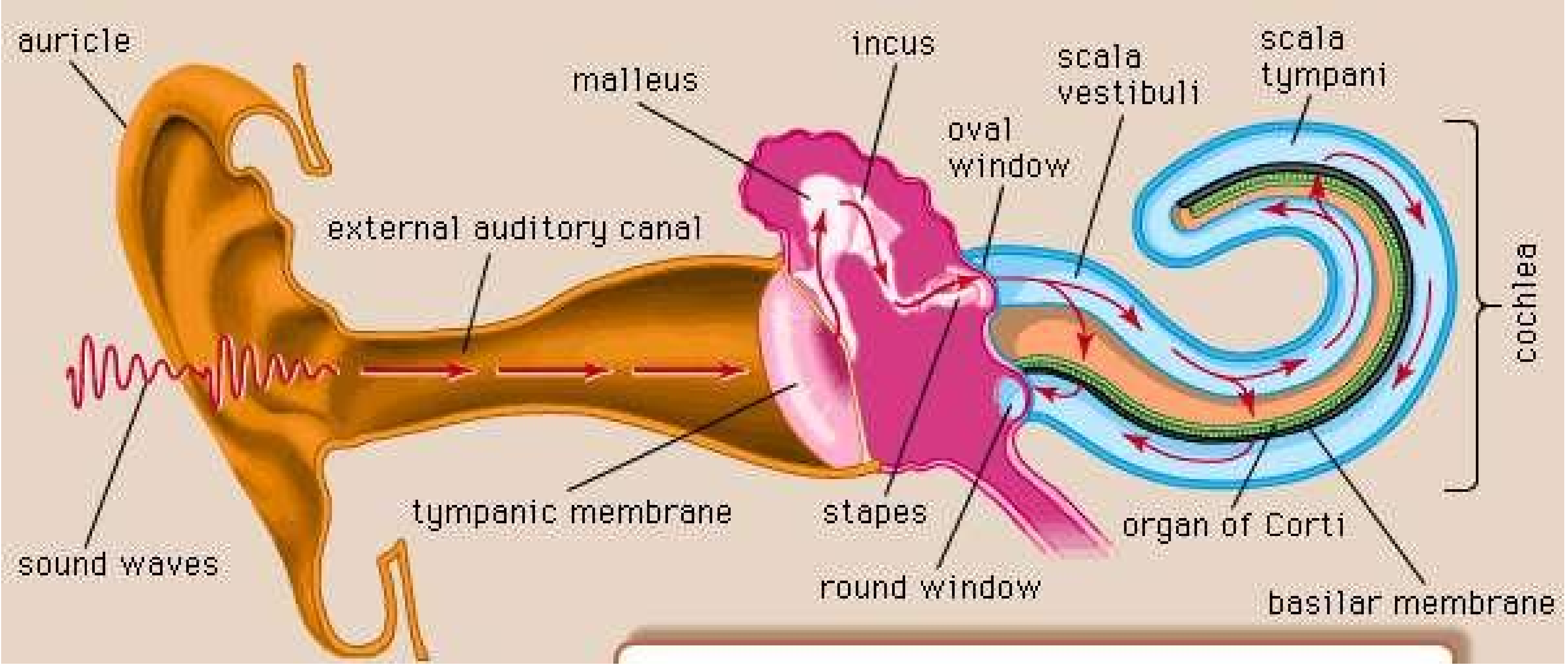




Note Frequency Chart

	Octave 0	Octave 1	Octave 2	Octave 3	Octave 4	Octave 5	Octave 6	Octave 7	Octave 8
C	16.35	32.70	65.41	130.81	261.63	523.25	1046.50	2093.00	4186.01
C#	17.32	34.65	69.30	138.59	277.18	554.37	1108.73	2217.46	4434.92
D	18.35	36.71	73.42	146.83	293.66	587.33	1174.66	2349.32	4698.64
D#	19.45	38.89	77.78	155.56	311.13	622.25	1244.51	2489.02	4978.03
E	20.60	41.20	82.41	164.81	329.63	659.26	1318.51	2637.02	5274.04
F	21.83	43.65	87.31	174.61	349.23	698.46	1396.91	2793.83	5587.65
F#	23.12	46.25	92.50	185.00	369.99	739.99	1479.98	2959.96	5919.91
G	24.50	49.00	98.00	196.00	392.00	783.99	1567.98	3135.96	6271.93
G#	25.96	51.91	103.83	207.65	415.30	830.61	1661.22	3322.44	6644.88
A	27.50	55.00	110.00	220.00	440.00	880.00	1760.00	3520.00	7040.00
A#	29.14	58.27	116.54	233.08	466.16	932.33	1864.66	3729.31	7458.62
B	30.87	61.74	123.47	246.94	493.88	987.77	1975.53	3951.07	7902.13

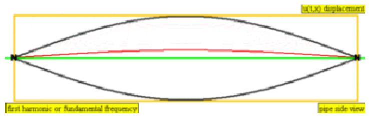
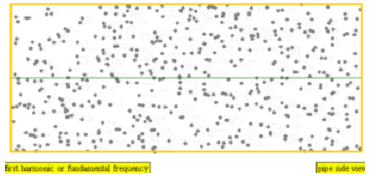
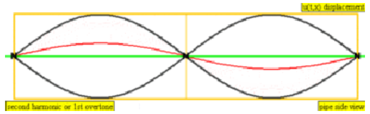
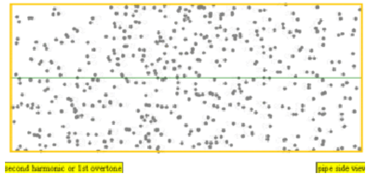
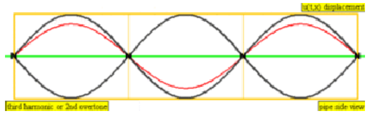
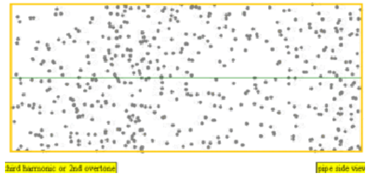
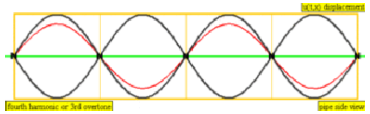
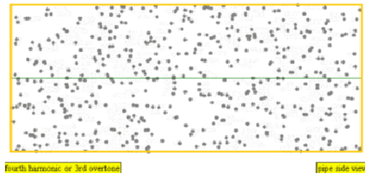
Frequency = how many times something happens



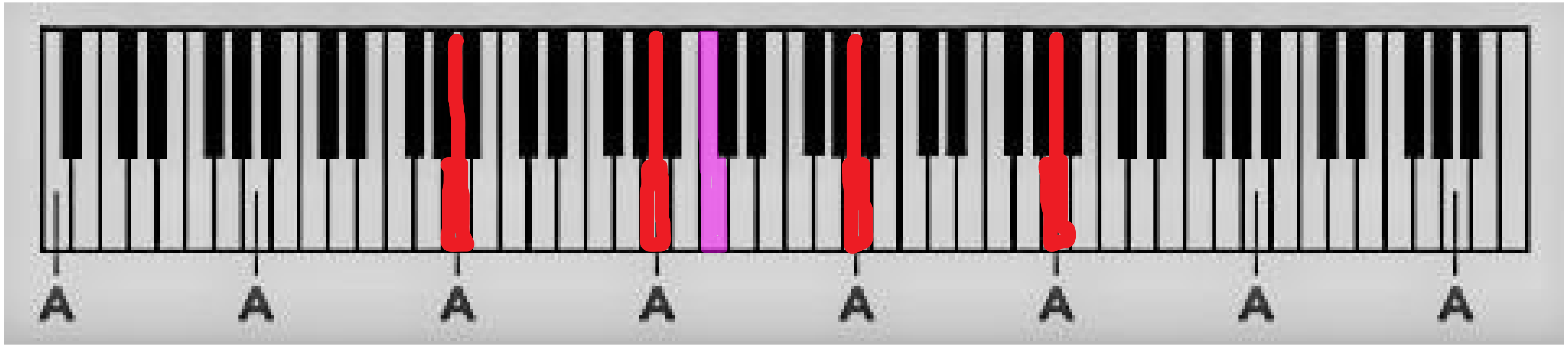
Sound Waves and the Ear

Complex Wave = Superimposed Frequencies

Source Harmonics = Funtamental Frequency + Overtones

Frequency	Order	Name 1	Name 2	Name 3	Wave Representation	Molecular Representation
1 × f = 440 Hz	n = 1	1st partial	fundamental tone	1st harmonic		
2 × f = 880 Hz	n = 2	2nd partial	1st overtone	2nd harmonic		
3 × f = 1320 Hz	n = 3	3rd partial	2nd overtone	3rd harmonic		
4 × f = 1760 Hz	n = 4	4th partial	3rd overtone	4th harmonic		

