Formant Tuning and Feedback in the Male Passaggio[1]

Wim Ritzerfeld and Donald Gray Miller

Abstract

It has been suggested that traversing the male (secondo) passaggio requires two important adjustments. When singing up the scale, first of all, the second harmonic (H2) needs to pass over the first formant (F1). After that, the timbre of the voice takes on a different, slightly 'darker' quality. This is the pitch where, in singer's jargon, the voice 'turns over.' Above the passaggio, in the optimal arrangement the second formant (F2) is then tuned to one of the higher harmonics, or, sometimes alternatively, the singer's formant cluster induces a dominant resonance in the approximate range 2.4 to 3.4 kHz. These two adjustments together produce the typical sound of the classical male upper voice. We have investigated the mechanisms underlying these adjustments, using the VoceVista software, to record EGG and microphone signals of the singers, and inverse filtering, using the DeCap software, to reveal the shape of the glottal airflow pulses. The results indicate that:

1. Tuning of the first formant to the second harmonic affects the shape of the glottal airflow pulses through non-linear feedback. Inverse filtering reveals that if H2 is below F1, the pulses are skewed to the right; if H2 is above F1, the pulses are more symmetrical. This could explain the change in timbre as the voice turns over. We will show that the difference in pulse shape is caused by the changing phase relation between the supraglottic pressure wave and the glottal airflow pulse as H2 passes F1.
2. The tuning of the second formant appears to have a negligible effect on the shape of the airflow pulses. Its most salient effect on the sound is the enhancement of one of the higher harmonics. The more precisely F2 is tuned to one of these harmonics, the higher its amplitude.
3. Successful second formant tuning not only depends on precise tuning, but also on the closed quotient (CQ). Certain values of the CQ turn out to be more favorable for F2 tuning than others. We have investigated the nature of this dependency and offer here an explanation of the phenomenon: F2 tuning turns out to be most effective if the glottis stays closed into the final cycle of the dominant enhanced harmonic.

Introduction

One of the challenges of the aspiring male classical singer is the successful negotiation of the *passaggio*. The passaggio is the range of pitches where the frequency of the first formant (F1) is near the second harmonic (H2) of the sung pitch for open vowels such as /E/, /O/ and /A/. When this happens, there is a strong tendency for the singer to keep F1 close to or higher than H2. This is often accomplished by raising the larynx and/or constricting the pharynx. The resulting sound resembles what is commonly known as a *yell* and it is considered undesirable in classical singing. During their training, singers must therefore learn to let H2 rise above F1 for the higher pitches. The sound then takes on a slightly darker quality. Above the passaggio, ideally, the second formant (F2) is tuned to one of the higher harmonics. Sometimes, alternatively, the singer's formant cluster induces a dominant resonance in the approximate range 2.4 to 3.4 kHz. [Reference]

In this study we have investigated two phenomena that play an important role in relation to the male passaggio:

1. Feedback of supraglottic pressure to the voice source, more specifically, the changes that occur in the shape of the airflow pulses through the glottis as the singer traverses the passaggio.
2. The effect of the Closed Quotient (CQ), especially on second formant tuning.

To this end, we have recorded microphone and EGG signals of several subjects, singing rising scales on the vowel /a/, using VoceVista. We then performed inverse filtering on the microphone signal, using the *Sopran/DeCap* software, designed by Svante Granqvist, to reveal the shape of the glottal airflow pulses. The resulting airflow waveforms were then displayed in VoceVista, together with the EGG signal.

Feedback of F1 resonance to the voice source

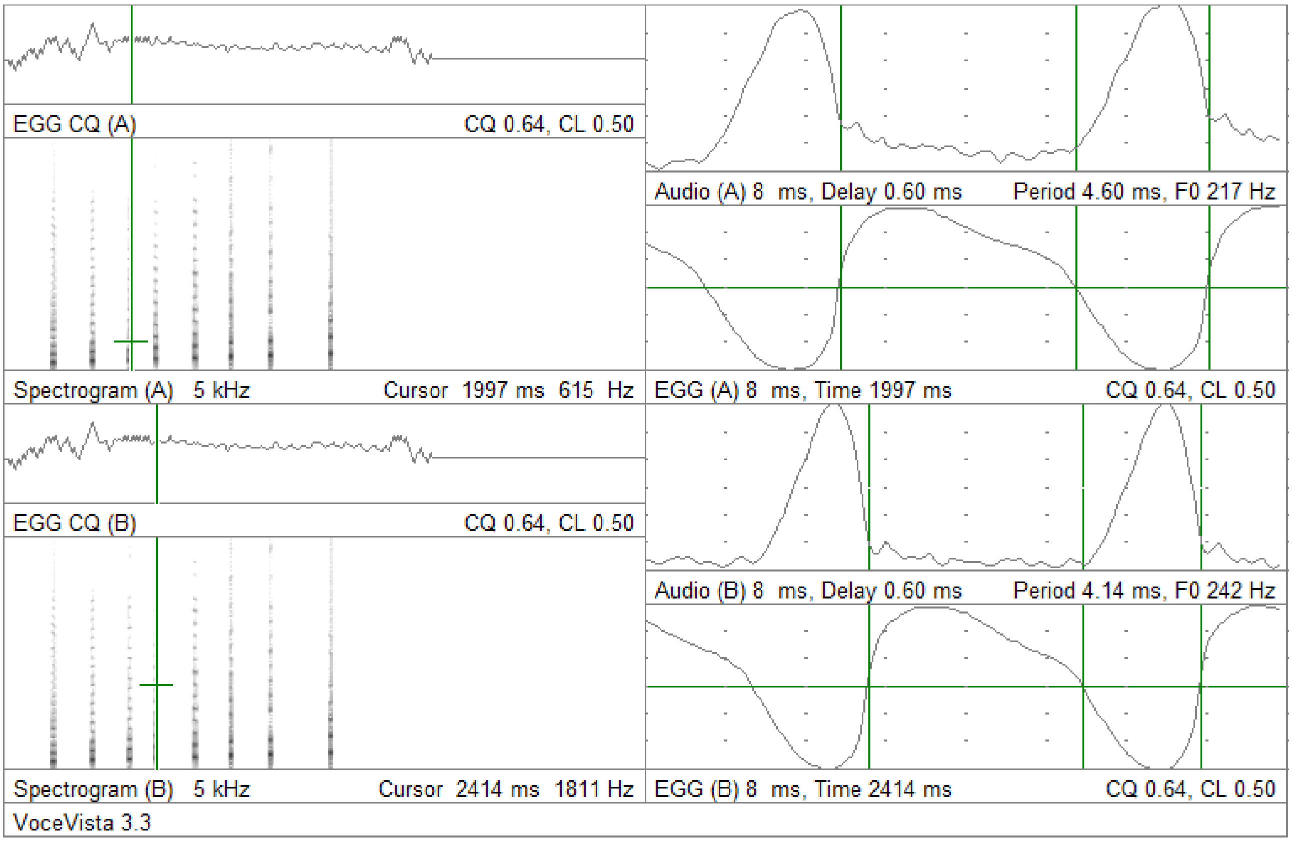

Figure 1: Scale F3-F4 [a] inverse filtered glottal airflow at A3 (upper) and B3b

In order to investigate the effect of feedback of vocal tract resonance to the voice source we have recorded scales from F3 to F4 or from F3# to F4# on two open vowels ([a] and [E]) from several singers, including the authors. Inverse filtering was applied on all pitches, using the Sopran/DeCap software and the results were imported into VoceVista for display and comparison.

Figures 1 thru 3 show the results of inverse filtering a scale from F3-F4 on [a] by one of the authors (WR). Figure 1 displays Glottal Airflow at pitches A3 (upper panel) and B3b (lower panel). Note that both pulses are skewed to the right (rise time is greater than fall time).Also note that airflow starts approximately at the knee in the EGG signal.

The Criterion Level (CL) of VoceVista (the horizontal cursor in the EGG panel) was set to 50%. As can be seen, with this setting, the open phase corresponds to the interval where the airflow signal is approximately flat (zero) and the open phase corresponds to the interval during which air is flowing through the glottis. The CQs obtained in this way are on the low side compared to the ones obtained through other methods. In order to avoid confusion we will call this particular CQ the "Airflow

CQ".

Now observe Figure 2, which displays pitches C4 (upper) and D4 (lower). At pitch D4 the glottal pulse is no longer skewed. It has become symmetrical and it exhibits a dent in the middle.

