Investigations of the Voce Faringea Using the VoceVista pro System

Alexander Mayr

The *VoceVista pro* system developed by Donald G. Miller, Richard Horne and Gerrie Goeree offers a non-invasive possibility for investigating the glottis function and vocal spectrum. This system consists of the computer software *VoceVista pro* and a small portable electro-glottograph. Two signals are measured for analysis: the audio signal is taken from a headset microphone, and the EGG signal is taken from two electrodes attached to the thyroid cartilage. The system can be used not only to record and analyze audio and EGG signals, but also as a feedback instrument in real time.

With the computer software employed, an audio signal recorded via microphone can be graphically represented in a spectrogram and a "power spectrum". The horizontal dimension of the spectrogram represents a temporal axis - sound samples between 4 and 20 seconds may be thus analyzed. The vertical dimension indicates the basic frequency (F_0, H1) with its attendant overtones, whose frequencies are always multiples of the basic frequency. In the lower left half, Figure 1 shows a spectrogram of one sung tone (B4) with a basic frequency of 500 Hz, and its attendant overtones (1000 Hz, 1500 Hz, 2000 Hz etc.). The volume (loudness) of the individual overtones can be recognized by color: a bright red color (this is in the grey/black depiction of Figure 1 a relatively dark waved line) shows a high decibel level, whereas a dark blue color (in Figure 1 a relatively pale waved line) a low decibel level. Frequency fluctuations due to vocal vibrato can also be clearly observed in the graph.

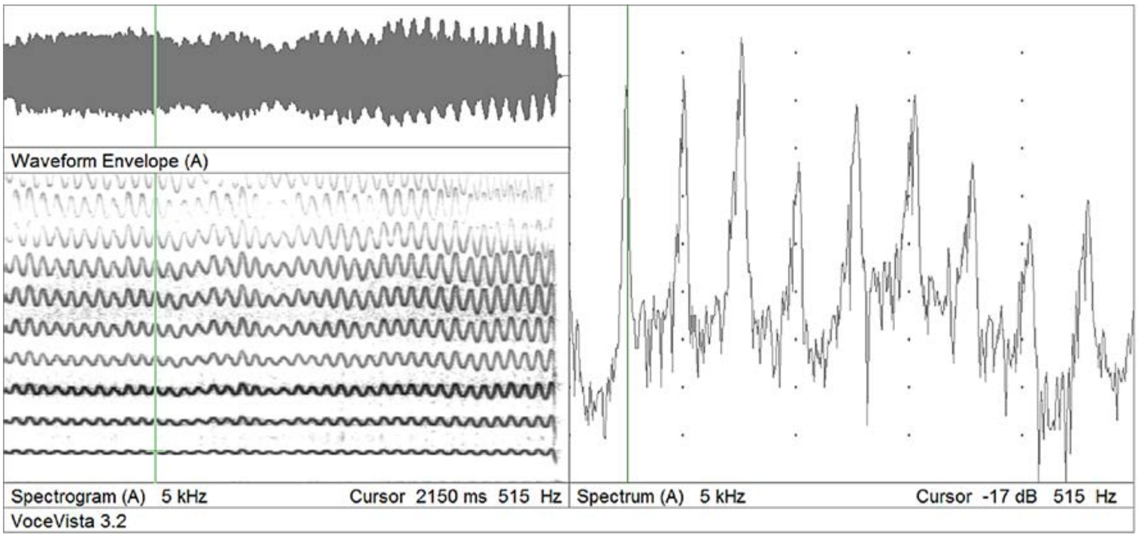
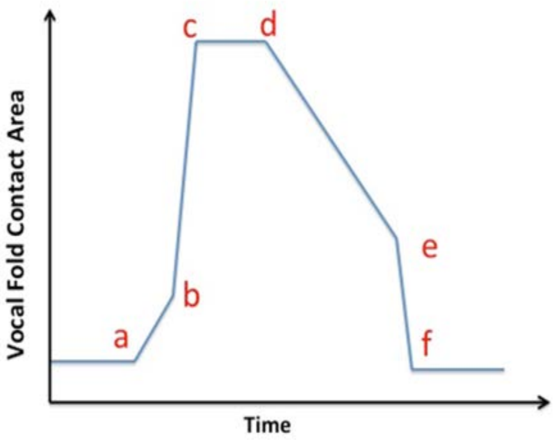