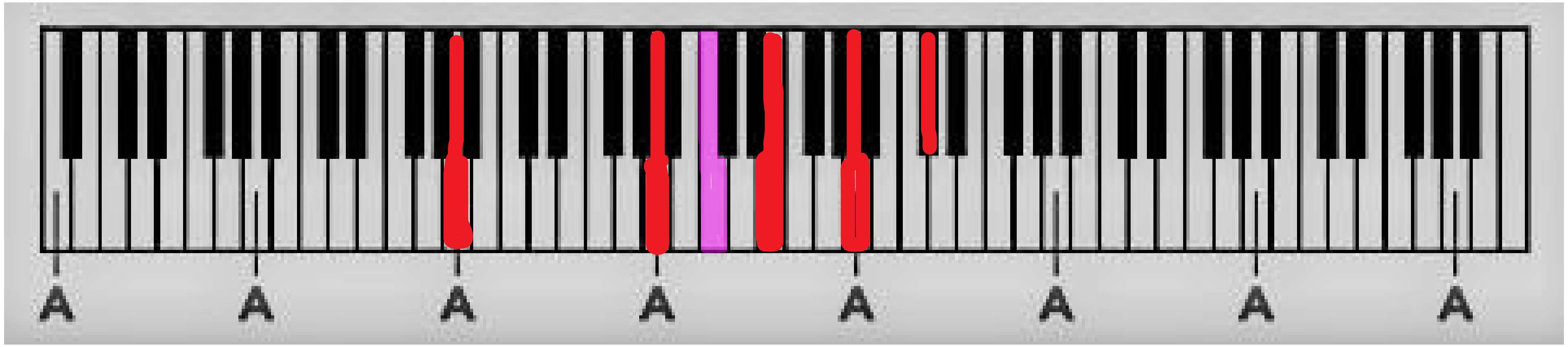
Octaves – Logarithmic Spacing



A2	A3	A4	A5
110	220	440	880
x	2x	4x	8x

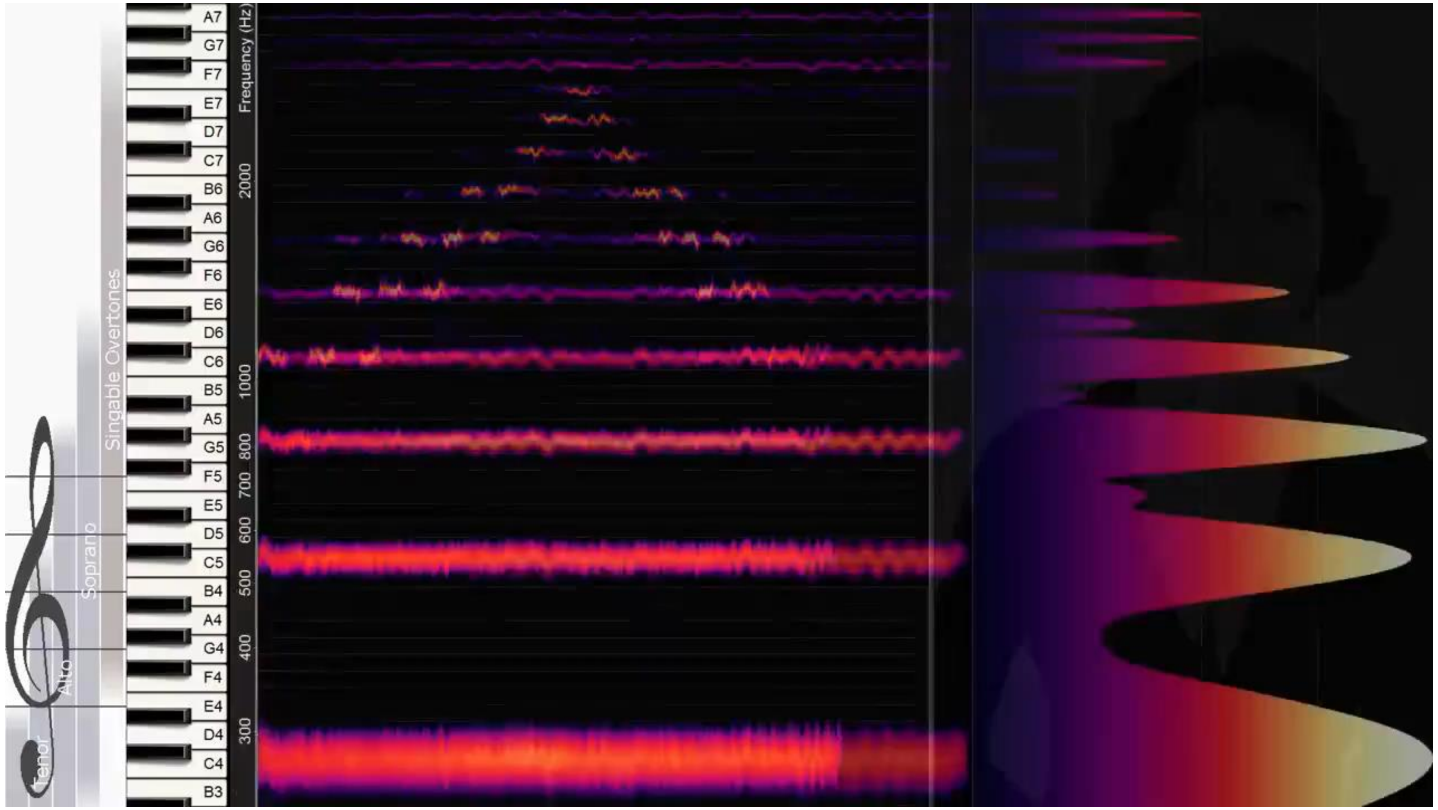
Each successive octave is *double* the number of cycles per second (Hz) as the previous octave. While the distance between them appears the same, there are twice the number of vibrations.

Harmonics – Simple Multiplier

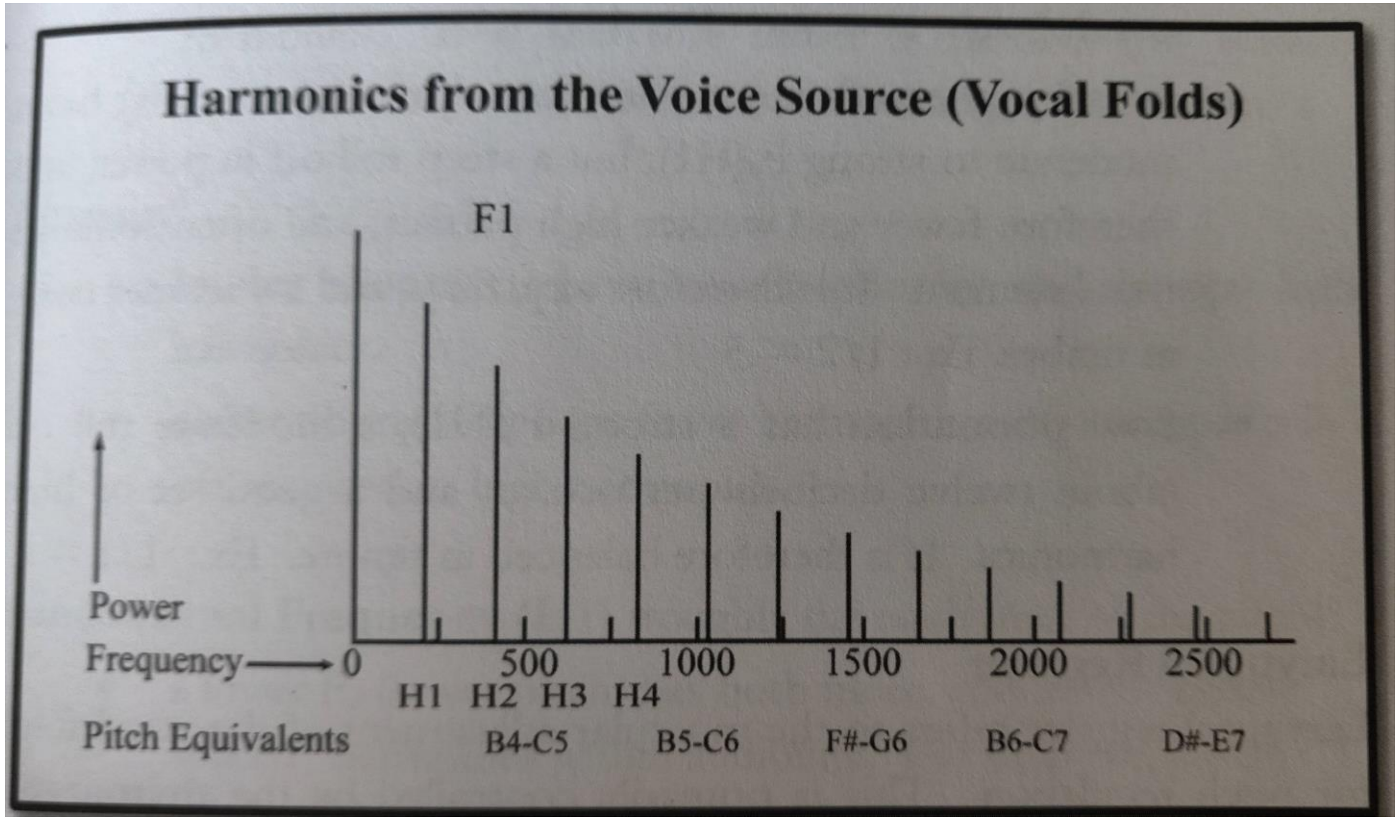


A2	A3	E4	A4	C#5
110	220	330	440	550
x	2x	3x	4x	5x
F0/H1	*H2*	*H3*	*H4*	*H5*

As harmonics are consecutive multiples of the fundamental frequency, the spacing appears closer as the harmonics get higher while the difference in cycles per second between them remains the same.

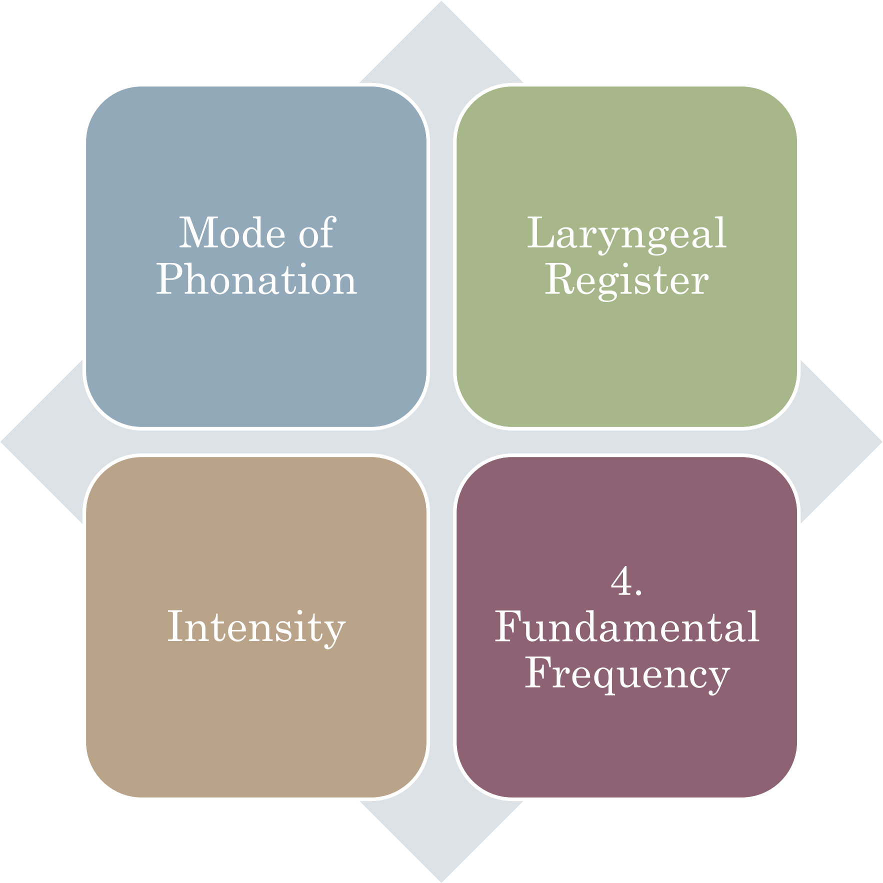


Isolating Source Harmonics Using Computerized Filter



Gradual Roll Off in Power: 12 dB per Octave

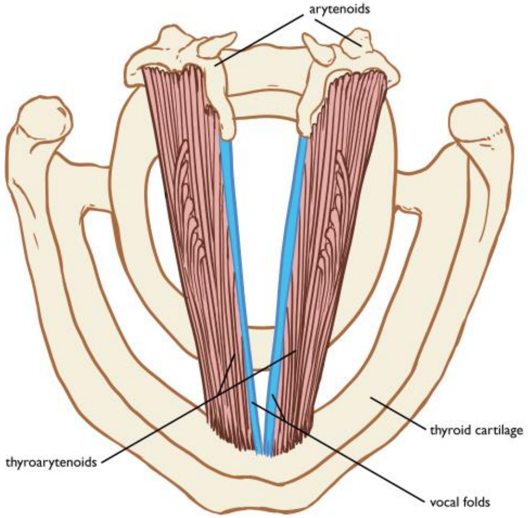
Factors Influencing Source Harmonics

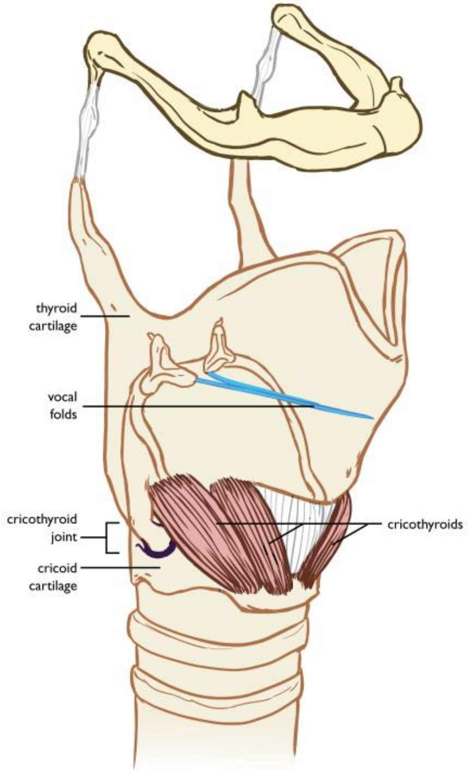


Mode of Phonation: relationship between how much air is supplied from the power source and how the flow of air is resisted by the sound source during phonation (closure and opening / vibration of the vocal folds)