Figure 3 displays the two highest pitches (E4 and F4). The pulses are nearly symmetrical at both pitches.

Discussion

Apparently the glottal pulses are skewed below D4 and symmetrical above D4. D4 happens to be the pitch where the first formant (F1) of the [a] vowel approximately aligns with the second harmonic (H2).

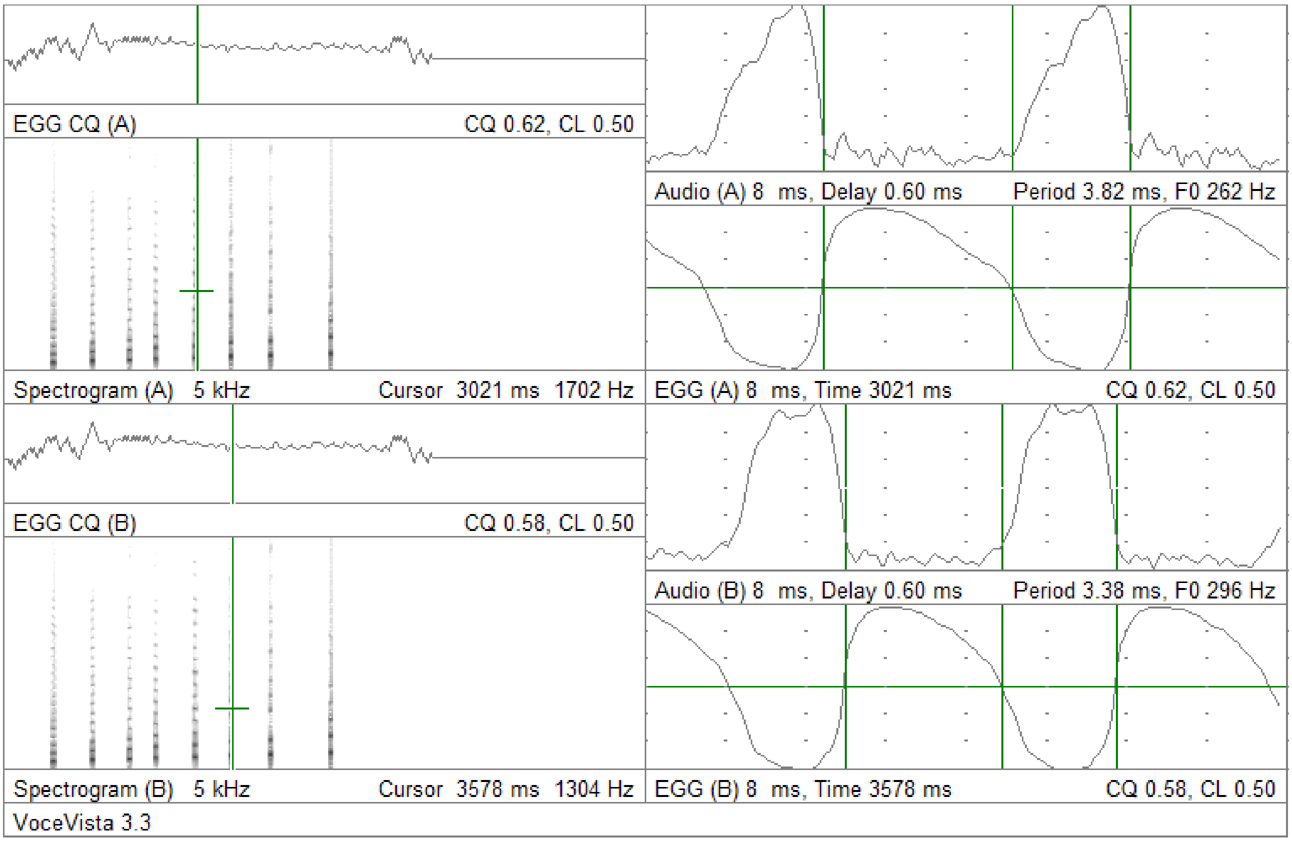

Figure 2: Scale F3-F4 [a] inverse filtered glottal airflow at C4 (upper) and D4

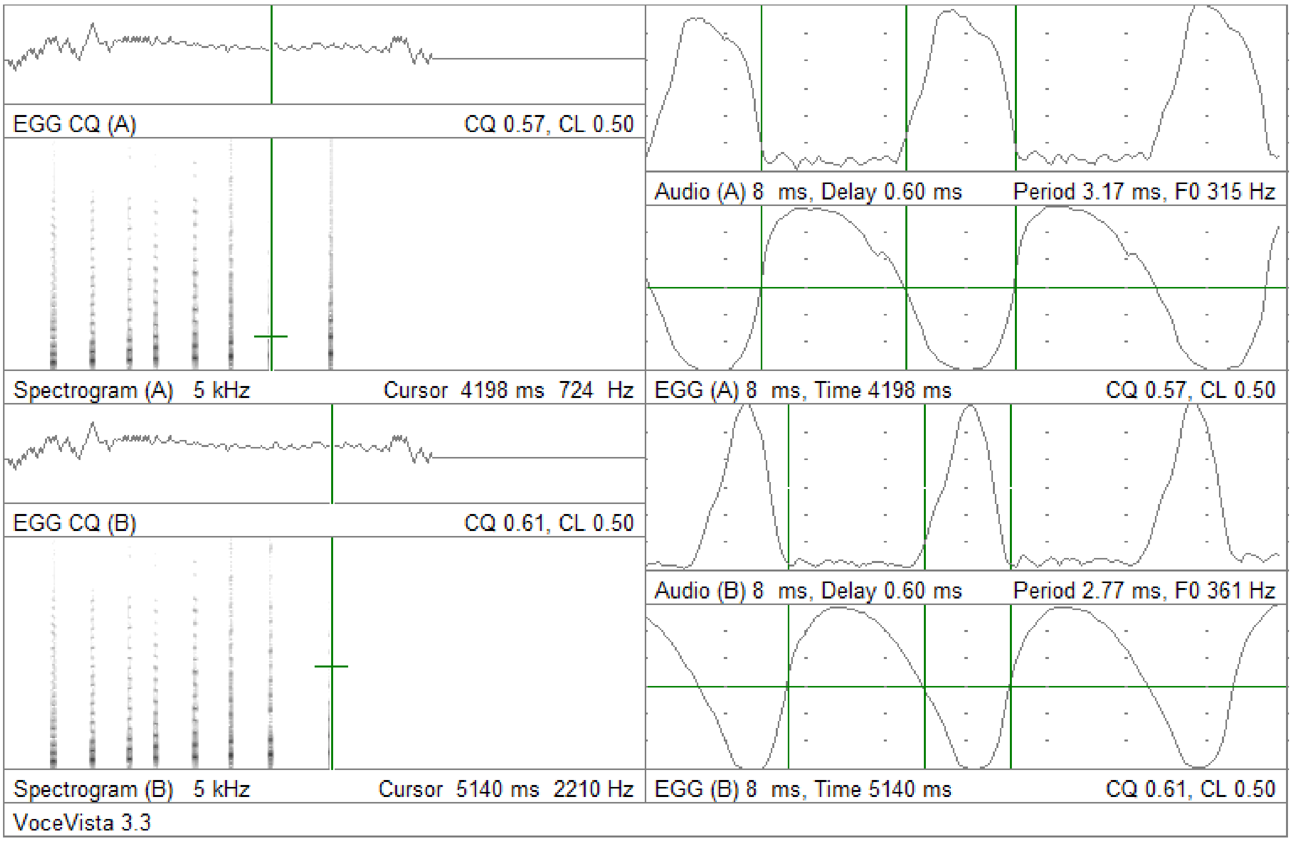

Figure 3: Scale F3-F4 [a] inverse filtered glottal airflow at E4 (upper) and F4

Intermezzo: timing of supraglottic pressure wave

In order to explain why the change from skewed to symmetrical pulses takes place exacly at the pitch where F1 aligns with H2, it is important to understand the time relation between the variable pressure above the glottis (supraglottic pressure, P_{supra}) and the source signal, which is essentially the airflow signal through the glottis (U_g). This timing relation can be derived theoretically by calculating *vocal tract input impedance* (Z_{in}) and, specifically, the *phase* of Z_{in} . We choose not to do this in the context of this article and, instead, show an example which explains the timing relation between P_{supra} and U_g in a practical way.

