Professional male singers' formant strategies for the vowel /a/

Johan Sundberg
Department of Speech, Music and Hearing, School of Computer Science and Communication, KTH, Stockholm, Sweden

Filipa M. B. Lã
Department of Communication and Arts, INET-MD, University of Aveiro, Portugal

Brian Gill
Department of Music and Performing Arts Professions
Steinhardt School of Culture, Education and Human Development
New York University, USA

Running title:

Male singers' formant strategies for the vowel /a/

Corresponding author:

Johan Sundberg, Professor, Department of Speech, Music and Hearing, School of Computer Science and Communication, KTH, SE-10044 Stockholm, Sweden

Phn +468 790 7873

Email pjohan@speech.kth.se

ABSTRACT

Certain spectrum characteristics have been identified as important for register equalization around the male *passaggio*, an effect ascribed to formant tuning although descriptions of formant tuning diverge. Eight professional singers sang scales including their *passaggio* range on different vowels, applying two formant tuning strategies as found in classical and non-classical singing. Formant frequencies were measured using inverse-filtering. Results revealed differences between the two strategies. For the classical tuning, systematic changes of formant frequencies with pitch were observed. For the highest note sung on /a/, F1 was below the second partial and F2 the vicinity of the third. Formant frequencies were quite close to or coinciding with a partial. Spectrum characteristics were achieved by different F1 and F2 values between singers.

Keywords

Formant tuning, Male singers, Spectrum, Harmonics

Professional male singers' formant tuning strategies for the vowel /a/

ABSTRACT

Certain spectrum characteristics have been identified as important for register equalization around the male *passaggio*, an effect ascribed to formant tuning although descriptions of formant tuning diverge. Eight professional singers sang scales including their *passaggio* range on different vowels, applying two formant tuning strategies as found in classical and non-classical singing. Formant frequencies were measured using inverse-filtering. Results revealed differences between the two strategies. For the classical formant tuning, systematic changes of formant frequencies with pitch were observed. For the highest note sung on /a/, F1 was below the second partial and F2 in the vicinity of the third. Formant frequencies were quite close to or coinciding with a partial. Spectrum characteristics were achieved by different F1 and F2 values between singers.

INTRODUCTION

Many singing teachers and classically trained singers focus on developing strategies to achieve maximum vocal output which compromises neither vocal health, nor evenness of timbre (1). A commonly reported strategy involves the adjustment of the lower formant frequencies to match certain partials with the goal being to gain vocal power.

This adjustment, which is achieved by a combination of phonatory and articulatory means, has been referred to as formant tuning (2). Acoustically, it can be described as tuning the first two formants to the vicinity of spectrum partials. As fundamental frequency (F0) rises, the distance between harmonic partials increases. This implies that there is a need for rather specific articulatory changes at high pitches in and above the male *passaggio* range, i.e., near the pitches of E4 to G4 (3). This formant strategy seems related to so-called covered singing (4). By adopting it, singers are assumed to avoid the risk of vocal hyperfunction and register breaks when singing high notes (4; 5).

Despite several studies on formant tuning amongst singers, it is still a question whether or not there are specific benefits of using this technique and, if used, how it would be applied to different vowels along the frequency range, particularly around the *passaggio*. Several studies have tried to answer this question; however, their results still do not provide definitive conclusions. On the one hand, Carlsson and Sundberg (6) synthesized two versions of a scale on the vowel /a/, one with constant formant frequencies and the other with a formant tuning strategy. They presented these stimuli pairwise to a panel of expert listeners and asked them what version they preferred. All listeners except one preferred the constant formant frequencies version. Hertegård and associates (4) analyzed formant frequencies in 11 professional classically trained male singers singing covered and open versions of the vowel /ae/ in their passaggio region. In the covered versions, the singers were found to lower F1 when singing in their passaggio. This effect was achieved by articulatory changes including lengthening of the vocal tract and a widening of the pharynx. Furthermore, the voice source was modified towards flow phonation, commonly regarded as a method to avoid hyperfunctional strain of the larynx (7). Perceptually, these modifications of formants and voice source produced an impression of darkening of the vocal timbre.

Using several methods, including audio, EGG and sub- and supraglottal pressure transducers, Miller and Schutte (2) analyzed the formant strategy of a professional baritone singer and found evidence of formant tuning that boosted the output sound pressure. These findings were later corroborated by spectral analysis of ingressive phonation and vocal fry phonation following singing (8). Schutte and coworkers (9) found further support for this by spectrum analysis of 80 commercial recordings of 34 different tenors' renderings of the final sustained B4-flat of the aria "Celeste Aida" from Verdi's opera Aida on the vowel /o/. They found that most frequently H3 (about 1400 Hz) was particularly prominent and they interpreted this as a sign that F2 was tuned to this partial. According to Miller (10), the tuning of either F1 or F2 to a harmonic partial is necessary for appropriate voice production of open vowels sung by males above the passaggio.