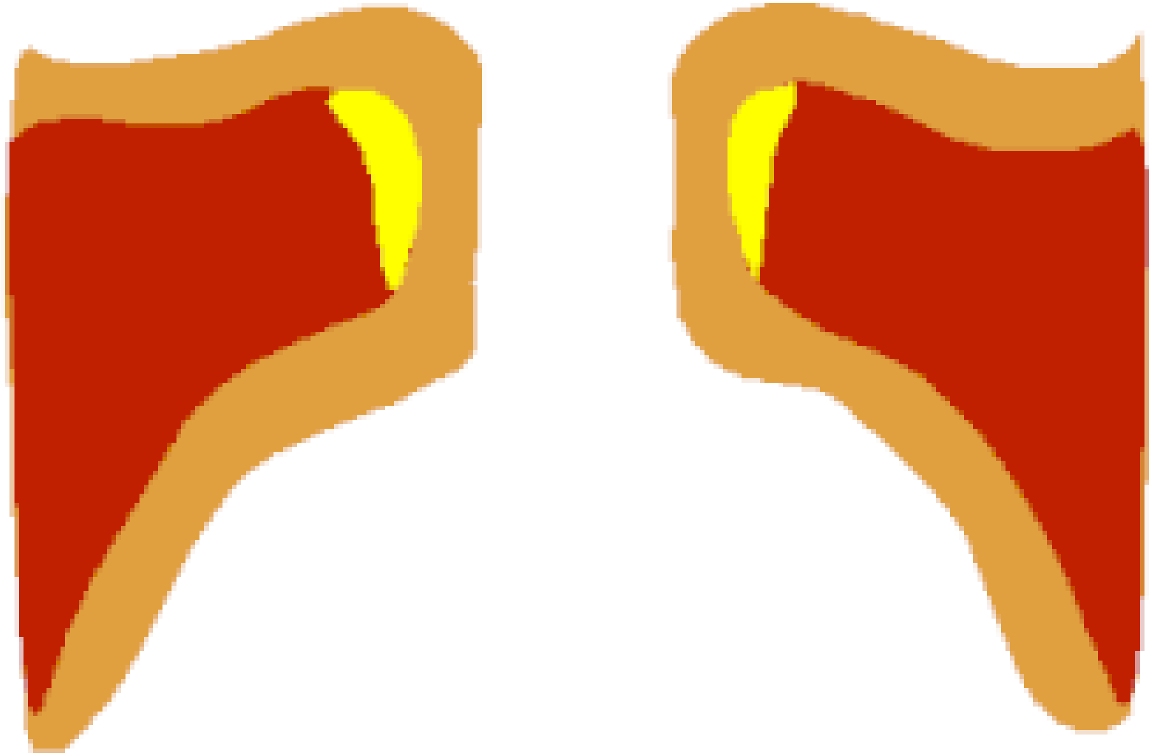
Hyper - adductive Phonation – Excessive Resistance

Hypo - adductive Phonation – Insufficient Resistance

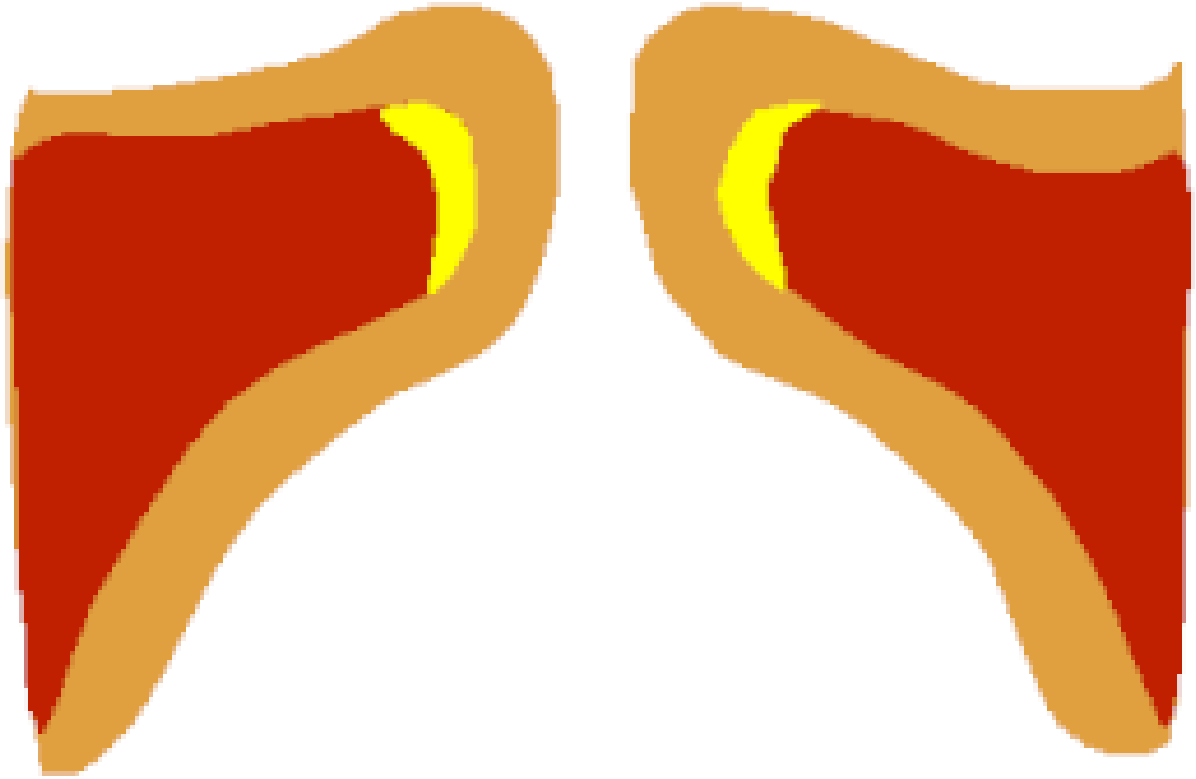




Laryngeal Register



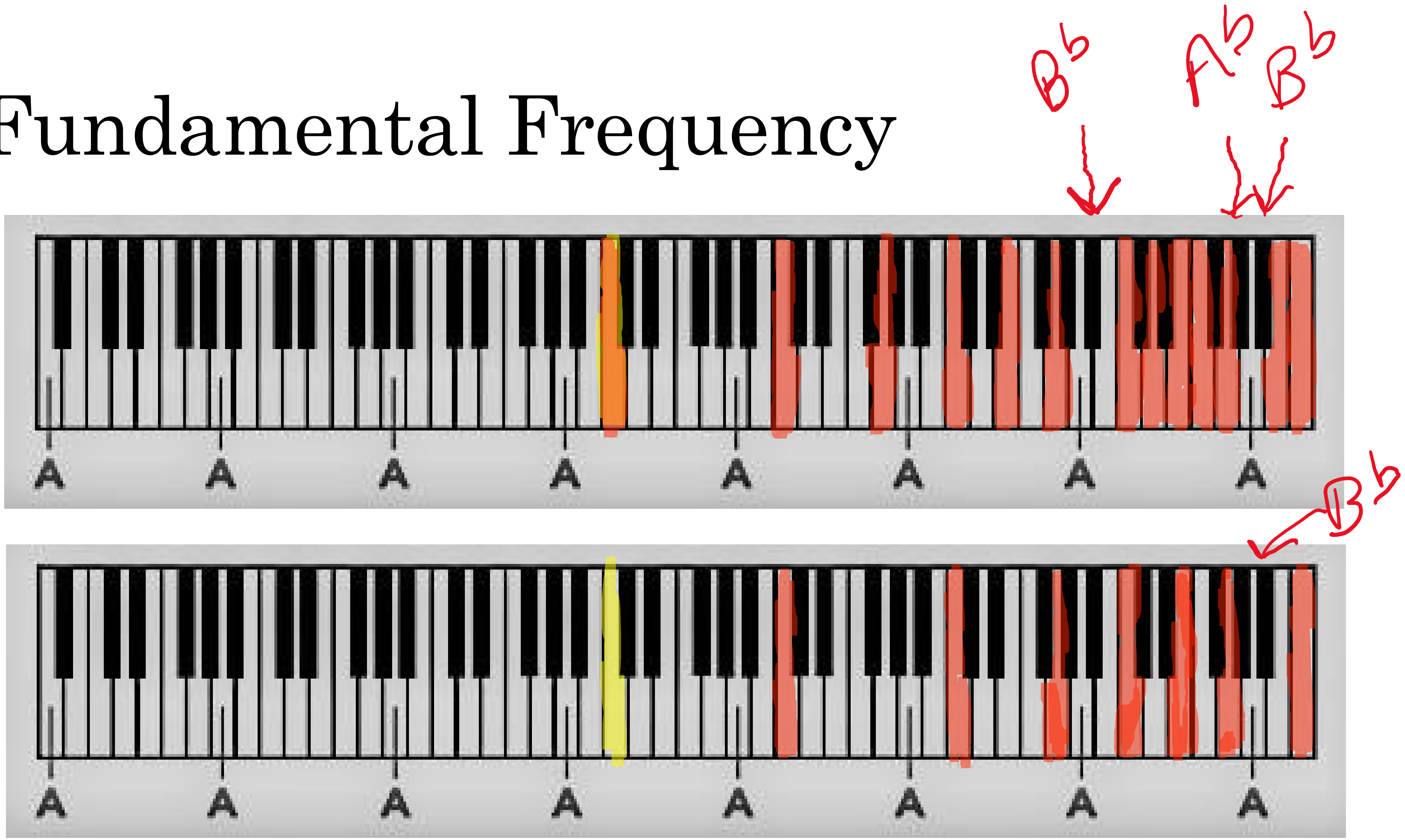
Vibrational Mode 1 or TA Dominant Phonation