Figure 1: Spectogram and power spectrum of a held tone (B4) in *voce faringea.*

The “power spectrum” represents here a brief moment of the sound sample. The axes shown indicate the frequency in Herz and amplitude in decibels. The power spectrum shows us which frequency components are dominant during this particular moment. When we move the cursor across the time axis of the spectrogram, any moment from the audio recording may also be shown on the power spectrum. In addition, using the VoceVista pro system, a long-term average spectrum (hereafter LTAS) can be constructed for sound samples up to 10 seconds in duration. According to Miller, in order to constitute a more realistic impression of the balance between individual frequency components, a typical vibrato rate of approximately 5 Hz could be normalized through the fixing of the average time between 200 to 300 milliseconds. The second signal to be analyzed by this computer software comes from the electro-glottograph. Between the two electrodes of the EGG, positioned laterally alongside the larynx, flows an electric current of minimal intensity and high frequency. The electrical resistance sinks when the vocal folds are closed, and rises when the glottis opens. Fluctuations in impedance while the vocal folds are vibrating are then represented in a curve form - a form describing the process of glottal closing.

Figure 2: Significant points of the typical EGG curve for a sung tone in modal register. (Kuang, Keating and Patricia, 2013.)



a: incipient contact between the lower levels of the vocal folds

b: incipient contact between the upper levels of the vocal folds

c: beginning of maximal contact between the vocal folds

d: end of maximal contact between the vocal folds; beginning of separation between the lower levels of the vocal folds

e: beginning of separation between the upper levels of the vocal folds

f: end of the opened phase and beginning of the opening phase

Through the EGG curve, certain characteristics of the glottis function can be seen. The contact/opened phases of the vocal folds can be calculated; the form of the curve indicates aspects of the vocal register employed. Figure 3 shows the audio and EGG signal of one sung tone in the vocal function of *voce faringea.* The two green outer indicators each represent the beginning of one period. The middle vertical indicator represents the beginning of the opened phase of the vocal folds.

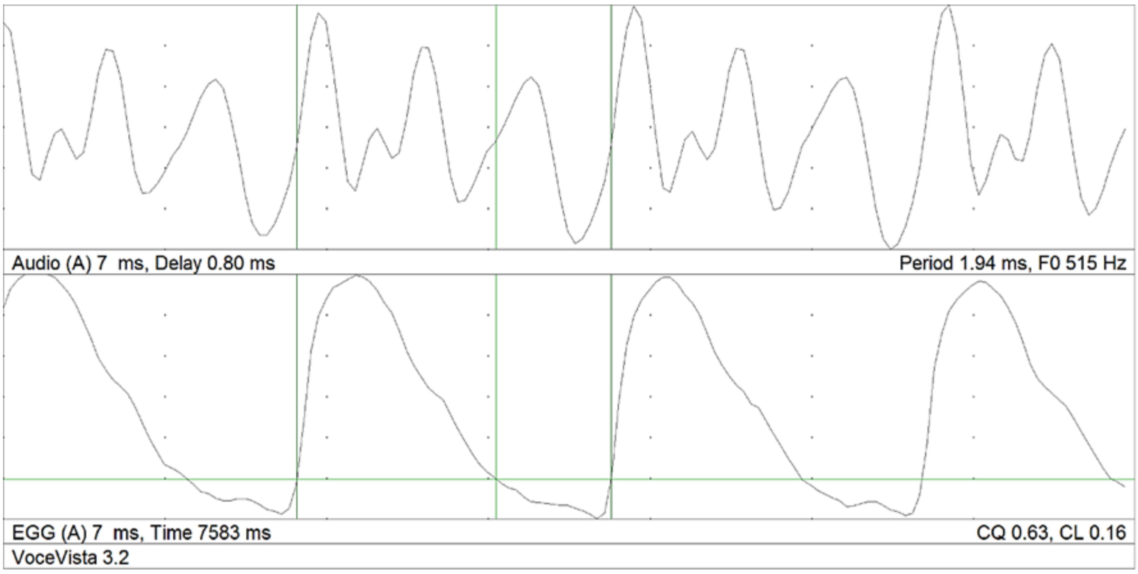


Figure 3: Audio and EGG signal of one tone (B4) sung in *voce faringea.*

However, the EGG alone cannot always deliver valid results in terms of contact quotients: for instance, phases of complete vocal fold closure or of nearly complete glottis closure cannot be distinguished from one another with absolute certainty. One strategy for nevertheless determining the moment of glottis opening is to compare the EGG signal with the audio signal. A dramatic falling off in acoustic intensity is a clear indication of the beginning of the glottis' opened phase. Since the glottis tends to remain closed until the audio signal maximum measures a late, local maximum, we can assume that such an acoustic event constitutes a further indication of the beginning of the opened phase. Employing a late, local maximum, the contact time and the time of the opened phase was determined in Figure 3. The horizontal indicator of the EGG (CL, criterion level) is determined by "manual" comparison of both signals.[1]