Neumann and associates (11) examined the spectra and contact quotient in 11 professional classically trained male voices. From the spectrum data they concluded that the singers tuned their F1 and F2 in back vowels to H2 and H4, respectively, when they reached the upper boundary of their 'chest' register (11). At the *passaggio* transition, the singers lowered their F1 below H2 such that it no longer boosted this partial; at the same time they shifted F2 from H4 to H3. The authors claimed that if this specific tuning strategy is not applied, the larynx position is likely to rise, causing "register violation". Using MRI and acoustic analysis Echternach (12) measured the vocal tract shape and the associated formant frequencies in premiere opera tenors. Like Neumann and associates, he found that for the vowel /a/, the singers tuned F1 to a frequency below H2 in their highest tones.

Titze (13) has analyzed, from a theoretical point of view, the consequences of tuning the lowest formants to spectrum partials. He applied a computerized model of the voice organ that allowed inclusion or elimination of non-linear source-filter interaction. The results suggested that sound intensity could be increased by 10 dB due to such interaction, if F0 was located in a zone of inertive reactance in the vocal tract. This condition occurs when F1 is just above a partial. When F1 is just below a partial, i.e., in the zone of capacitive reactance, the opposite effect occurs. The author concluded that for the purpose of producing maximum acoustic power, tuning F1 to a frequency slightly higher than that of one of the three lowest spectrum partials is advantageous. Further, certain combinations of vowel and F0 may necessitate the reversed situation, namely that F1 is located at a frequency just below one of these partials. In such cases, Titze claims that avoiding non-linear source-filter interaction is required. In the light of the theoretical model, Titze and Worley (14) presented further support for this conclusion by interpreting photographs of operatic and pop singers' mouth openings.

Further theoretical work with the model showed that frequency jumps and other instabilities may occur at high F0, when formants are near harmonics. The results also suggested that register breaks may occur unexpectedly, when unfavorable inertive-compliant conditions are combined with a narrowing of the epilaryngeal tube. On the other hand, vocal efficiency can increase by an order of magnitude if the cross-sectional area of that tube is narrowed from 3 to 0.2 cm^2. Thus, contrary to some previous work (9; 10) Titze concludes that tuning a formant to the frequency of a harmonic would be disadvantageous, since non-linear source-filter interaction would then produce instability of vocal fold vibration (15).

Summarizing, reports on formant tuning strategies in and above the male *passaggio* range and how it affects phonation diverge. Hence, a systematic analysis of the formant strategies applied by professional male singers in and above the *passaggio* region seemed indicated. Here a preliminary report will be given on formant frequencies analyzed by inverse filtering in classically trained male vocal artists applying two different tuning strategies.

METHODS

Participants and equipment

Eight classically trained professional male singers, three tenors and five baritones, age range between 23 to 42 years, volunteered as participants, See Table 1. One was an international opera singer, three were national opera singers, and the other four were full-time graduate voice students.

All singers were recorded in a sound-treated studio in the Steinhardt School of Culture, Education and Human Development at New York University (NYU). Recordings were made using a combination of a digital Laryngograph microprocessor and a Glottal Enterprises MS-110 computer interface, thus allowing simultaneous recording of audio, electrolaryngograph (ELG) and flow signals. Flow was collected by means of a Rothenberg flow mask. The audio channel was calibrated by recording a sine tone, the sound pressure level (SPL) of which was measured by a sound level meter held next to the recording microphone. This SPL value was announced in the recording. Flow was calibrated using the Glottal Enterprises flow calibration system, provided with a flow-head device. The three recorded signals were digitized and sent over a USB contact into a PC equipped with the Speech Studio software; thus, audio, ELG and airflow signals were obtained as separate tracks of digital wav files.

Vocal task

The singers were asked to sing ascending/descending major scales, extended up to the ninth so that they included their *passaggio* range, i.e., ≈ E4 – G4, depending on voice classification. These scales were sung on the vowels /ae/, /a/, /u/ and /i/. The singers were asked to apply two different formant tuning strategies. First, they used a vocal approach commonly regarded as appropriate in classical singing (performed twice). In this study we will refer to this strategy as classical. They were also asked to sing one ascending/descending scale using a tuning strategy commonly accepted in musical theatre and popular music. Henceforth, we will refer to this strategy as non-classical, as it is not commonly accepted in classical music. Singers arrived with their voices already warmed up. A reference pitch was provided at the beginning of each scale.