In 1984 Schutte and Miller [reference] have carried out an investigation in which P_{supra} and also P_{sub} (subglottic pressure) were measured directly by two pressure transducers, mounted on a catheter, which was inserted through the nose into the singer's pharynx. One pressure transducer was located below the glottis and the other above the glottis. One of the experiments involved having the singer vocalize a vowel series on [i] – [I] - [e] – [E] – [a] – [O] – [o] - [u] on pitch B3. Note that in this vowel series the first formant frequency (F1) gradually rises from a low value at [i] to a high value at [a] and gradually drops back to a low value at [u].

Figure 4 (lower panel) shows the Psupra and EGG signals for the vowel [E], where F1 > H2. Note that the second maximum of P_{supra} appears *early* in the open phase. In Figure 5 (vowel [e]) the first formant frequency (F1) is lower than the second harmonic (F1 < H2). Now the second maximum of Psupra appears *late* in the open phase. This behavior is consistent with the theoretical behavior of vocal tract input impedance and it holds for any formant that is close to a harmonic. It can be summarized as follows:

- If F1 > H2, the maxima and minima of H2 appear sooner ("are shifted to the left"). For the mathematicians, the second harmonic of P_{supra} receives a positive phase shift with respect to the second harmonic of Glottal Airflow.

- If F1 < H2, the maxima and minima of H2 appear later ("are shifted to the right"). For the mathematicians, the second harmonic of Psupra receives a negative phase shift with respect to the second harmonic of Glottal Airflow.

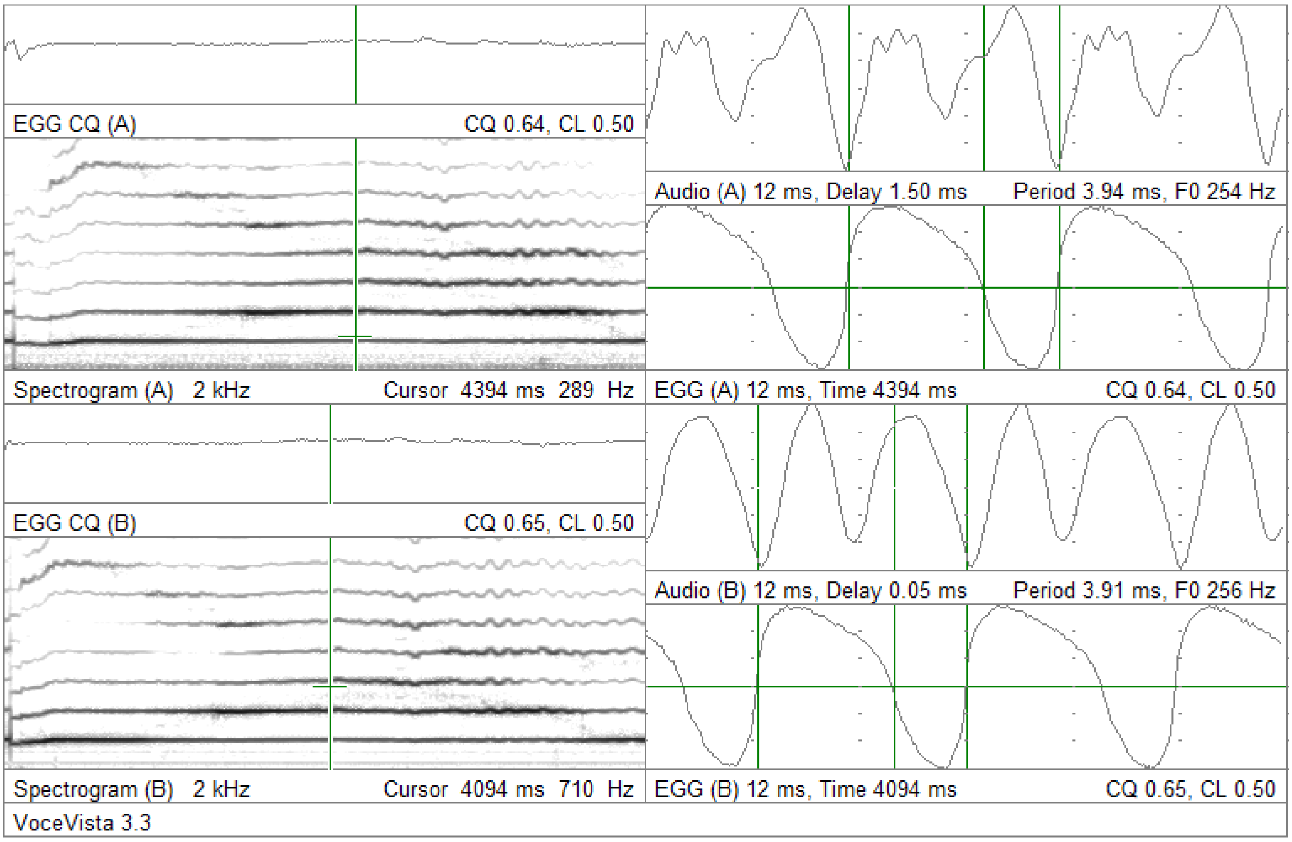

Figure 4: Supraglottic pressure (lower panel) and integrated Audio (upper panel) (F1 above H2)

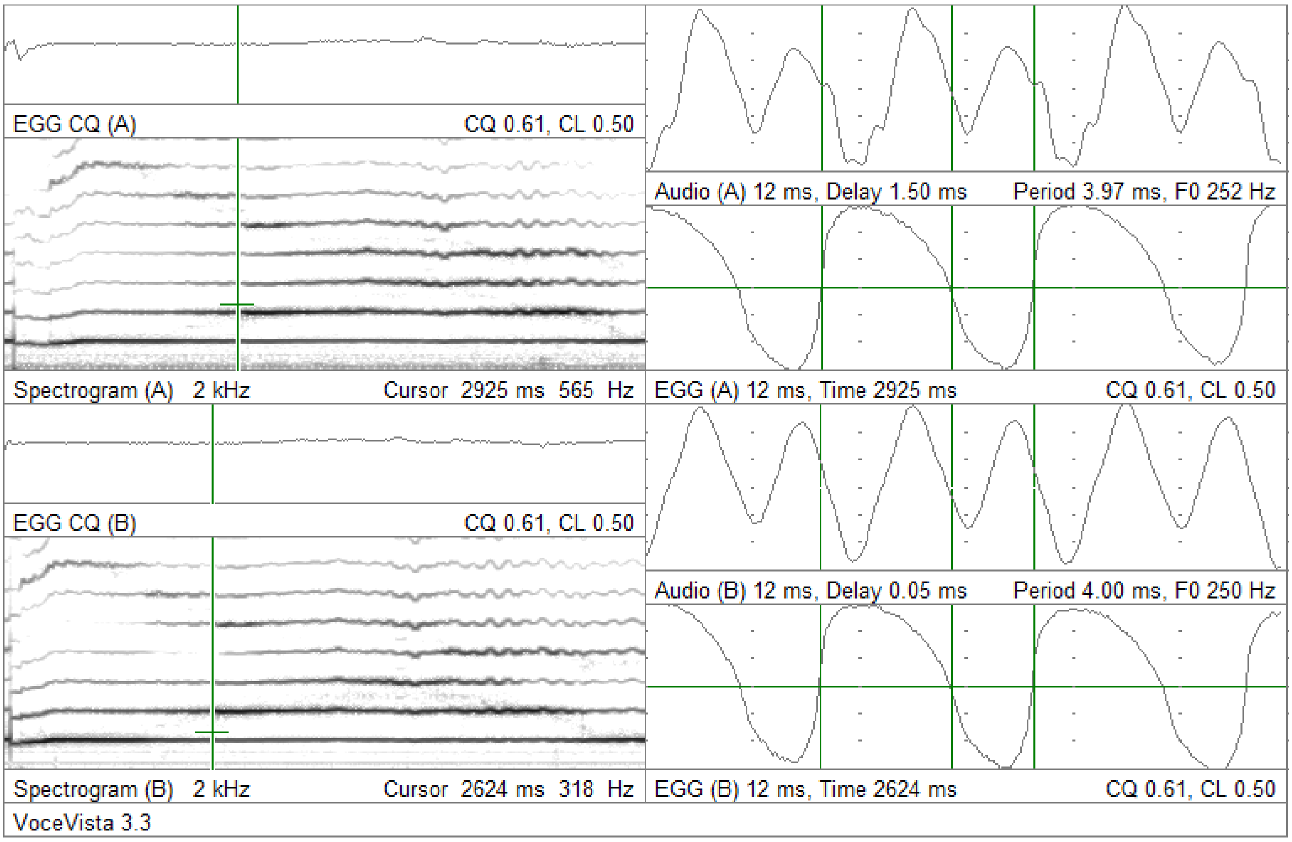

Figure 5: Supraglottic pressure (lower panel) and integrated Audio (upper panel) (F1 below H2)