The Point of Departure

Muscular contractions in certain regions of the vocal tract, as described for the formation of the falsetto register by Bennati, Colombat, Sieber and other authors in the 18th and early 19th centuries, have a considerable influence on the vocal timbre in the phonation of falsetto tones. Each small alteration of the vocal apparatus results in an alteration of the sound spectrum. "Modeling" the vocal apparatus through partial broadening or narrowing, through opening and closing the access to nasal cavities, through alteration of the tongue's position or lengthening/shortening the entire vocal tract through lifting or lowering the larynx—all these activities accordingly modify the formant structure of the sound. Formants function as a kind of acoustic filter, determining which tones in the overtone series are amplified. Still,

[1] Miller D. G., *Resonance in Singing, Voice Building through Acoustic Feedback,* 2008, pp. 40-44.

historical[2] and contemporary[3] sources confirm that modifications in the form of the vocal tract usually also lead to tensile alterations in the laryngeal apparatus.

During research on the vocal registers of modal, falsetto and *voce faringea* - conducted together with Johan Sundberg at the KTH in Stockholm - I was able to demonstrate clear correlations between the parameters of glottis function and vocal register. Some of the especially relevant parameters for this study, such as vocal fold closing quotient (Q_{closed}), the relationship in intensity between the two lowest partial tones, range and decibel level of formats F1 and F2, as well as the intensity of the singer‘s formant cluster, can also be demonstrated and documented with the *VoceVista pro* system. An advantage of this system is its portability and independence of institutional infrastructure (e.g. sound studios). This system allows me as an investigative artist the chance to demonstrate and analyze - through relatively simple means - specific characteristics of phonation.

Methods and Research Structure

In the final phase of my research project on the reconstruction of the *voce faringea*, I used the *VoceVista pro* system as a device for real-time feedback. To this end, I selected sound samples I had recorded that represented specific characteristcs of the *voce faringea* and in particular its divergences from countertenor falsetto.

These sound samples were recorded using a headset microphone with omnidirectional access and linear frequency. In addition, the contact between my vocal folds during phonation was measured by the EGG. These signals were both transferred to the computer (Microsoft Surface Pro) by means of an audio interface (Tascam US-144 Mk2). During this process, the level of volume for the audio and EGG was set so that distortion and strain could be avoided. In order to attain a better comparison, these signals were then normalized to a level of 85% using the software *Adobe Audition.*

Figure 4: The author of the present work. Experimental setup with the portable VoceVista electro-glottograph and the connected headset microphone.

[2] See Pétrequin and Diday, commentator of the anonymous translator, 1844, p. 291.

[3] Hanayama et al., in their study *Metallic Voice: Physiological and Acoustic Features,* point out the anatomical link between the pharyngeal resonator (*velopharynx*) the aryepiglottic fold (*plica aryepiglottica*), and the *m. arytenoideus obliquus*. Contractions of the *m. palatoglossus* and the *m. palatopharyngeus* would thus continue through the *plica aryepiglottica* and the muscular fibers of the *m. arytenoideus obliquus*, running diagonally along the back of the arytenoid cartilage, and result in an increased adduction of the vocal folds. (Hanayama, Camargo, Tsuji, and Pinho, 2009)

Results of EGG Measurement

Not only my investigation carried out together with Johan Sundberg but also several comparative studies[4] confirm the Q_{closed} value to be an important experimental indicator of modal register and falsetto register. According to Donald Miller, in modal register a closed-quotient of more than 50% can be expected; in falsetto, this value is usually under 40%.[5] Using the recorded sound examples, the results I attained during my research stay in Stockholm should now be confirmed employing the *VoceVista pro* system. Accordingly, we should see that in the *voce faringea*, on the average higher Q_{closed} values should be measured than for countertenor falsetto. An analysis of the EGG curve form should further confirm that vocal fold vibration in the *voce faringea* evidences a clear mucosal wave that is typically very weakly defined or absent in falsetto.

Figure 5 shows the spectrograms (left) as well as the audio and EGG signals (right) for a sung tone of the fundamental frequency 504 Hz (H4) in voce faringea (above) and a falsetto tone of 501 Hz (below). Both sung tones were sung on the vowel “a”. The point in time when the glottis closing began can be clearly measured for both cases. The beginning of the opened phase is falsetto can also be clearly measured. It is, however, considerably more difficult to detect the contact time of the vocal folds in voce faringea, since the EGG curve for the area representing the opened phase assumes a very flat form.