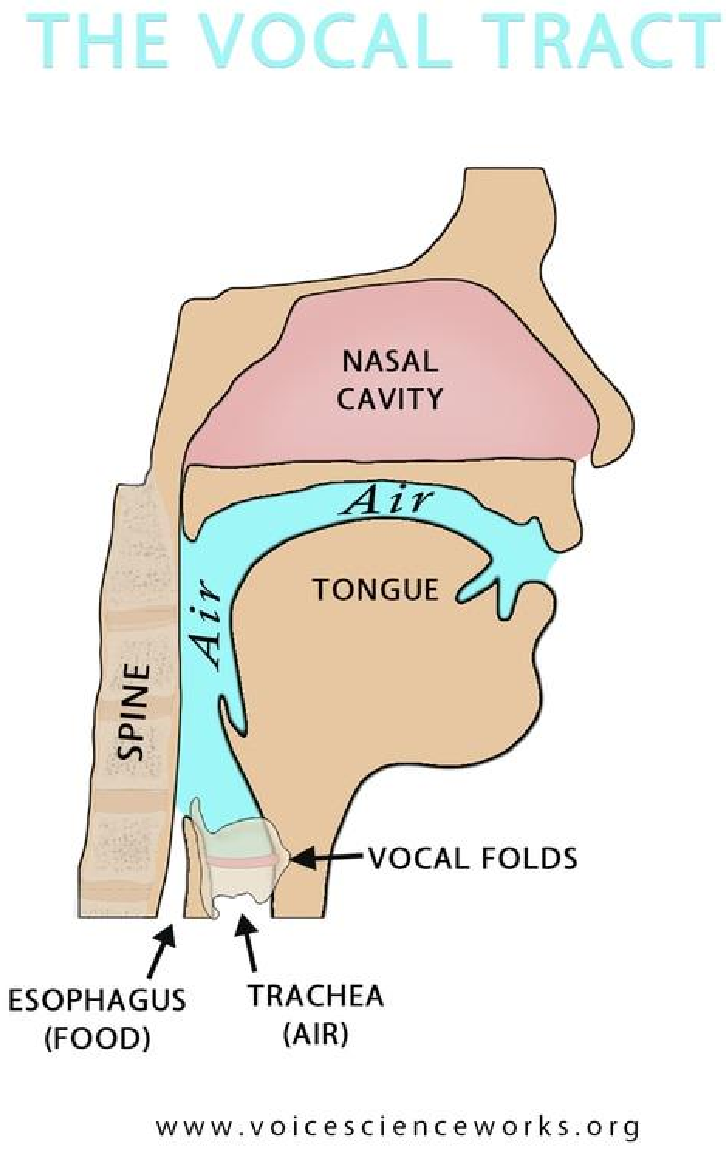


Vibrational Mode 2 or CT Dominant Phonation


Intensity: Loud vs. Soft

Fundamental Frequency

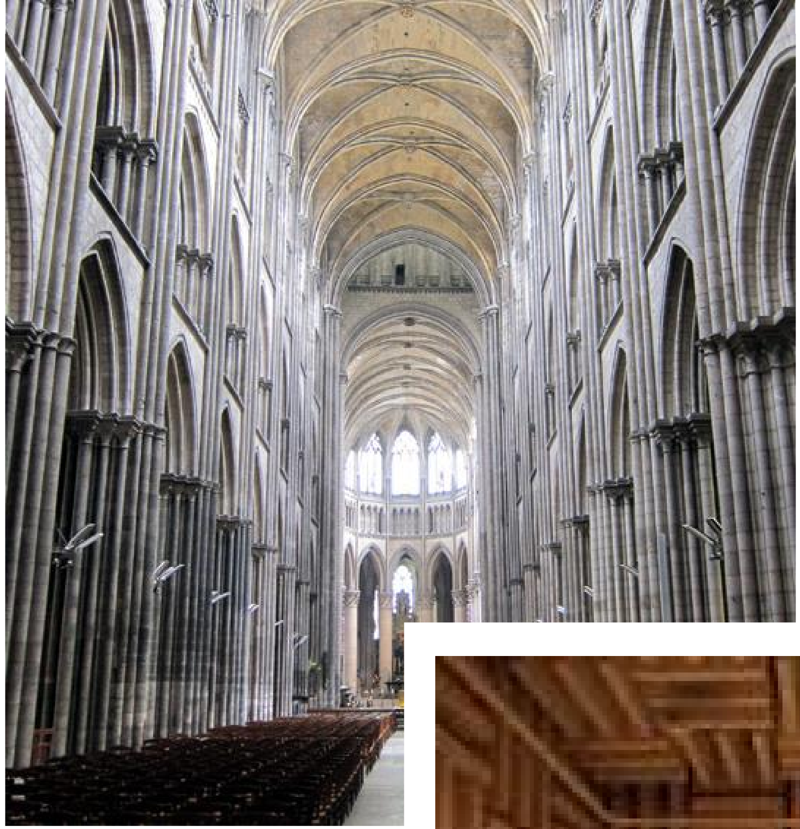


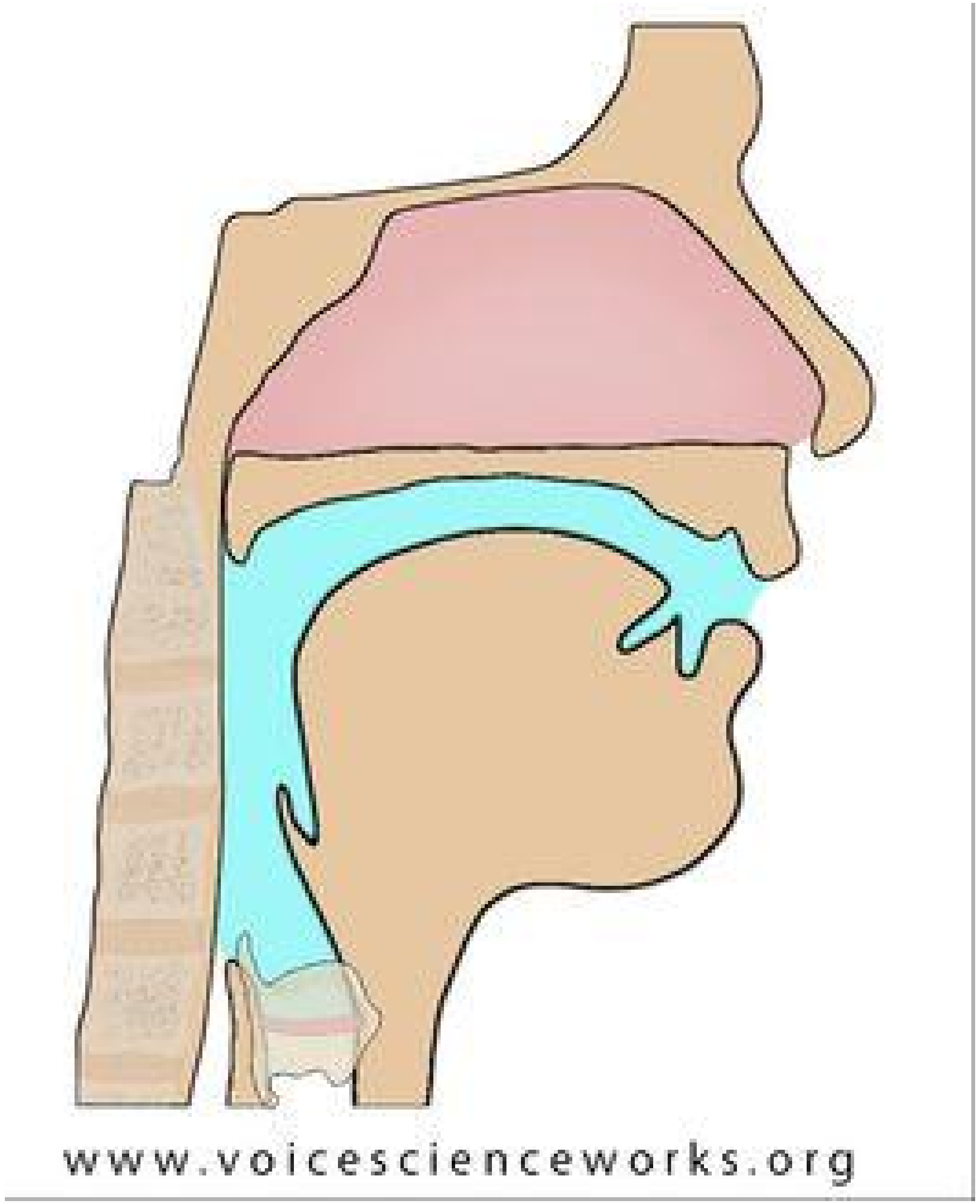


Resonance Defined

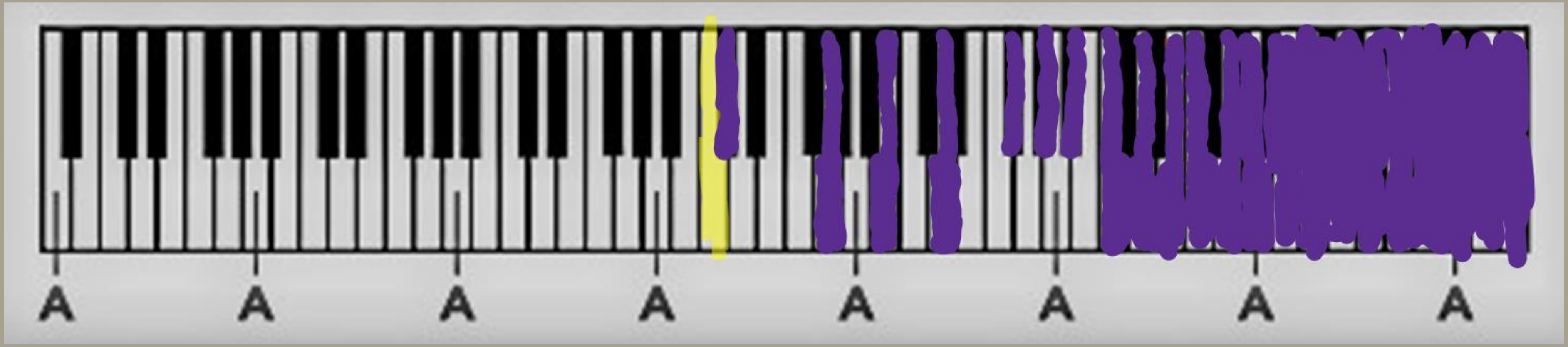
How the Resonator / Filter Works

The Vocal Tract – A Special Resonator

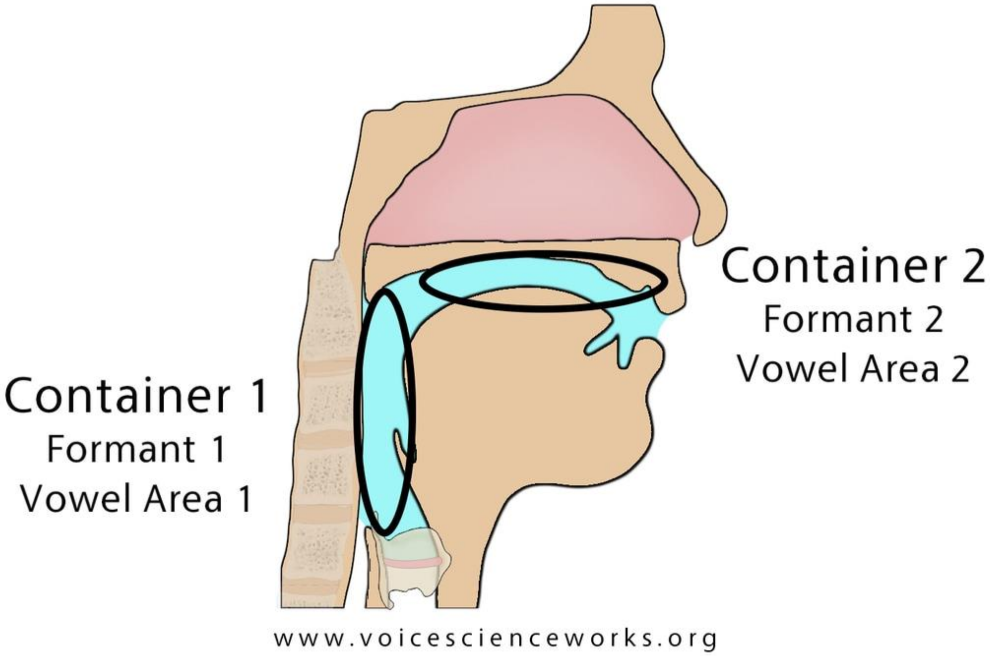




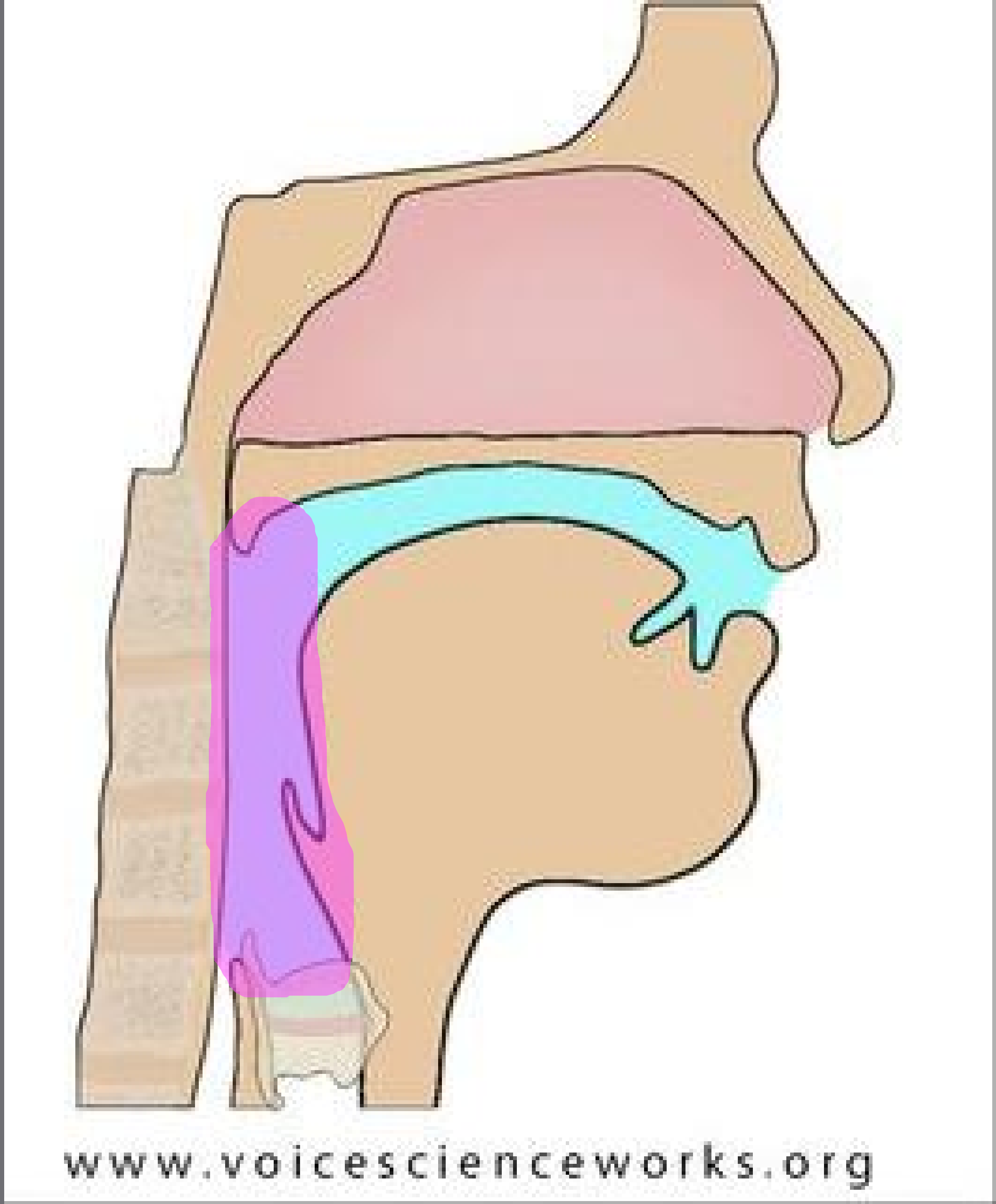
Tube or Tract Resonance



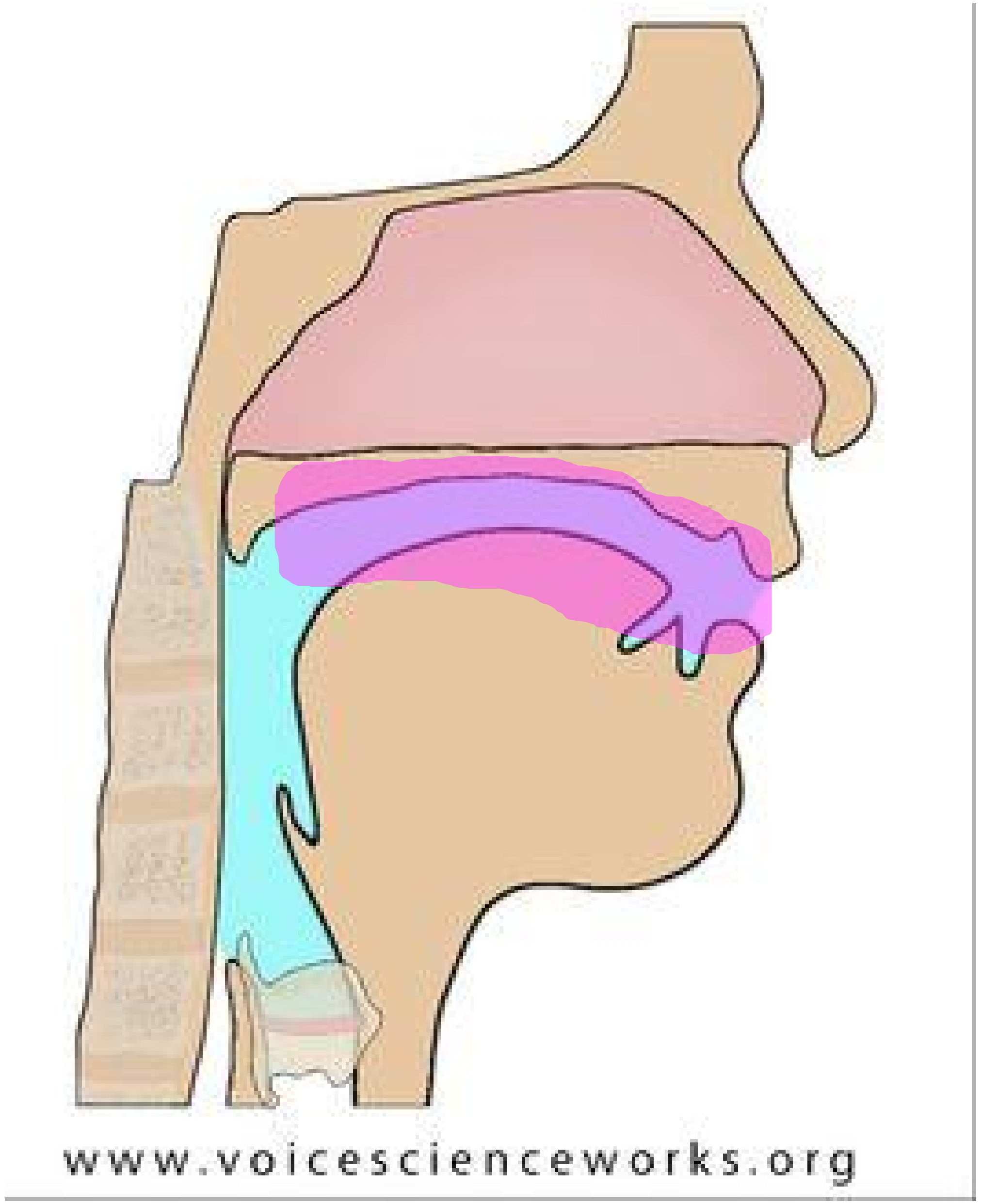
Quarter Wave Resonator:
¼, ¾, 5/4, 7/4, etc.