In our present investigation we have not measured P_{supra} directly. The only signals recorded were microphone (Audio) and EGG signals. The Audio signal does not correctly represent P_{supra}. The reason for this is that the sound waves that travel from the singer's mouth to the microphone are *spherical* waves, while the waves inside the singer's vocal tract are *planar* waves. It has been shown [reference] that, as a result of this, the pressure wave that reaches the microphone (P_{audio}) is proportional to the *time derivative* of the airflow wave (U_o) through the singer's mouth. In other words: the microphone signal can be found by differentiating the airflow signal through the mouth. In formula:

$$P_{audio} = K * dU_o/dt$$

Where P_{audio} is the sound pressure, measured by the microphone, U_o is oral airflow (the airflow through the mouth) and K is a constant.

We therefore hypothesize that a better representation of P_{supra} (without directly measuring it) can be found by performing integration (the reverse of differentiation) on the microphone signal. Doing this for the experiment outlined above yields the signals in the upper panels of Figures 4 and 5. as can be seen, they are quite similar to the P_{supra} signals in the lower panels. Whether the integrated audio signal can generally be used as a substitute for the P_{supra} signal remains to be investigated, but in the case where the first formant is close to the second harmonic, the correspondence seems quite good.

Explanation of glottal pulse skewing for H2 < F1

Applying the hypothesis that the Integrated Audio signal can be used as a substitute for the P_{supra} signal to the F3-F4 scale on [a] can bring us closer to an explanation of the pulse skewing phenomenon described above. In Figure 6 the integrated Audio signal (upper panel) and the inverse filtered glottal airflow signal (lower panel) are plotted together for pitch C4.

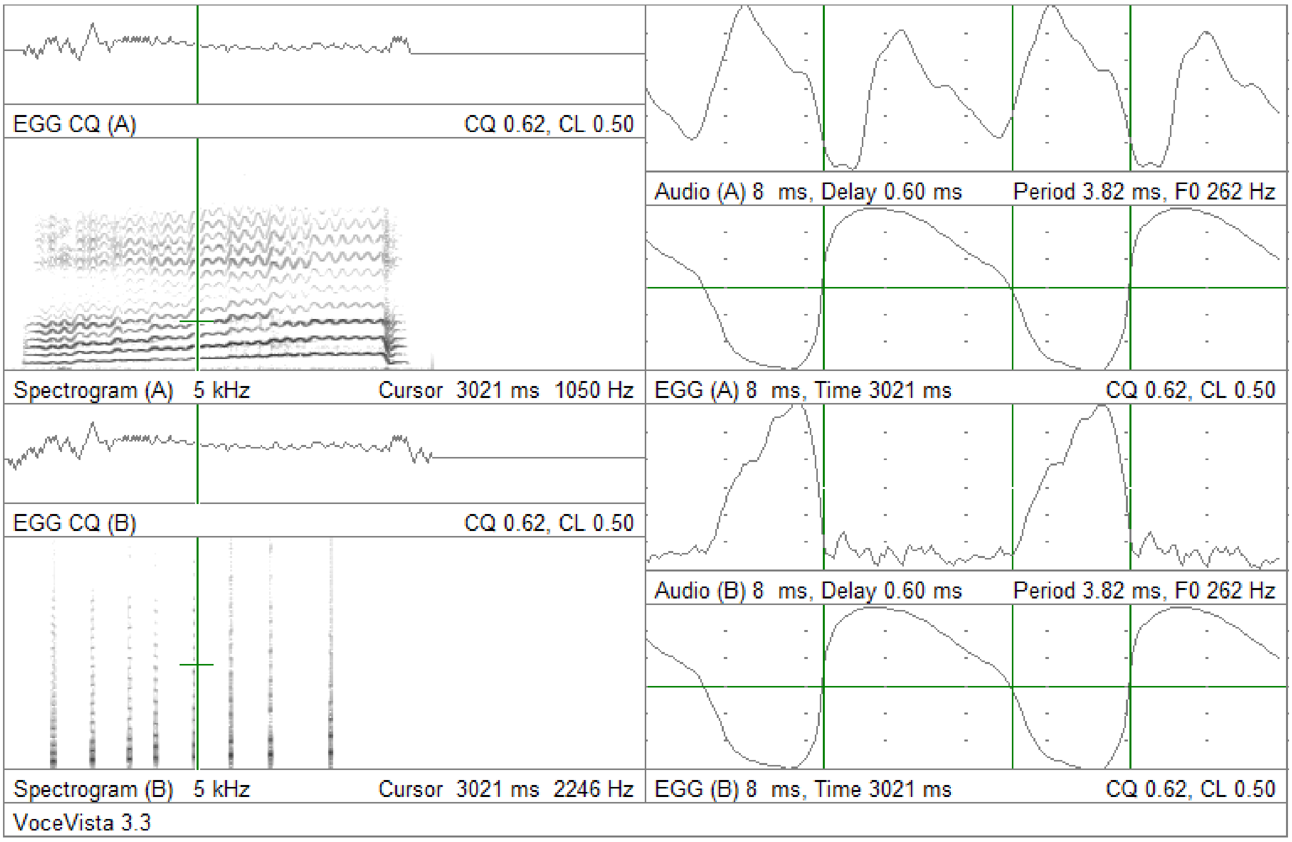

Figure 6 Scale F3-F4 [a] Integrated audio signal (upper) and inverse filtered glottal airflow at C4

If we assume that the integrated Audio signal is a good representation of supraglottic pressure (P_{supra}), we can conclude that P_{supra} is high during the time that glottal airflow is rising, while it is lower during the time that glottal airflow is falling. This would explain why the glottal pulse rises slowly and drops quickly, since a higher P_{supra} decreases airflow while a lower P_{supra} increases it.

In Figure 7 we see the Integrated Audio signal (upper panel) and the inverse filtered Glottal Airflow signal (lower panel) at pitch D4, where H2 ≈ F1. Here, we see a dent in the middle of the pulses Figure 8 shows what happens at pitch E4, where H2 > F1. Now, the pulses are deformed on the right side.

A plausible explanation for the phenomenon that pulse skewing takes place when H2 < F1 seems to be that, while H2 < F1, the second harmonic of P_{supra} is shifted to the left with respect to the Glottal Airflow maximum and thus the second maximum of P_{supra} coincides with the rising edge of glottal airflow, slowing down the buildup of airflow (Figure 6). When H2 ≈ F1, the second maximum of P_{supra} approximately aligns with the maximum of Glottal Airflow, causing a dent in the middle of the pulses (Figure 7), As soon as H2 rises above F1, the second harmonic of Psupra is shifted to the right and thus the Psupra maximum coincides with the falling edge of glottal airflow (Figure 8). Note that in Figure 8 Integrated Audio does not align with the dent in glottal airflow. We assume that this is due to the fact that Integrated Audio is not a perfect representation of P_{supra}.

Finally, Figure 9 displays Integrated Audio and Glottal Airflow for pitch F4, where H3 ≈ F2. Note that the glottal pulses are not skewed in this case, nor are they deformed in any other way. Up until now we have consistently made this observation, also in other examples where F2 aligns with H3 or H4. We conclude that F2 resonance probably causes no significant feedback to the Glottal Airflow signal.