Data analysis

Listening test

The total number of recorded scales amounted to 288 - classical: 8 singers x 3 starting pitches x 2 takes x 4 vowels; non-classical: 8 singers x 3 starting pitches x 1 take x 4 vowels. However, as three participants agreed with co-author BG that they failed to produce the typical examples of the classical tuning strategy, 180 scales were deemed useful for a listening test. Considering this high number of stimuli, it seemed appropriate to include only a selection of scales, randomly chosen. Thus 28 scales were used for the listening test. Again in order to limit the duration of the test, only thirteen stimuli were repeated, producing a total of forty-one scales; the duration of the test was about 11 min.

The listening test, stored in the form of a digital computer wav file, was sent as an email attachment to a panel of six expert listeners, who were all singing teachers at the university level in the United States. The stimuli were presented in a numbered order and with a pause of 5 seconds in between. The panel members were asked to rate, along 100 mm visual-analogue scales, how successful the singer was with regard to vocal tract tuning in transitioning into the pitches above the *passaggio*. In this way quantitative estimates were obtained of how typical the example was with regard to the tuning strategies used in classical vocal music.

Stimuli receiving a mean rating above 65% of full VAS length (top third) were accepted as examples of the classical version, whereas, those with a mean rating below 34% (bottom third) were regarded as clear examples of non-classical

These boundaries resulted in a material including five pairs of classical and non-classical versions of the scale in which both versions were produced by the same singer. Of these pairs three were sung on the vowel /a/, one on /ae/ and one on /i/. Analyses were limited to the vowel /a/. In addition to the three paired versions of /a/, an /a/ rated in the middle of the scale (44%) was included in the material.

Inverse filtering

The formant frequencies were determined by inverse filtering the audio signal by means of the custom made program *Decap* (Svante Granqvist, KTH). The program applies the classical equations to calculate the transfer function corresponding to the given combination of formant frequencies and bandwidths. The input signal is filtered with this transfer function inverted. This implies that the effects of the vocal tract transfer function are eliminated from the input signal. In the present application the program displayed (1) the inverse filtered audio waveform in terms of a flow glottogram, i.e., as flow versus time, (2) the derivative of the ELG signal (Lx), delayed by a time interval corresponding to the travel time of the sound from the glottis to the microphone i.e., vocal tract length estimated at 18 cm plus the mouth to microphone distance, (3) the spectrum of the input audio signal and (4) the spectrum of the flow glottogram, see Figure 1. The program displays these results of the filter settings in quasi-real time.

Figure 1 about here

A ripple free closed phase and a source spectrum envelope as void of local minima as possible were used as the criteria for setting the frequencies and bandwidths of the inverse filters. After completing the tuning of the filters, their frequencies and bandwidths were saved in a file. These files were then opened and analyzed in the excel program.

F0 was extracted from the Lx signal by the FoX program of the Soundswell signal workstation. Using its Histogram option, the frequencies of the individual scale tones were measured as the average over a time interval comprising a set of complete vibrato cycles.

RESULTS

Listening test

The consistency of the raters participating in the listening test was evaluated in terms of the average and the standard deviation of the absolute difference between the first and second ratings of replicated stimuli, as well as in terms of the Pearson's correlation between these ratings, see Table 2. The mean difference varied between 5.5 and 18.3 mm and the correlation between 0.75 and 0.99.

The results of the listening test are shown in Figure 2 in terms of the mean ratings, rank ordered along the abscissa. The span of the means is 10% to 89% of the available rating scale. All of the stimuli intended as classical obtained higher mean ratings than those intended as non-classical. In this sense, these results corroborated the initial classification of the scales.

Figure 2 about here

To assess whether there were significant differences between the ratings for the classical and non-classical examples included in the listening test, a non-parametric paired sample test –Wilcoxon- was performed, see Table 1. This particular test was used because the data showed a skewed distribution. Classical versions were scored significantly higher than non-classical versions for all singers, except for singers 1 and 3, for the vowels [u] and [a], respectively.

Formant frequencies

Figure 3 shows the formant frequencies used by the singers. In general, there were distinct differences between the classical and non-classical versions. In the classical version, F1 showed a trend in all singers to fall below H2 at some point near the highest notes of the scale. At some scale tones in the classical version, all singers tuned F1 almost exactly to H2. Singers 1 and 3 kept their F1 almost constant throughout the scale. F2 showed a weak trend to increase with increasing F0 and ended close to H3 in all singers: it was tuned just above H3 for two of the singers, directly on H3 for one singer, and just below H3 for the remaining singer.

Figure 3 about here

For three singers the classical and non-classical have quite similar formant frequencies for the three lowest scale tones but separate for the higher ones. In the non-classical version, F1 showed a tendency to rise with increasing F0. For the top pitch, it was almost identical with H2 in three singers and higher than H2 in one singer. F2 was higher in the non-classical than in the classical versions for three of the singers and almost identical with the classical version in one singer, although it was higher in the top note. Singer 2 tuned his F2 exactly to H3, both in the classical and non-classical version of his highest note.