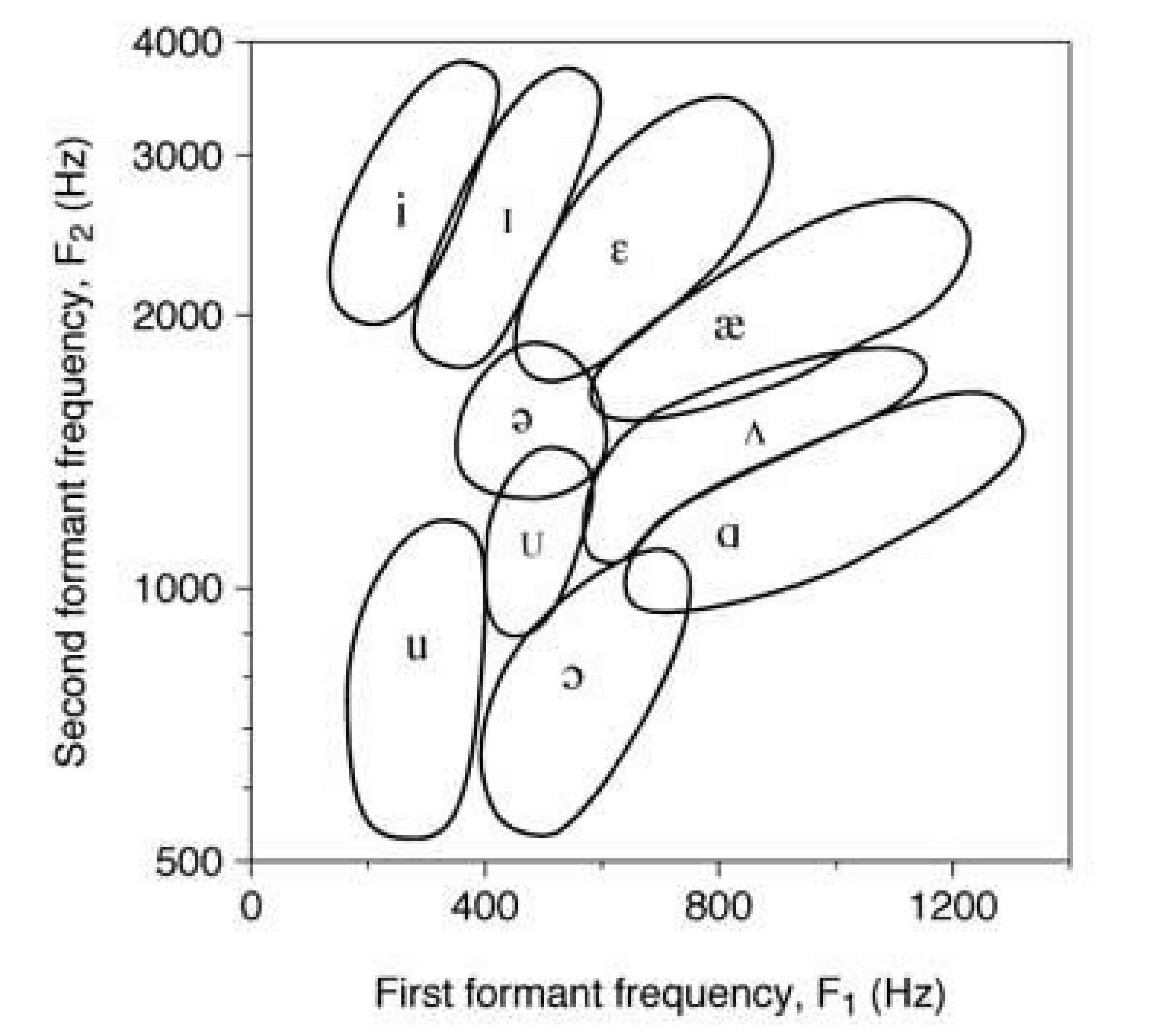
Vowel or “Formed” Resonance



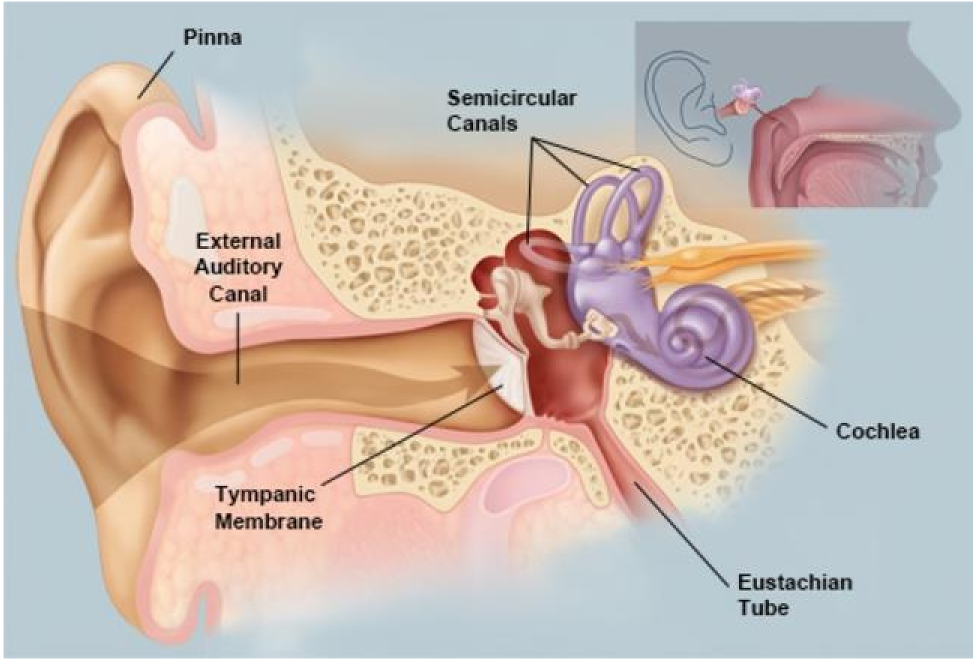
F1: Depth, Open / Close,& Efficiency



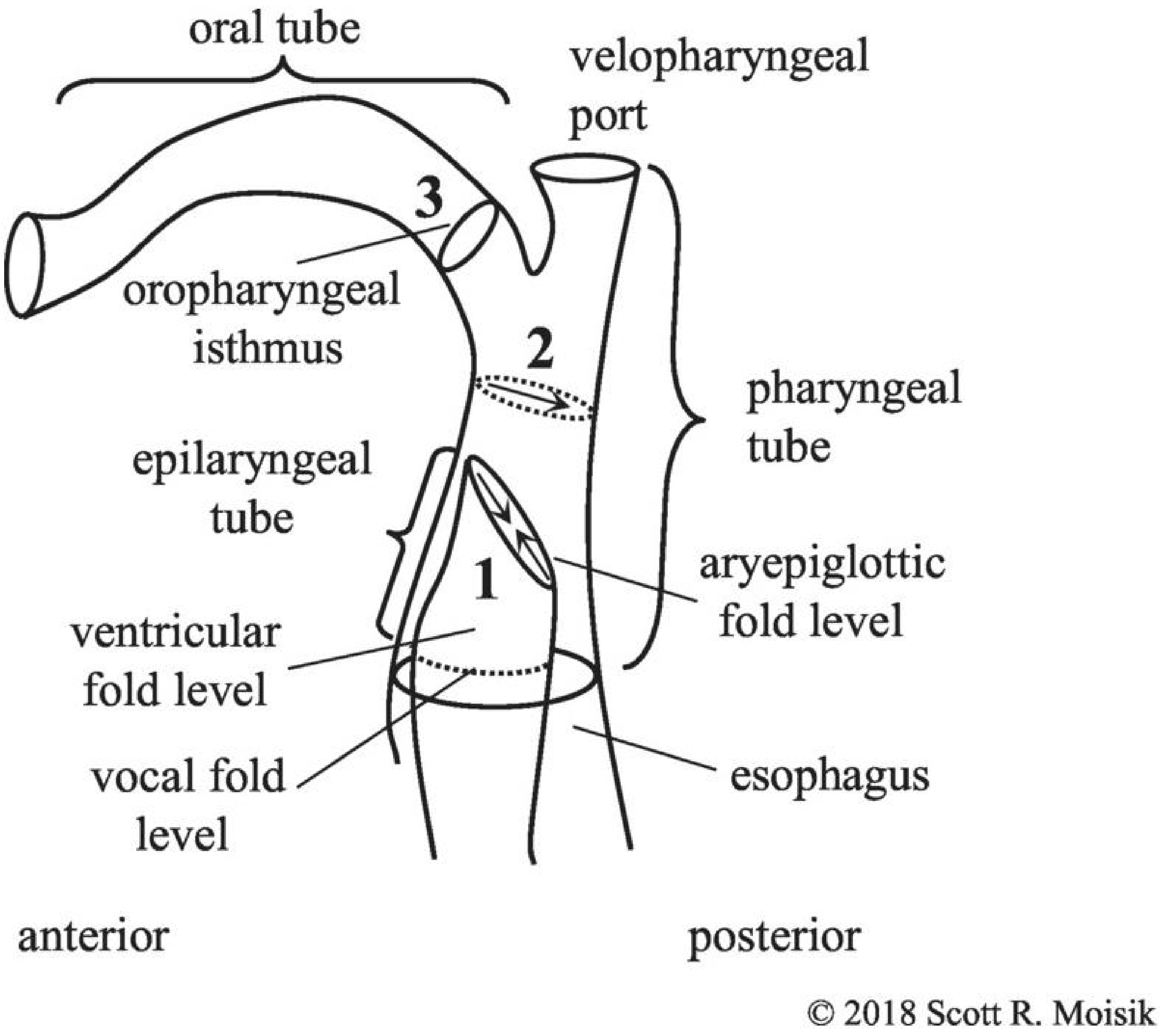
F2: Clarity, Front / Back, & Uppper Voice Strategies



Traditional Concept of Vowel Resonance



Singer’s Formant: a most unique resonance



Epilarynx: a 6 to 1 Ratio?

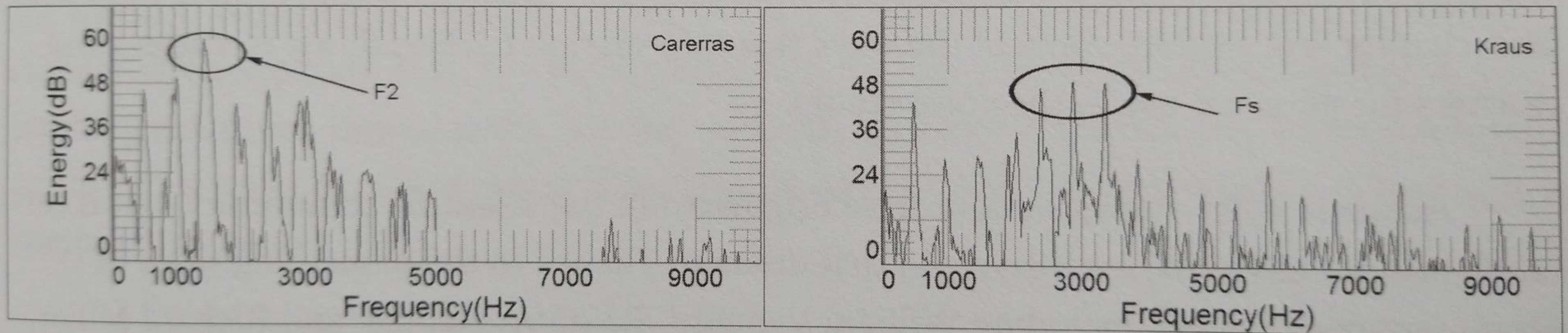
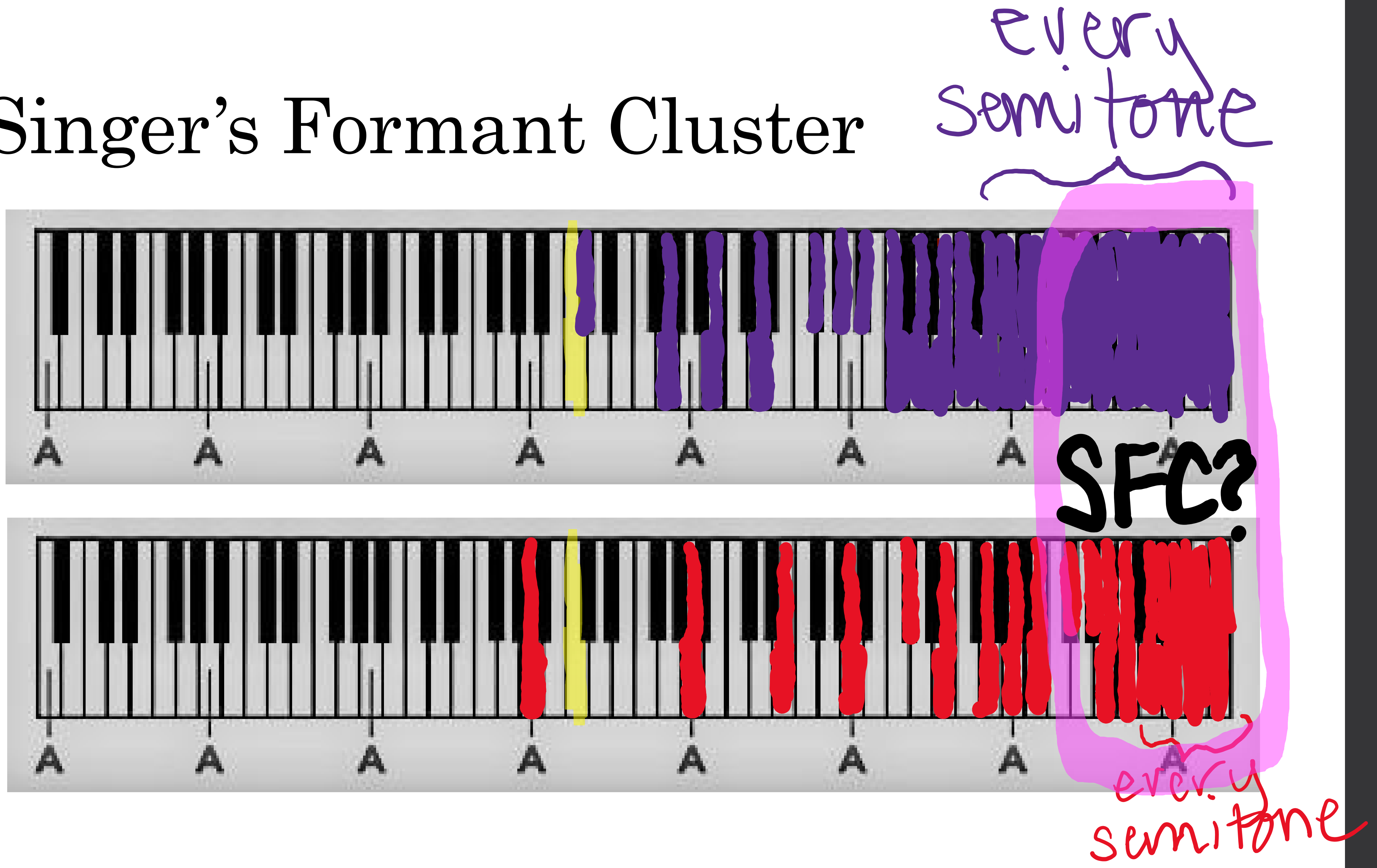


Figure 5.11: Spectra of Carerras and Kraus

Not only High Female Voices

Singer's Formant Cluster

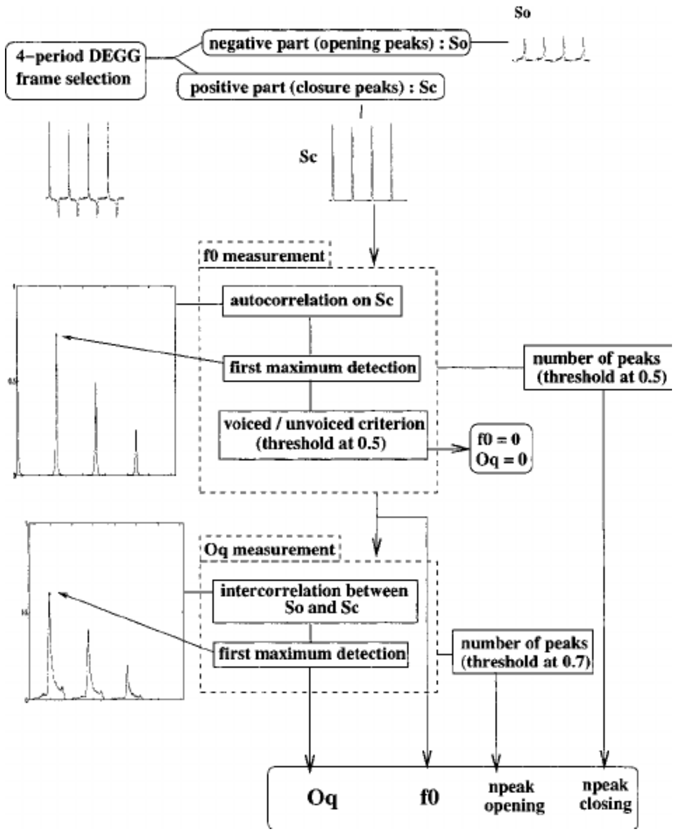


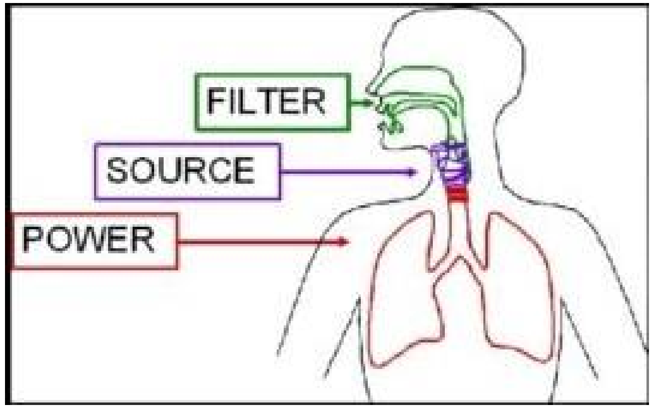
Vocal Tract Length and the SFC

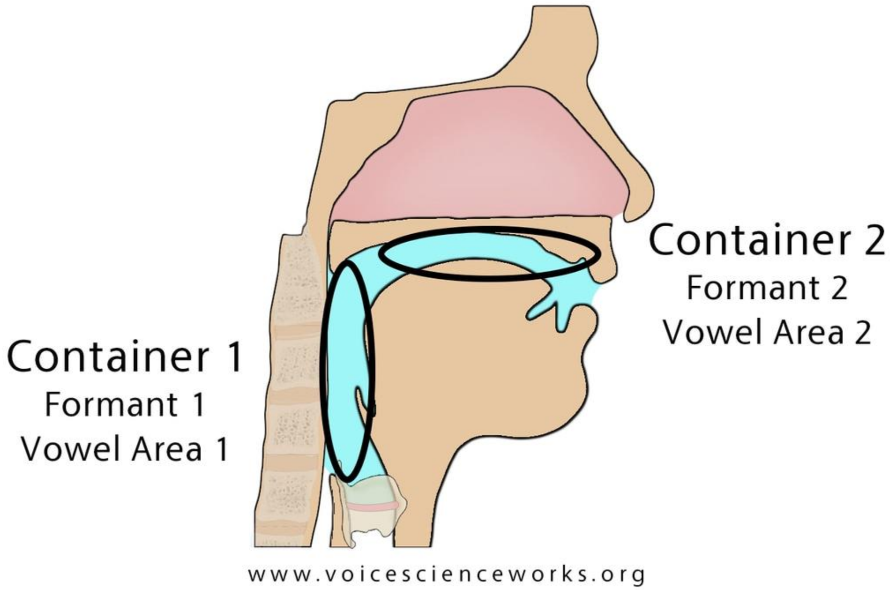
Long Tract: lower voice optimal?