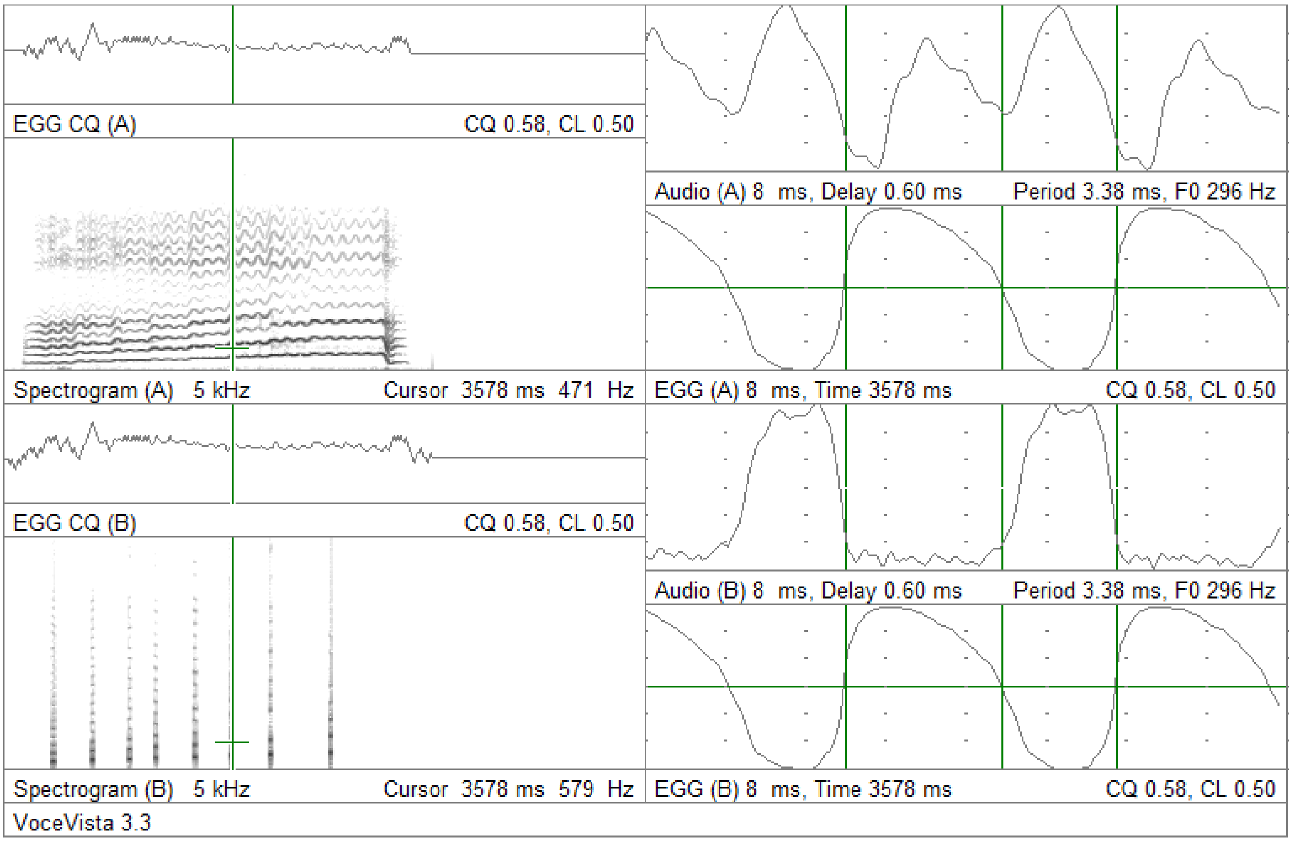

Figure 7: Scale F3-F4 [a] Integrated audio signal (upper) and inverse filtered glottal airflow at D4

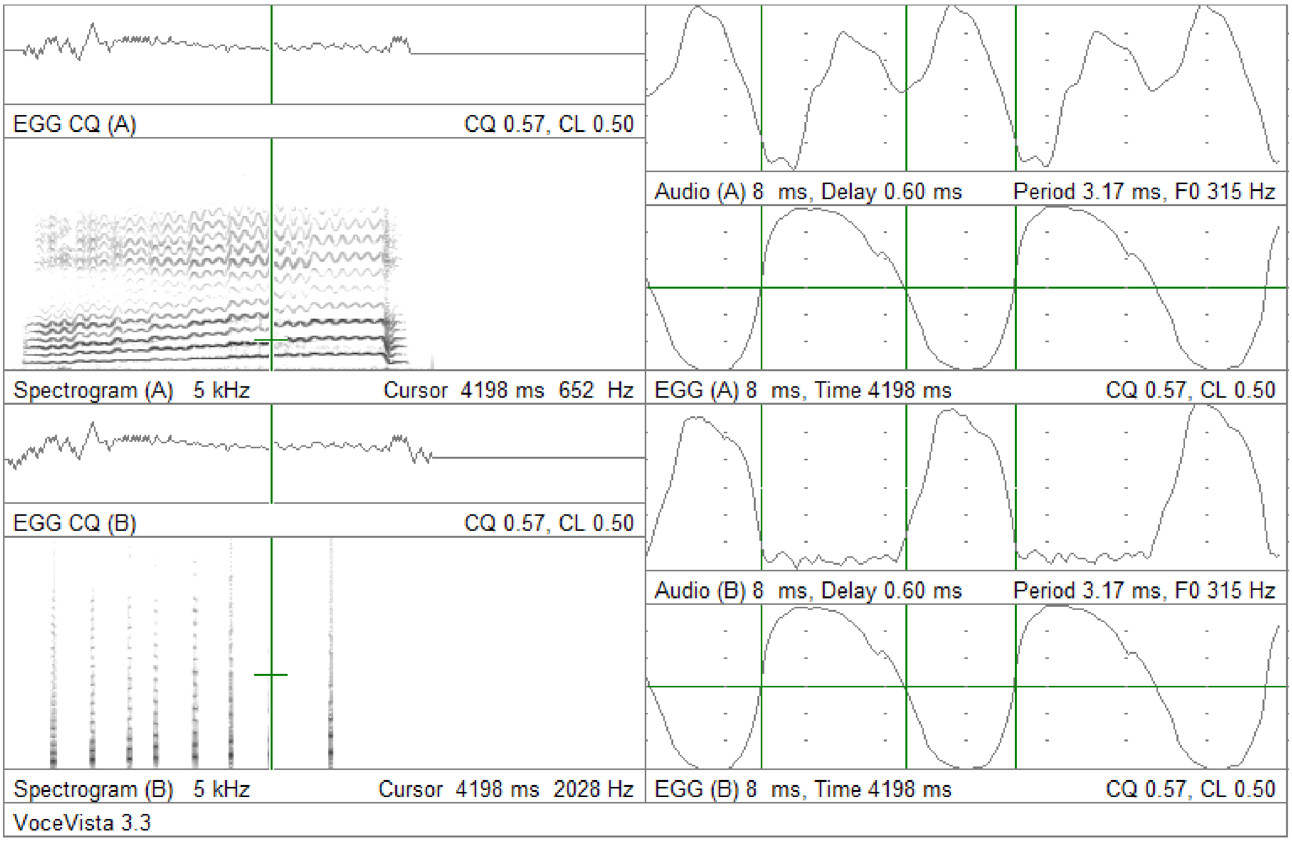

Figure 8 Scale F3-F4 [a] Integrated audio signal (upper) and inverse filtered glottal airflow at E4