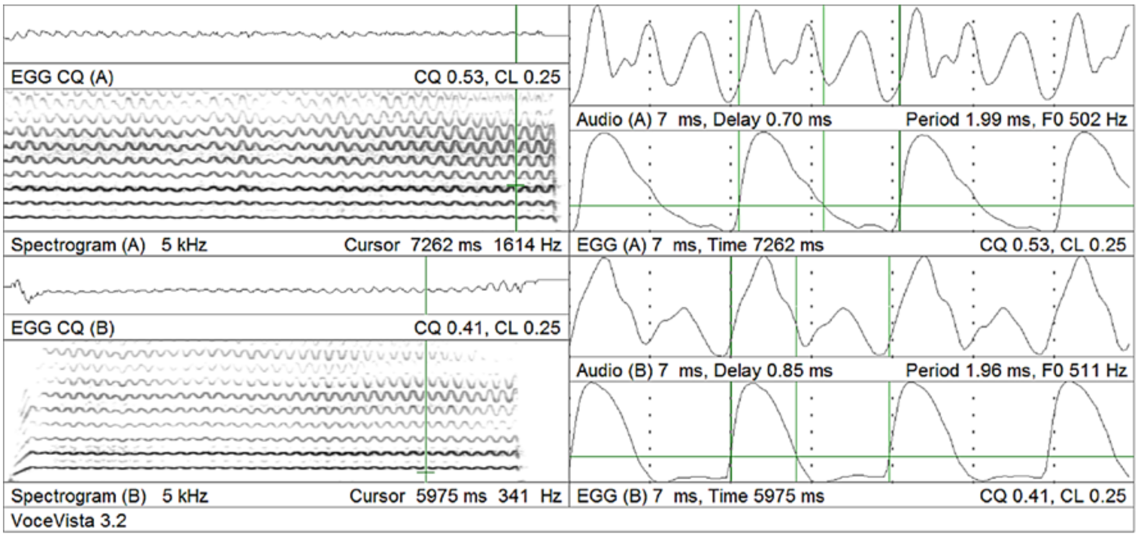


Figure 5: Spectrograms (left). Audio signal and EGG signal (right) of the *voce faringea* (above) and falsetto (below). Audio examples can be heard at www.voce-faringea.com.

For comparison between the two signals, a criterion level of .25 was chosen. The values of the contact quotients fluctuate with vocal vibrato between 58% and 48% in *voce faringea* and between 46% and 36% in falsetto. The Qclosed values in Figure 5 - 53% for *voce faringea* and 41% in falsetto - correspond to values measured in the two sound samples. This means that the glottis in *voce faringea* is, on the average,

[4] Högset and Sundberg, 2001, Salomao and Sundberg, 2008.

[5] Miller D.G., *Resonance in Singing, Voice Building through Acoustic Feedback*, 2008.

12% longer closed than for the corresponding period in falsetto. Particularly conspicuous is the difference in curve shape for the two EGG signals: the wave line of the EGG exhibits a clear knee shape for the *voce faringea*. This shape indicates a vocal fold vibration with a phase displacement between the body of the vocal folds and their cover. This phase displacement causes a prolongation of the contact period and an enlargement of the vocal folds' vertical contact area. The mucosal wave that occurs in *voce faringea* is extremely atypical for a type of phonation based in the falsetto register. Through the greater longitudinal tension of the vocal folds during moderate to no vocal activity, the vocal folds vibrate in falsetto with a reduced substance and without any phase displacement. For this reason, the EGG wave form of the falsetto note exhibits no knee shape.

The Q_{closed} values were calculated for the study *Flow-Glottogram and EGG Parameters in the Vocal Registers: Modal, Falsetto and Voce Faringea*, using data from an EGG measurement and an inverse filtering of the audio signal. In this method, accuracy of measuring contact quotients is presumably somewhat higher than with the *VoceVista pro* system. Nevertheless, as regards the relation between falsetto and *voce faringea* values, we can consider our results as valid. The average Q_{closed} value for *voce faringea* was approximately 3 points higher than for falsetto in the Stockholm study. This results in a difference of contact time of about 9%. The Qclosed difference for the 2014 *VoceVista pro* study is about 22%. In the *VoceVista* EGGs, the knee shape in the curve form also has a more characteristic appearance than in the EGG of the 2012 study. Differences in vocal timbre were also clearer to observe. Whereas the vocal timbre with a broadened pharynx exhibited a mezzo-like, soft timbre, the timbre acquired more sharpness and a metallic, penetrating character with contracted faucial head. Comparison of data from 2012 and 2014 supports my personal conviction that I have in fact managed to considerably refine my technique towards developing *voce faringea* during the last two years of my artistic research process.