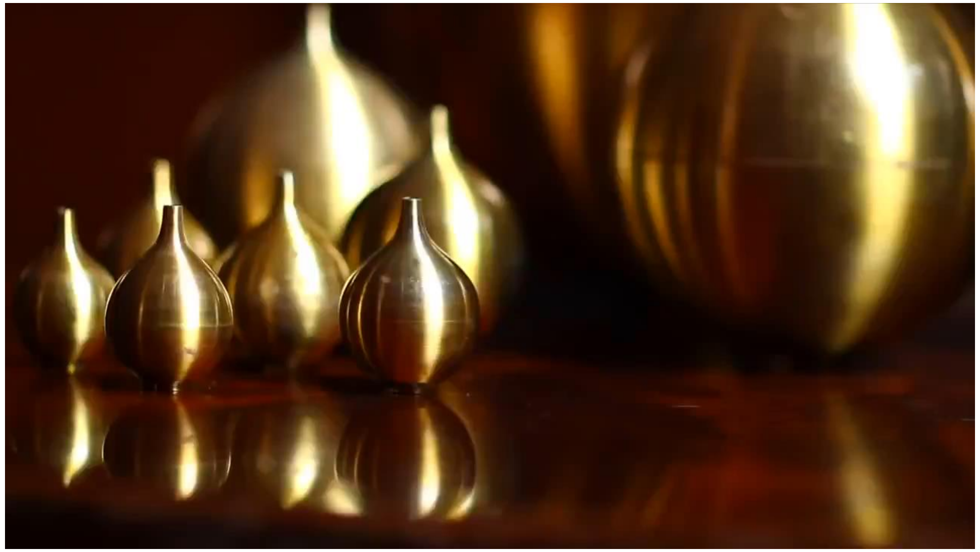
Short Tract: higher voice optimal?