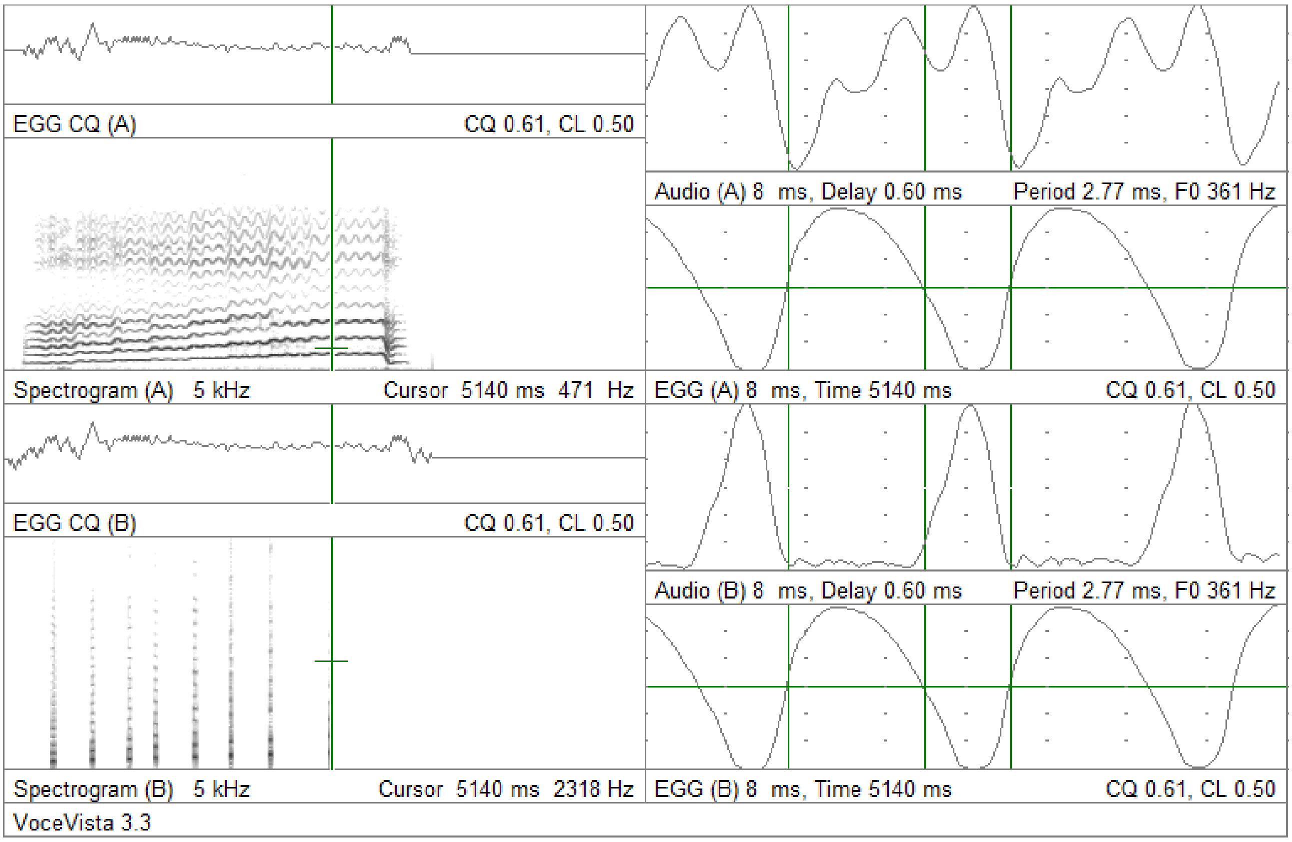

Figure 9 Scale F3-F4 [a] Integrated audio signal (upper) and inverse filtered glottal airflow at F4

Dependency of second formant tuning on the Closed Quotient

In the previous section we have shown that, once the pitch rises above the *secondo passaggio,* which is the pitch where the 1st formant aligns with the 2nd harmonic, the glottal pulses are no longer skewed. The direct consequence of this is that the amplitude of the higher harmonics is reduced to some extent. Also, the amplitude of the second harmonic is reduced, since F1 no longer aligns with H2. Together, the effect is that the sound above the *secondo passaggio* tends to be darker, less 'resonant' and generally weaker than the sound below the *secondo passaggio*. It has been shown (references!) that, in order to compensate for this loss of resonance, many singers use the strategy of tuning the 2nd formant (F2) to a higher harmonic, which is typically H3 for the vowel [a] and H4 for the vowel [E]. This stategy requires careful adjustment of the vowel and can be difficult to achieve for some singers. As we have recently discovered, the succes of F2 tuning also depends on the CQ.

In Figure 10 the spectrogram and power spectrum are shown of two attempts to sing F4# on the vowel [a] by a baritone (WR). In the upper panel the 3rd harmonic is dominant, while in the lower panel the 3rd harmonic is much weaker.

Figure 11 shows the result of inverse filtering the microphone signal at the same points in time as in the two vocalizations shown in Figure 10. It is clear that the CQ in the upper panel is higher than in the lower panel. The vocalization with the higher CQ produces a dominant 3rd harmonic, while in the vocalization with the lower CQ, the dominant 3rd harmonic is lacking. It looks as if the amplitude of the 3rd harmonic is dependent on the CQ and, possibly, on the shape of the glottal airflow pulses. The pulses in the *low CQ* case have a CQ of ≈50% and they are symmetrical. It is known from the theory of Fourier series that a signal with this shape must have a weak 3rd harmonic and a strong 2nd harmonic (cf. the section *Theory and optimal condition* below).

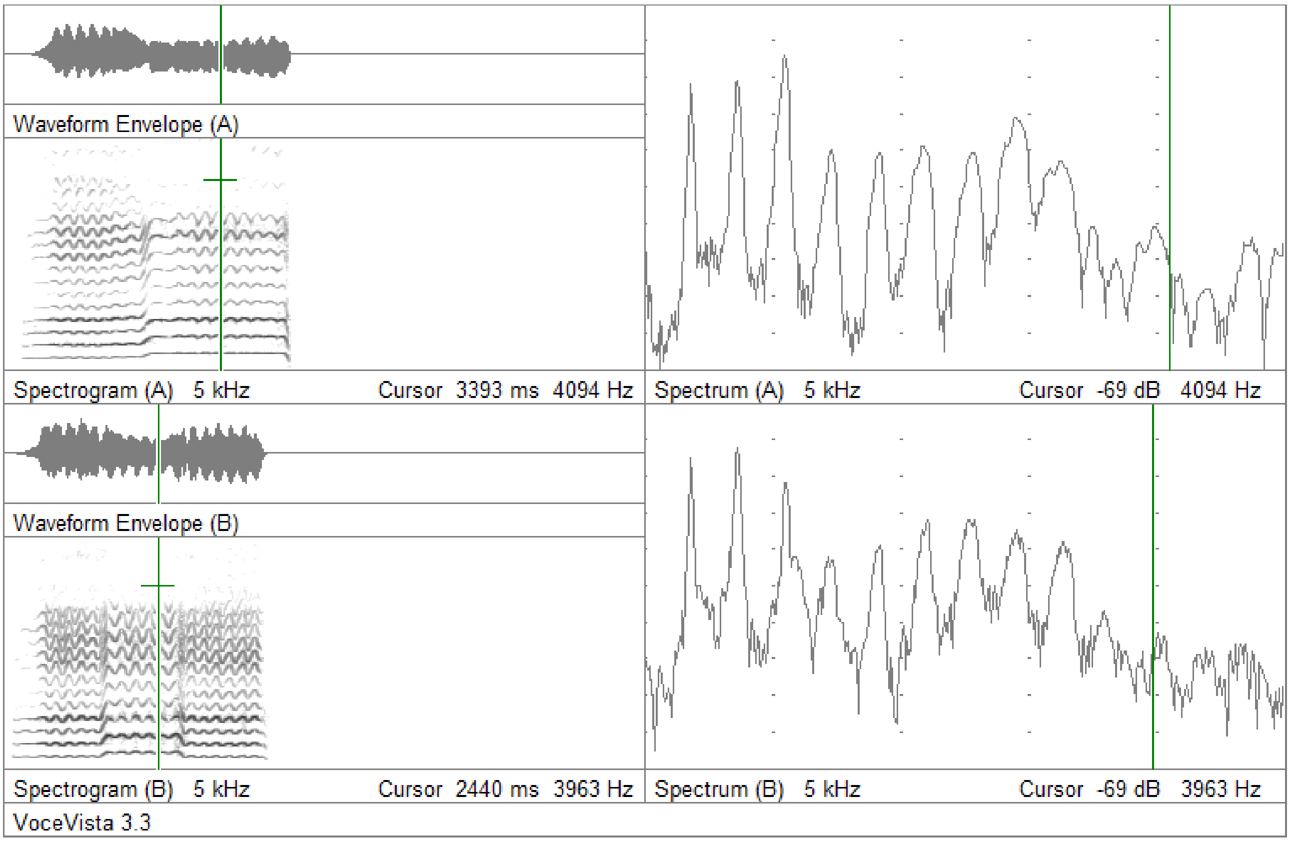

Figure 10: Spectrogram and Power Spectrum of F4# [a] with high CQ (upper) and low CQ (lower)