Divergences in the Power Spectrum

The power spectrum shows the moment-to-moment relative power of the overtone series for one tone. Figure 6 shows the simulation of a primary sound spectrum, without alterations due to resonance in the vocal tract. We can clearly see the decreasing decibel strength (6 to 12 dB per octave) through the partial tones. Naturally, the low overtones are stronger than the higher ones, which also means the first two formants F1 and F2 are acoustically more strongly featured. When this sound spectrum is filtered through the vocal tract, displacements occur according to the strength of individual frequency areas.

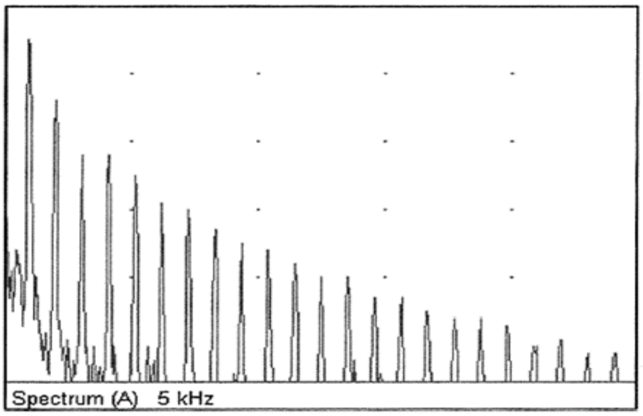


Figure 6: Simulated primary sound spectrum of one tone (A3). (Miller, D.G., *Resonance in Singing, Voice Building Through Acoustic Feedback*, 2008, p. 25)

Vocal tract shape gives preference to the production of sounds within certain frequency areas - the areas of the formants. The two lowest formants F1 and F2 are determined by the vowel being sung, the formants F3, F4 and F5 are related to the vocal timbre, constituting the singer's formant cluster for classically trained opera singers.

The form of the power spectrum thus shows the resonance influences of the vocal tract and, accordingly, the formant structure. Figure 7 shows two phonetic categories related to F1 and F2: vowels can be closed or open in articulation. Closed vowels have a relatively low first formant, whereas open vowels have a high first formant. The second category is related to the form of the mouth cavity and tongue position.

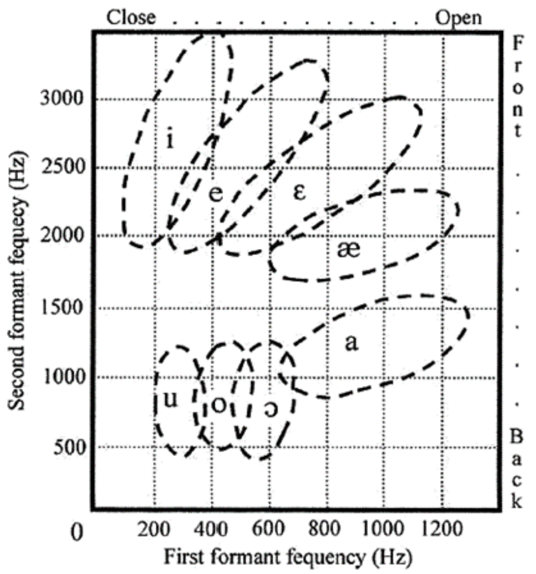


Figure 7: Illustration of the formant regions of the vowels. The axes show respectively the frequency ranges of the first (horizontal) and the second formant (vertical). (DG Miller, Resonance in Singing, Voice Building through Acoustic Feedback, 2008, p 27)

With the articulation of the vowel "*I*", the tongue forms a curve relatively far forward in the mouth, with its midsection close to the hard palate. In order to articulate the vowel "*A*", on the other hand, the tongue curves further back in the mouth. Through subtle alteration and adaptation of the open/closed and anterior/posterior dimensions, it is possible to alter the formant frequencies of the vowels so that partial tones will be added or overlapped and acoustic strength within these areas of frequency significantly strengthened.[6]

By means of the power spectrum, differences in the vocal tract's shape for *voce faringea* and falsetto as well as diverse strategies for increasing sound in both modes of phonation can be demonstrated, and the corresponding singer's formant clusters can be examined for *voce faringea* and falsetto using long-term average spectra. Finally, the power spectrum for *voce faringea* when it is predominantly falsetto can be compared with what Donald Miller has characterized as *mezza voce (voix mixte*,

[6] Miller, D.G., *Resonance in Singing, Voice Building Through Acoustic Feedback*, 2008, pp. 24-28.

according to the current definition of that term), which is predominantly modal in vocal function for male opera singers.