Acoustic / Resonance Theories

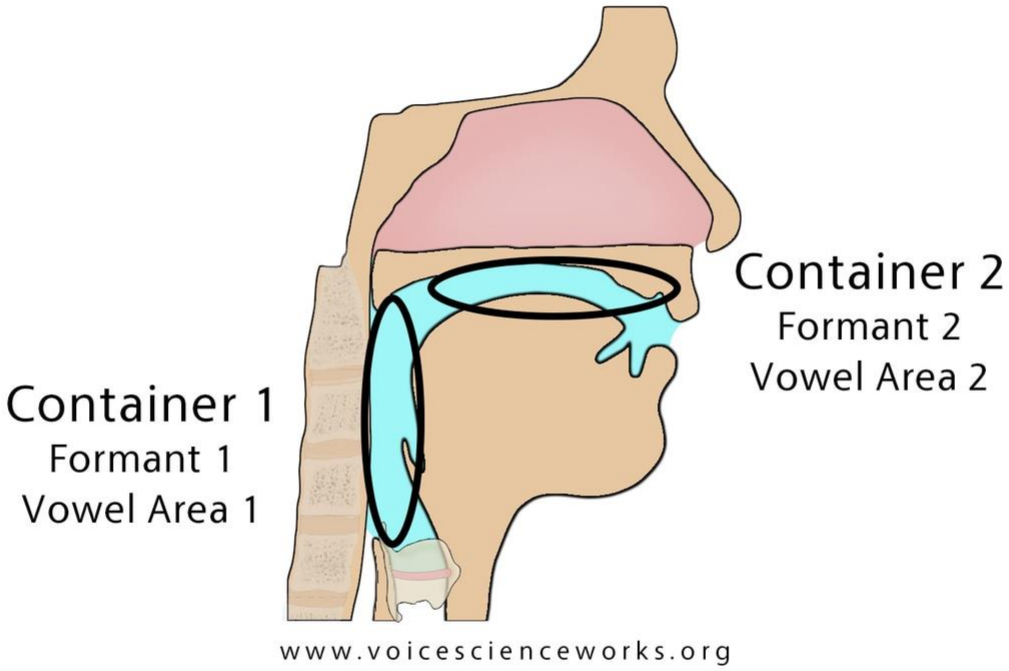


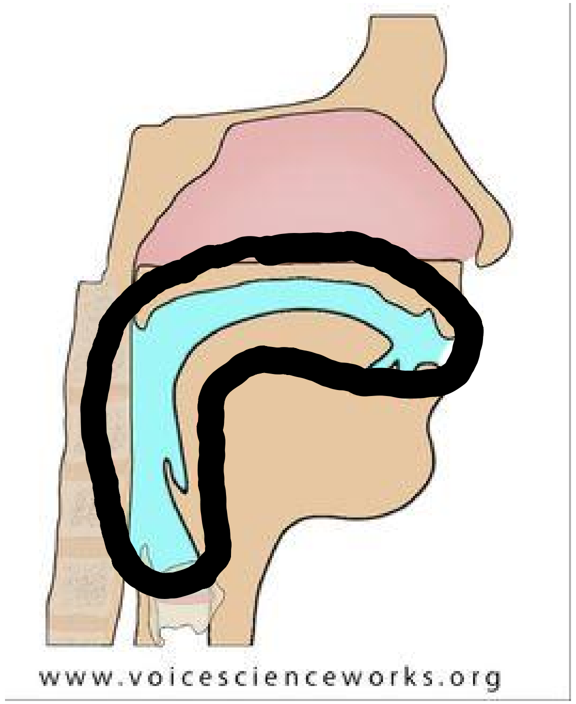




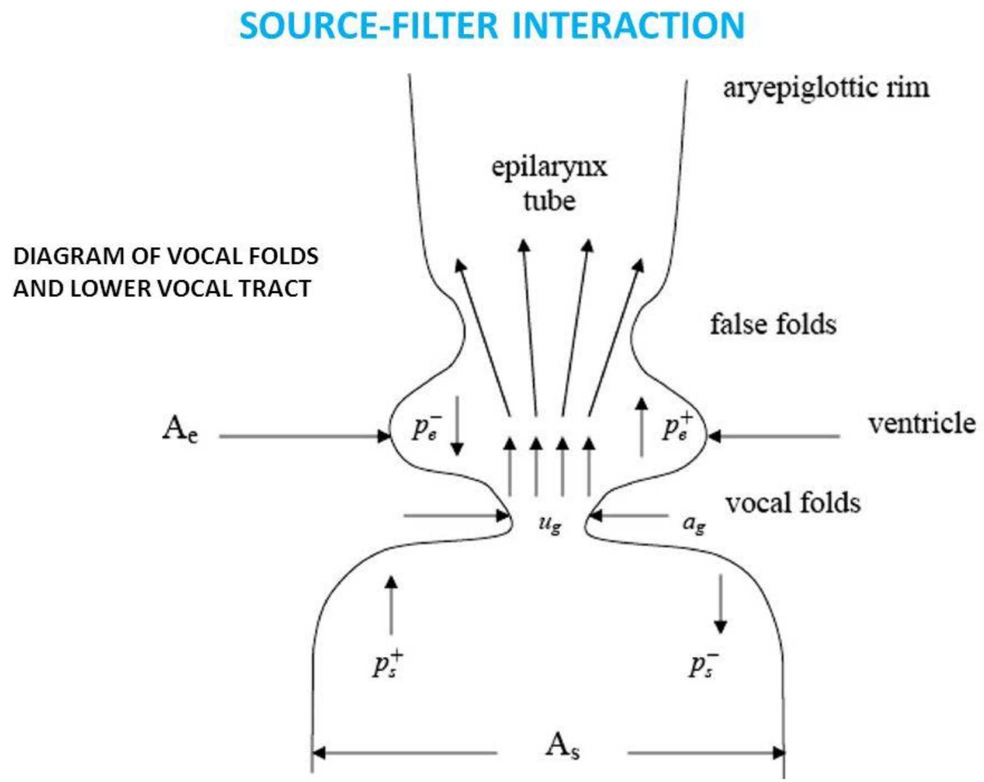


Coupled Helmholtz Resonators

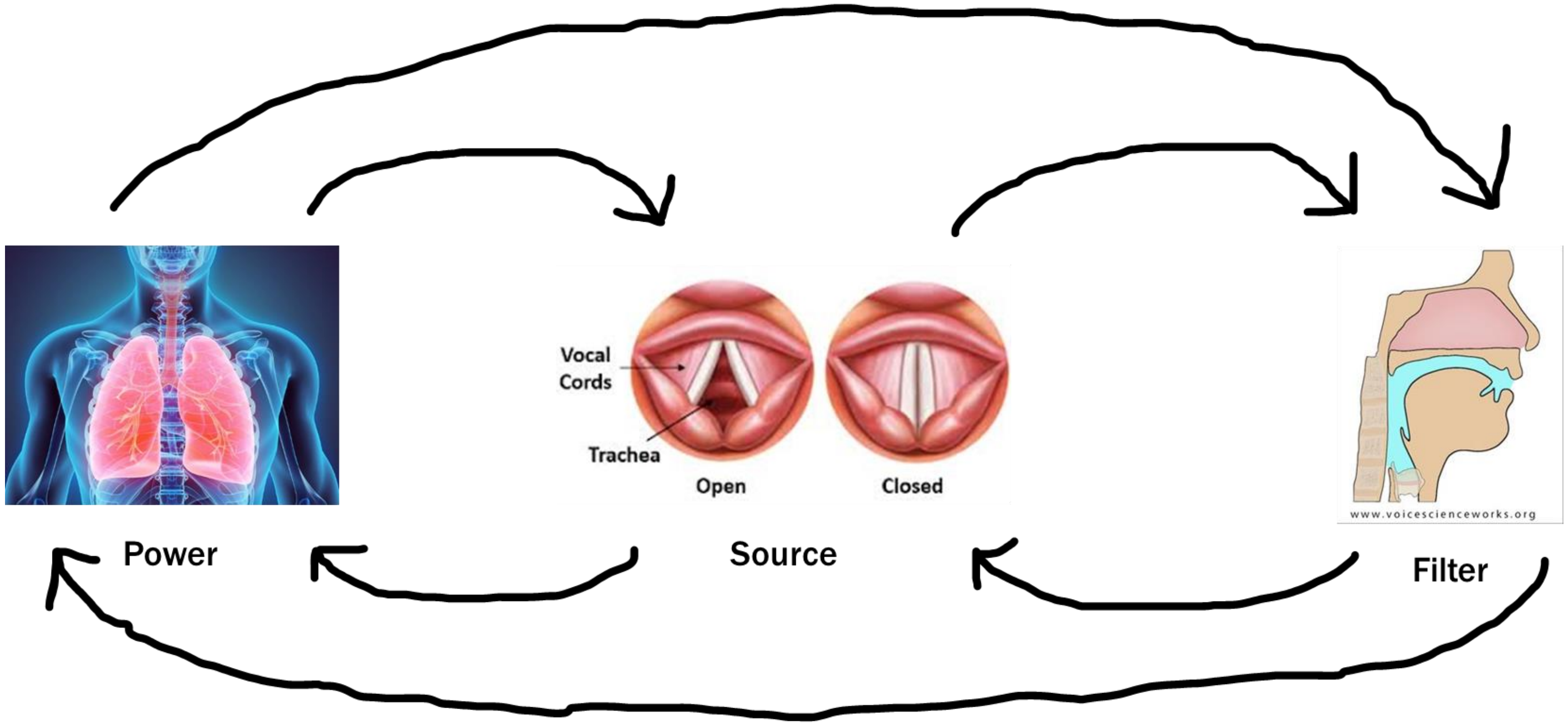




Linear Source – Filter Model



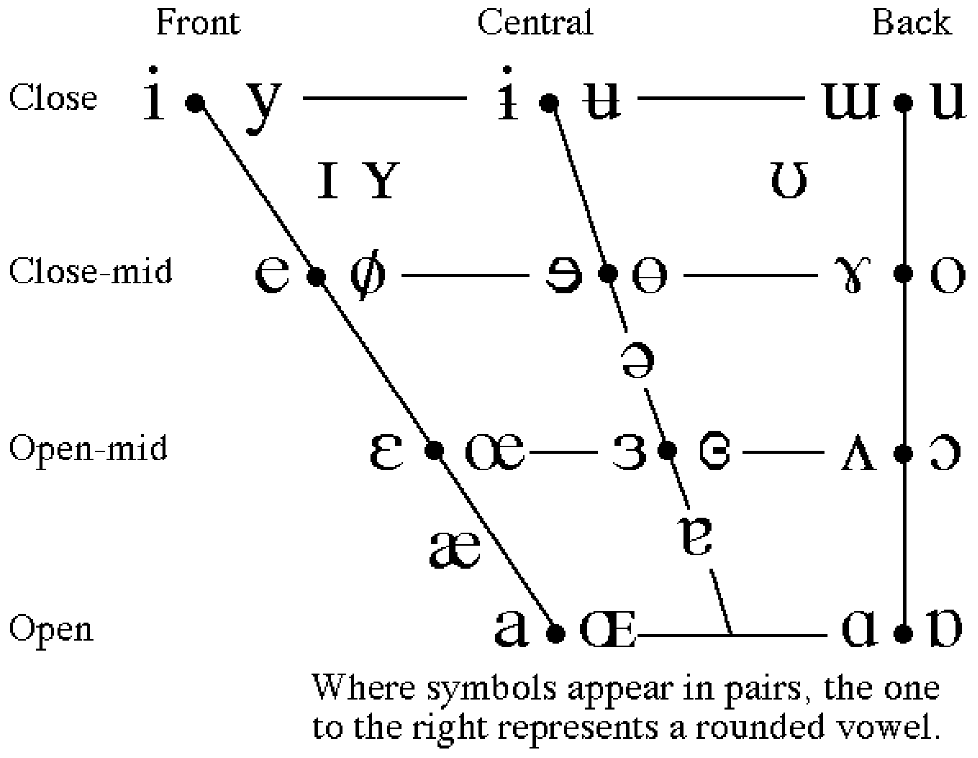
Non-linear Source – Filter Model

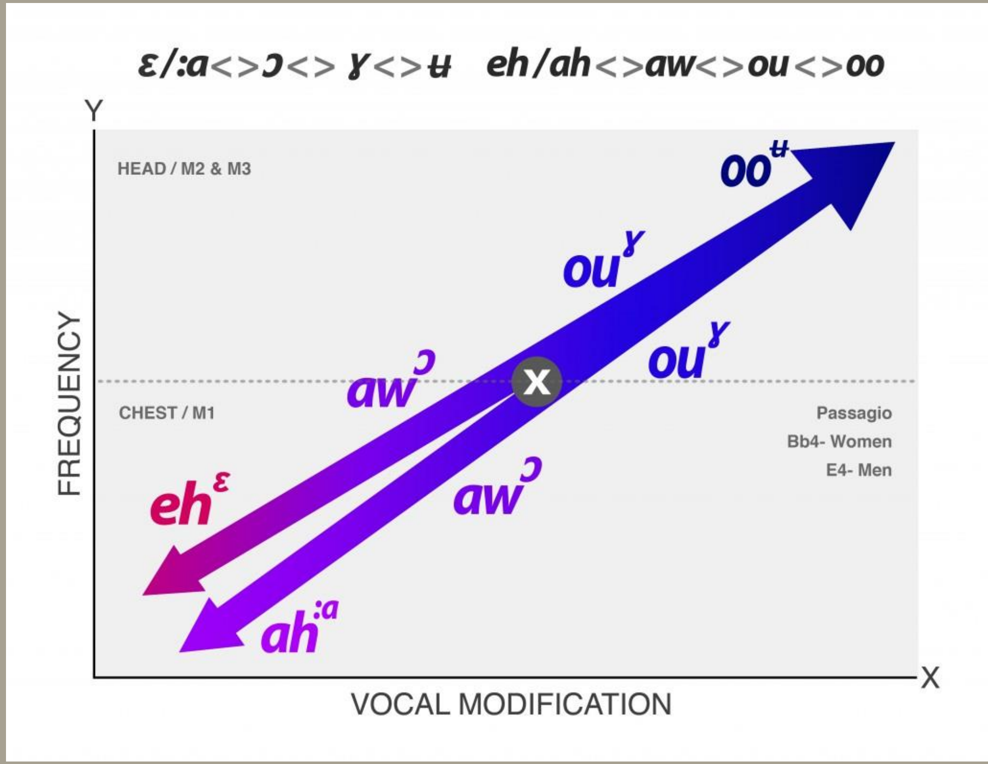


Interactions

Efficient vs. Inefficient Interactions

Acoustics + Psychoacoustics

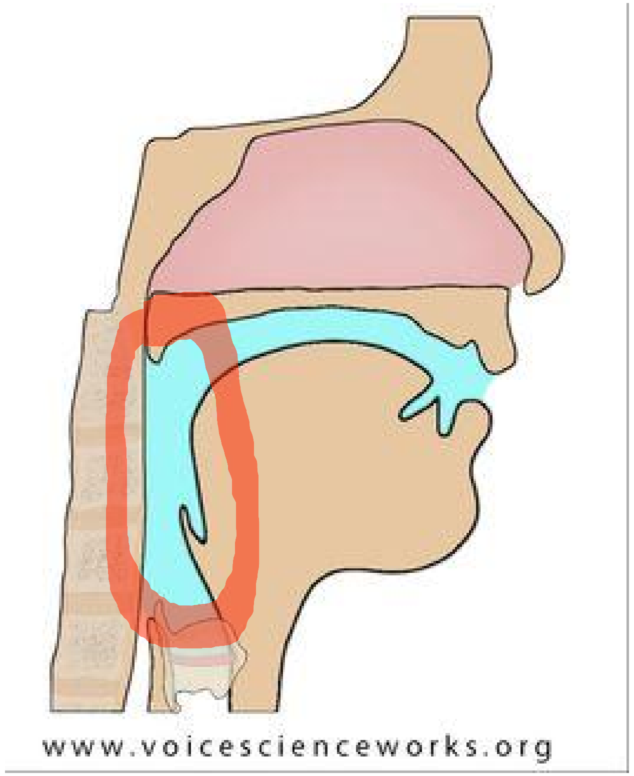




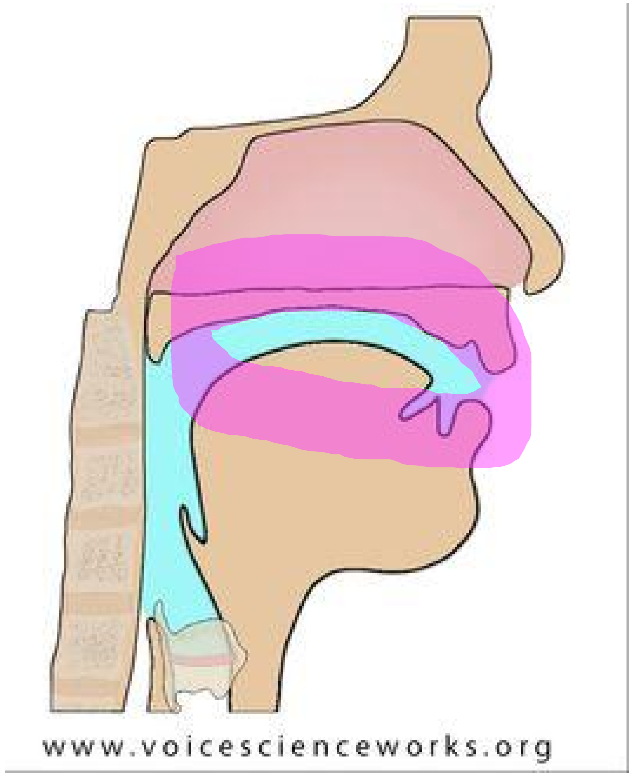
Absolute Spectral Tone Color

Non-Linear Model + ASTC