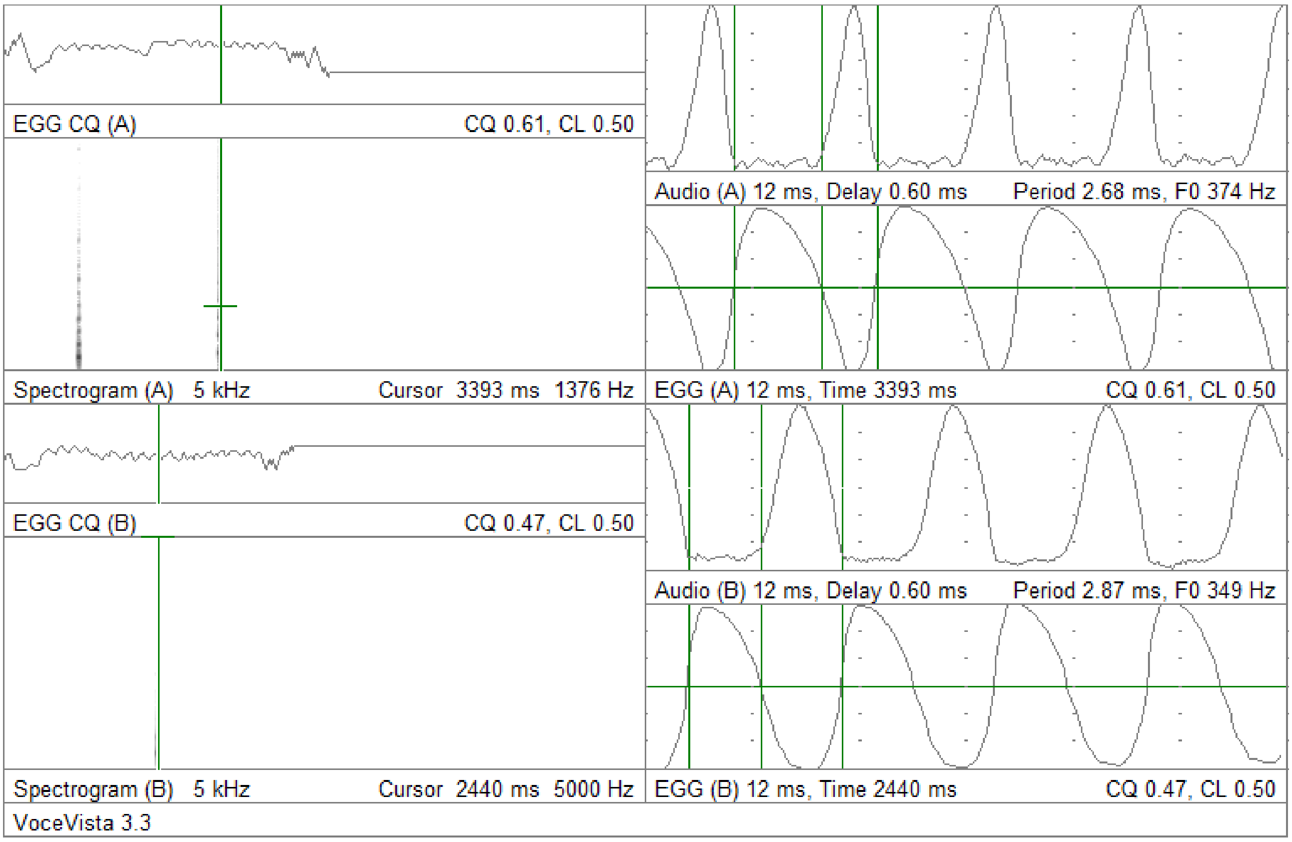

Figure 11 Inverse filtered glottal airflow of F4# [a] with high CQ (upper) and low CQ (lower)

Obviously, the difference in amplitude could also be due to different tuning of the formants, so it is important to make sure that it is the pulse shape and/or the CQ that makes the difference and not formant tuning. A look at the formant frequencies (derived from Inverse Filtering) and the harmonic frequencies tells us that:

- F2=1090 Hz and H3=3*373 Hz=1122Hz (a difference of 32 Hz) for the *high CQ* case
- F2=1139 Hz and H3=3*349 Hz=1047 Hz (a difference of 92 Hz) for the *low CQ* case.

The tuning of F2 to H3 is indeed better for the *high CQ* case, but it is unlikely that this is responsible for the large difference in amplitude of the 3rd harmonic.

Since we have the glottal airflow signal (derived by Inverse Filtering) available in VoceVista, we also have the possibility to display the spectrum of that signal. This has been done in Figure 12. Figure 12 clearly shows the weak 3rd harmonic (lower panel) of the glottal airflow signal in case of the low CQ.

Apparently, an unfavorable pulse shape and/or CQ can reduce the amplitude of the 3rd harmonic of the glottal airflow signal to such an extent as to make F2/H3 tuning fail. A combination of a symmetrical pulse shape and a CQ of ≈50% seems to be such an unfavorable condition.

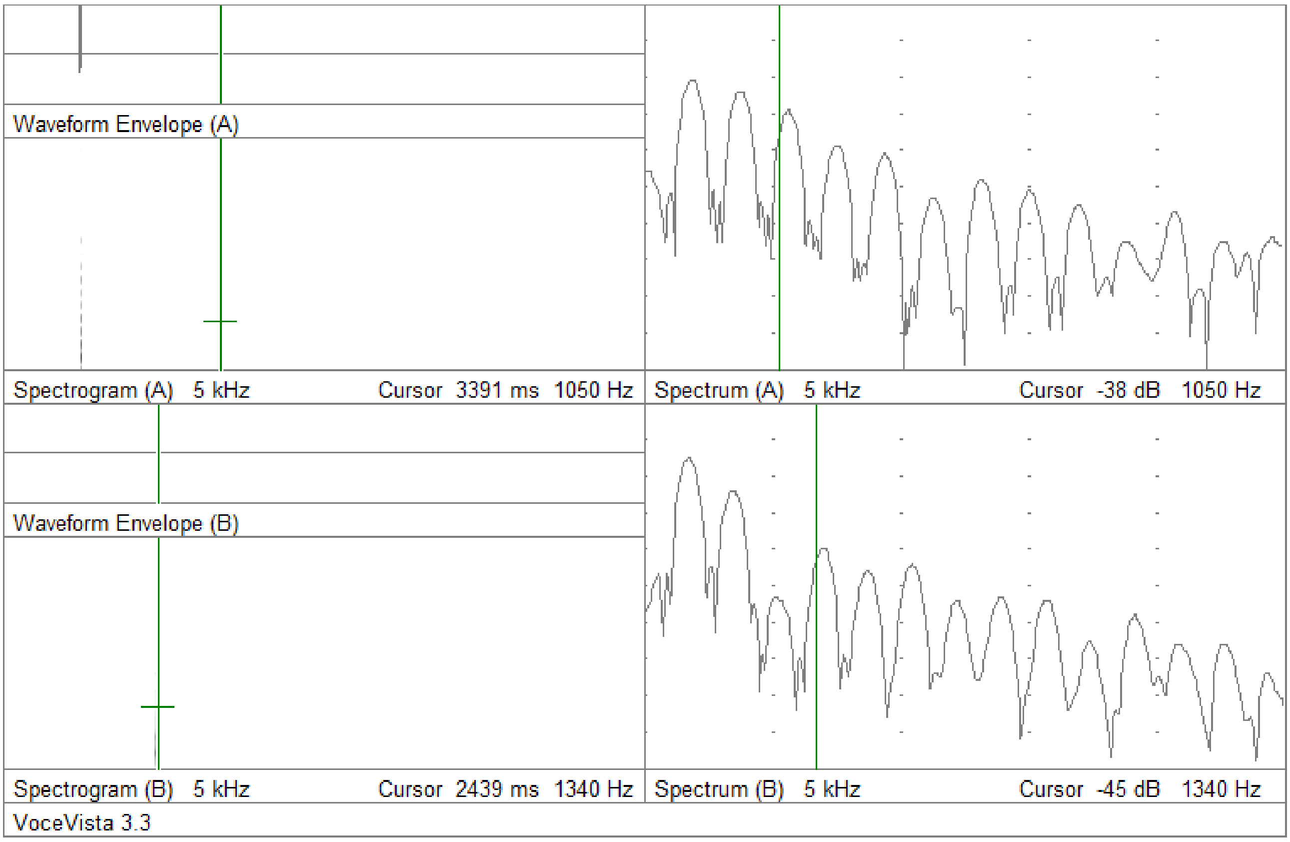

Figure 12 Spectrum of glottal airflow of F4# [a] with high CQ (upper) and low CQ (lower)

Theory and optimal condition

Our statement that symmetrical pulses with a CQ of 50% will produce a weak 3rd harmonic can be backed up by the theory of Fourier series. We start by symplifying the pulse shape in the lower panel of Figure 11 to the pulse shape of Figure 13, the positive part of a sine wave. Then we create a periodic signal by repeating this pulse shape and we calculate the power of the 3rd harmonic of this periodic signal as a function of CQ. The result is shown in Figure 14.