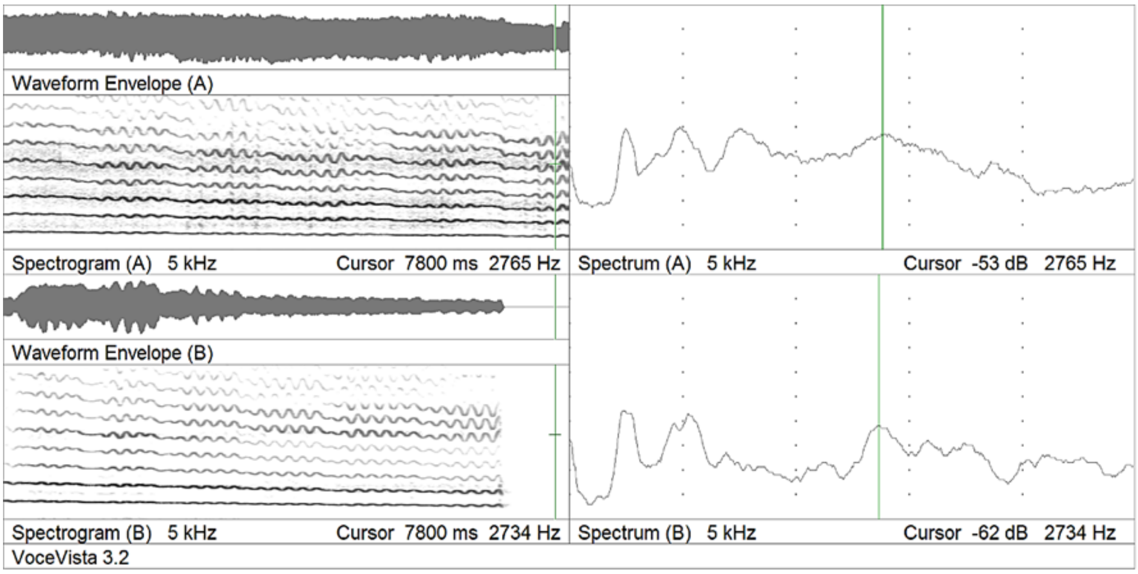


Figure 8: Long-term average spectrum (LTAS) for a series of tones in *voce faringea* (above) and falsetto (below). Audio examples can be heard on www.voce-faringea.com

For comparing the strength of the singer's formant clusters, a chromatic scale from D5 to A4 on the vowel A was sung first in *voce faringea* and then in falsetto. The power spectrum average was set to 8000 ms. The LTAS for the *voce faringea* (Figure 8, above) exhibited a maximum of -53 dB of the singer's formant cluster at approximately 2765 Hz. In the LTAS for the falsetto (Figure 8, below), the highest decibel value was also in the area of formant F3 - yet this was 9 dB less than in the *voce faringea*. This is presumably due to varying resonance strategies between both register mechanisms.

In falsetto - particularly in the highest range - the overlapping of base frequencies and the first formant F1 lead to a considerable amplification of the sound. This is due on the one hand to the comparatively wide amplitude of H1 in LTAS, but also to the relative strength of the fundamental frequency in the first section of the spectrogram (Figure 8, lower left) compared with the higher overtones. In the *voce faringea*, on the other hand, the singer strives to overlap the second formant with the third partial tone in order to amplify the sound. Altered positions in the larynx and in the vocal tract[7] (compared, that is, with falsetto) also contribute to strengthening of the singer's formant cluster.

The setting of formant frequencies to individual partial tones occurs through shifts on both axes, which represent the open/closed and anterior/posterior dimensions of the vowels.

[7] These altered positions as falsetto is approached are described in more detail in Chapter 3.1.2 of the author's dissertation as regards phonation and resonance systems in *voce faringea.*