F1 is responsible for warmth, roundness, depth, efficiency, and for /u o/ primary vowel identity.



F2 is responsible for clarity, degree of brightness, efficiency, and for /ɑ a e i/ primary vowel identity.



What's next?

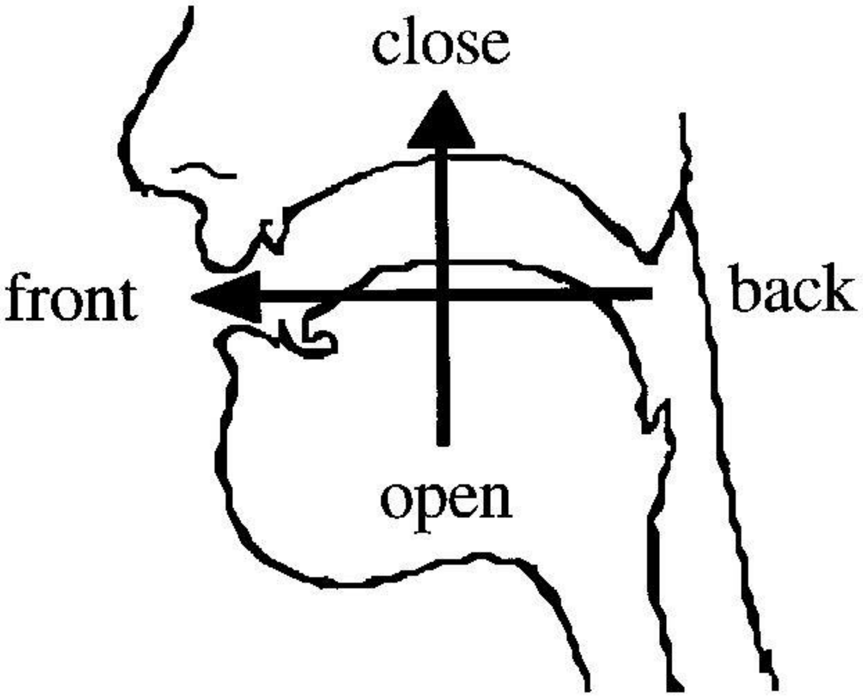
Terminology

Synonyms:

Source Harmonics – Partials – Overtones

Inaccurate or Outmoded:

Hypo- and Hyper-Functional Phonation



Borrowed Terminology

Similarities *and* Differences

Speaking

Singing

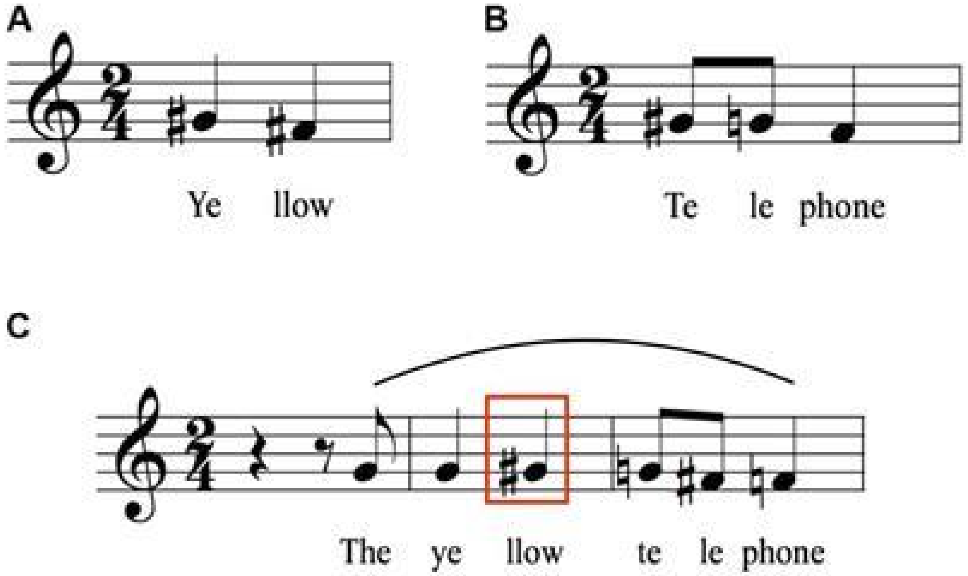
Projection

Speaking

Singing

Intonation

Speaking



Singing

Duration

Speaking

Singing

igne natura renovatur integra