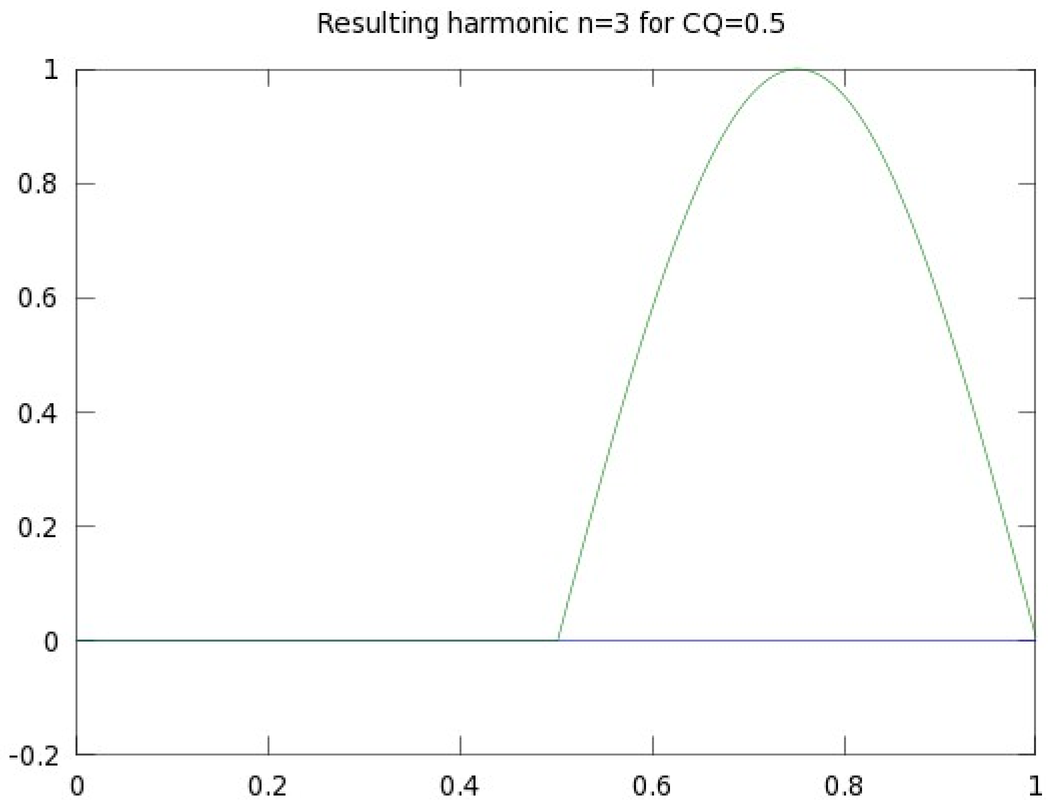

Figure 13: Symmetrical glottal pulse with CQ=50%

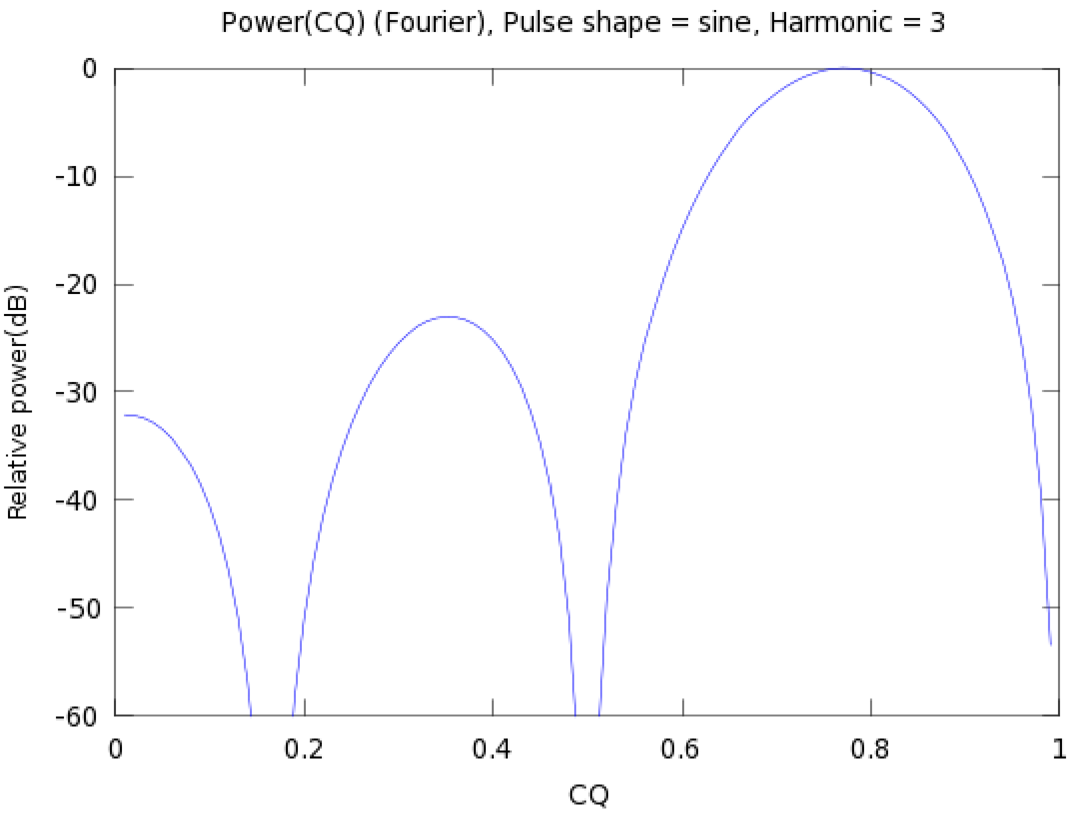

Figure 14: Relative power of the 3rd harmonic in dB for the pulse of Figure 13

For a CQ of 0.5 the power of the 3rd harmonic is zero. This confirms our finding that nearly symmetrical pulses with a CQ of 50% produce a weak 3rd harmonic. A similar calculation reveals that for the idealized pulse shape of Figure 13 also the 5th, 7th , 9th harmonic etc. have zero power. This is clearly not true for the 'real' pulse shape resulting from Inverse Filtering (Figure 11, lower panel).

Figure 14 also gives us an idea of the most optimal CQ for symmetrical pulses i.e. the CQ that yields the strongest possible 3rd harmonic. This optimal CQ lies between 70% and 80%. The CQ in the vocalization with the higher CQ (Figure 11 upper panel) was only 61%. Thus, Figure 14 suggests that the amplitude of the 3rd harmonic could be further improved by raising the CQ to over 70%, provided this is physically possible for the singer.

Another important suggestion that follows from Figure 14 is that lower values of the CQ can also be favorable for the amplitude of the 3rd harmonic (e.g. CQ's between 30% and 40%). This situation is often encountered in female voices in the lower part of their light register.

Conclusions

1. We have found that the glottal airflow pulses are shaped differently above and below the passaggio. Below the passaggio, we see that the pulses are skewed to the right. Above the passaggio, the pulses turn out to be nearly symmetrical. We think this is due to a difference in phase (timing) of the supraglottic wave below and above the passaggio. The absence of pulse skewing above the passaggio possibly explains the darker quality of the sound in this pitch range, as symmetrical pulses are likely to produce weaker high harmonics than skewed pulses. This phenomenon could also explain why many inexperienced singers insist on keeping their first formant above the second harmonic, even at the expense of raising their larynx, simply because this adjustment gives them a brighter sound.
2. The success of second formant tuning above the passaggio depends on the Closed Quotient (CQ). If F2 is tuned to H3 (as for the vowel /a/), a CQ of around 50% results in a weak 3rd harmonic and thus it may not be possible to achieve a spectrum with a dominant 3rd harmonic. The (theoretical) optimal CQ for F2/H3 resonance is 70-80%, suggesting that singers who can produce exceptionally strong harmonics by using second formant tuning, are probably able to produce exceptionally high closed quotients. On the other hand, a CQ of 30-40% is also favorable for generating a relatively strong H3. Therefore, many female singers can easily produce a dominant H3 in the lower part of their light register.
3. We have found no evidence of feedback to the voice source in the case of F2 tuning i.e the pulse shape found for F2 > H3 does not differ from the pulse shape found for F2 < H3.