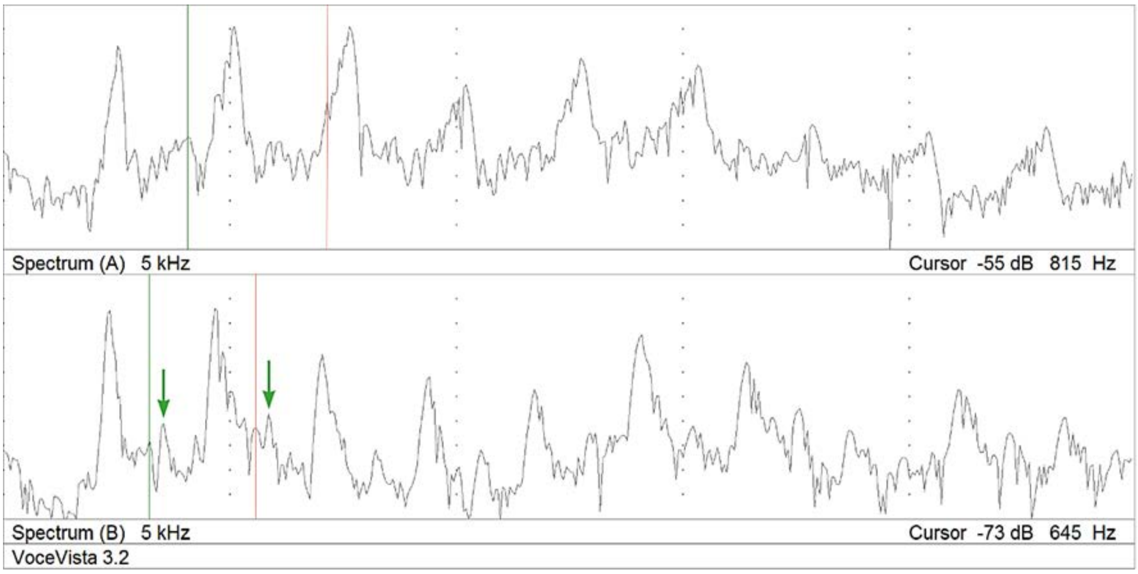


Figure 9: Excerpt from the LTAS (tone pitch: B4) for *voce faringea* and falsetto in a power spectrum. The green indicators show the first formant F1, the red vertical lines mark the second formant F2. The two small green arrows in the lower power spectrum indicate subharmonic partials. Such subharmonies often fall into the area between two partial tones and must not be confused with formants.

Figure 9 shows an excerpt from the two LTAS for *voce faringea* and falsetto in a power spectrum. The green vertical markers indicate the frequency of the first formant F1, and the red markers show the position of the second formant in the spectrum.

Here it can be clearly observed that in *voce faringea*, the third partial tone is strengthened through tuning of F2 to the frequency area around 1440 Hz. This resonance strategy of modifying the vowel to achieve an overlapping of the second formant with H3 or H4 is often typical for tenor voices in modal register.[8] The second formant in falsetto, on the other hand, is approximately 200 Hz above the fundamental frequency for the area between H2 and H3 – F1. The decibel level of H3 is 20 dB higher in *voce faringea* than in falsetto (-17 to -37 dB). The vocal color of the A vowel seems darker or more closed, and the tongue curve occurs further back in the mouth than for *voce faringea*. Accordingly, when a tenor sings the vowel A in falsetto we hear a tendency towards the vowel O, whereas when he sings A in *voce faringea* we hear a tendency towards the vowel AE (Ä).

These data clearly demonstrate that the differences between falsetto and *voce faringea* in vocal color and timbre correlate with specific positions in the vocal tract. In countertenor falsetto, the pharynx and nasal cavities are widely broadened and the larynx retains a low position. Analysis of the sound sample for falsetto register confirms that here, the resonance strategy employed is also used by female singers when singing in their higher ranges (upper middle register and head register).

[8] In one recent study, Sundberg, La and Gill were unable to demonstrate formant-tuning for classically trained opera singers. (Sundberg, La and Gill, 2013)

Accordingly, the timbre exhibits a round, soft and prevailing feminine character. For phonation in *voce faringea*, on the other hand, the vocal tract is somewhat shortened by raising the larynx minimally higher than in modal *or* falsetto register. In particular, modifications in the vocal tract close to the *isthmus faucium* can be linked to the tuning of the second formant and a pervasively metallic character in the sound.

In comparison to the predominantly falsetto *voce faringea,* the predominantly modal *mezza voce* evidences somewhat lesser average Q_{closed} values (see Figure 10). During sound production in *mezza voce*, I was able to observe a relatively strong contractive tension in the posterior area of the vocal folds, whereas in *voce faringea* the medial tension in the anterior area of the glottis felt considerably less during *voce faringea* phonation. Results of EGG measurement in *mezza voce* correspond to those in Donald Miller's study, *Soft Phonation in the Male Singing Voice.* Miller investigated characteristics of *mezza voce* in two tenors (one lyric, one dramatic), and found in one of the recordings he made that there were high contact quotient values even though a stroboscopic examination showed that the vocal folds had not completely closed during this light form of phonation.

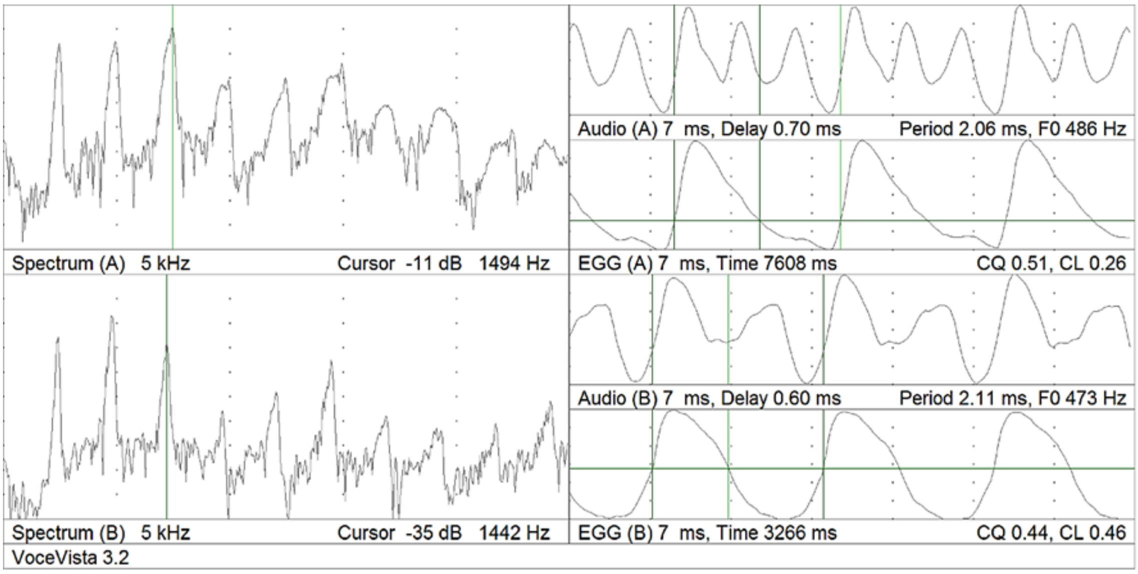


Figure 10: Power spectrum and EGG of one sung tone (B-flat 4) in *voce faringea* (above) and in *mezza voce* (below). Audio recordings can be heard at www.voce-faringea.com.

In *mezza voce,* broadening of the entire vocal tract and stretching of the epilaryngeal collar lead to a significant reduction in the otherwise dominant second formant F2. Figure 10 shows a decibel level for H3 that is 24 dB higher in *voce faringea* than in *mezza voce.* The reason for the weakness of the third partial is that the second formant is no longer in the area of influence of H3, and thus can no longer amplify it. In addition, the power spectrum for *voce faringea* also demonstrates considerably stronger partials in the area of the singer's formant cluster. During the course of experimentation with diverse register mechanisms, it became evident that the shift from modally dominated *mezza voce* to falsetto-dominated *voce faringea* can be accomplished relatively simply. In this sense, *mezza voce* could be applied as a middle level in shifting from *voce faringea* to modal register and vice versa.

References

Hanayama, E. M., Camargo, Z. A., Tsuji, D. H., & Pinho, S. M. (2009). Metallic Voice: Physiological and Acoustic Features. In: *Journal of Voice, Vol. 23, No. 1*, p 62 - 70.

Högset, C., & Sundberg, J. (2001). Voice source differences between falsetto and modal registers in counter tenors, tenors and baritones. In: *Logopedics Phoniatrics Vocology Nr. 26*, p 26-36.

Kuang, J., Keating, & Patricia. (2013). Glottal articulations in tense vs. lax phonation contrasts. In: *The Journal of the Acoustical Society of America, Vol. 134, Issue 5*, p 4069.

Mayr A. (2014). Die *voce faringea*, Rekonstruktion einer vergessenen Kunst. doctoral dissertation at the University for Music and Performing Arts Graz, 2014

Miller, D. G. (2008). *Resonance in Singing, Voice Building through Acoustic Feedback.* Princeton: Inside View Press.

Pétrequin, J. P., & Diday, P. (1844). Ueber den Mechanismus der Fistelstimme (Falsettstimme). Von Herrn Pétrequin, Oberarzt am Hotel Dieu zu Lyon. und Herrn Diday. In: L.-F. v. Froriep, *Notizen aus dem Gebiete der Natur- und Heilkunde, gesammelt u. mitgetheilt von Ludwig Friedrich v. Froriep: Nr. 19 of Vol. 29*. Berlin: Lossius. p 289 - 286, 321 - 328, 337 – 344.

Sundberg, J., Lã F. M. B. & Gill, Brian P. (2013). Formant Tuning Strategies in Professional Male Opera Singers. In: *Journal of Voice*, *Vol. 27, No. 3*. p 278-288.