The three top tones of the scale were all in the singers' *passaggio* region. These tones are generally considered to be of particular interest. The column plots in Figure 4 show how far, in semitones, F1 and F2 were from their nearest partials in the classical and non-classical versions of the scales. As previously mentioned, in the classical versions F1 is moved from above to below its nearest partial (H2) in all singers. In the top note, F1 is more than 4 semitones below H2, except for singer 3, where it is only 1.4 semitones below. However, it may be relevant that his mean rating in the listening test was quite low (44%). Also, singer 3 increased the distance of his F2 to H3, ending 4 semitones below for the highest note, while the other singers moved F2 closer to H3 (within 1.5 semitones). The non-classical versions showed a great deal of inter-individual variation, but for the top note, F1 is very near H2 for three of the singers.

Figure 4 about here

Summarizing, the singers' formant tuning strategies seem complex and diverse and showed some inter-singer variability. Nevertheless, there seemed to be a common denominator in terms of the rising spectrum envelope over the three lowest partials for the top note of the scale. This is illustrated in Figure 5a, b, c. d, showing Decap displays. The panels present the flow glottograms, the Lx derivative and the spectra. Note that these graphs illustrate how the input audio signal can be explained as the consequence of the formant frequencies and bandwidths used in the inverse filtering, combined with the voice source spectrum shown..

Figure 5 about here

While a rising spectrum envelope of the three lowest audio spectrum partials was a common denominator for all classical versions, H2 was the strongest of the three lowest partials in the non-classical versions. It can further be noted that the flow glottograms possess the typical characteristics. They present (1) a ripple free closed phase, (2) more or less skewed flow pulses, (3) synchrony between the ceasing of the air flow and the positive peak of the derivative of the Lx signal, and (4) synchrony between the onset of the air flow and the and negative peak of the derivative of the Lx signal. Finally, the source spectrum envelopes slope smoothly, except for singer 1, who tuned F2 to H3. In this case, the nearest higher partial (H4) appeared attenuated in the source spectrum (Figure 5a).

The levels of the ten lowest source spectrum partials are shown for the highest pitch in Figure 6. The frequencies of the F1 and F2 are pointed by arrows. The lines represent approximations of the main overtone spectrum slope, derived as trendlines for partials 2 – 10. In the classical versions the partials are closer to these lines than in the non-classical versions. In other words, the source spectrum comes closer to the predictions of the source filter theory of voice production in the classical versions. Also, it is noteworthy that partials coinciding with a formant do not deviate more from the lines than their neighboring partials.

Figure 6 about here

These results imply that the audio spectrum can be explained by combining a normal voice source spectrum with the formant frequencies shown in the graphs. The graphs illustrate how proximity (above or below) or exact match between F1 and F2 and their nearest harmonics, respectively, define the rising envelope of the three lowest spectrum partials.

DISCUSSION

Measuring formant frequencies at high F0 is a difficult task, so a question of prime relevance is to what extent our results are reliable. As mentioned, inverse filtering was used for formant measurements. This method seemed particularly appropriate, since it allows a more accurate determination of F1 and F2 than other methods, such as spectrography or linear predictive coding. It is based on two assumptions: (1) the transfer function of the vocal tract can be predicted, given the frequencies and the bandwidths of the formants (16); and (2) the voice source is defined as the pulsating transglottal airflow. The strength of inverse filtering is that is shows both waveform and spectrum. This offers a double guide for tuning the frequencies of the inverse filters, i.e. the formant frequencies. On the other hand, the reliability of this method decreases with increasing F0. The highest pitches analyzed in the present study were near 400 Hz. Still, there are reasons to believe that our formant data are reliable, as (1) the flow glottograms showed typical characteristics found for lower F0 values, (2) the source spectra showed the smoothly falling envelope, basically void of peaks and troughs, which is typically found when analyzing vowels at lower F0 values, and (3) the formant frequencies did not show major changes between adjacent scale tones.

In addition, the vocal behavior of the singers and the listening tests provided support for the assumption that the data gathered were representative. The participants were skilled singers and tried to follow the instructions accurately. If participants and the singing teacher were not satisfied with the examples provided of the classical and non-classical versions, they were re-recorded. Also, the results of the listening test showed clear cases of classical and non-classical versions, and these were selected for analysis.

The main findings of this investigation were that F1 and F2 differed between the two formant tuning strategies, especially concerning the highest notes. For the classical version, singers tuned F1 to a frequency below H2 at or near G4 (383 Hz), whereas for the non-classical versions, it was higher. For all non-classical versions, F2 tended to be higher than in the classical versions.

The present results can now be compared with those reported in earlier investigations. As mentioned in the Introduction, Miller (10) regards the tuning of F1 or F2 to certain partials as a necessary attribute of appropriate singing technique in open vowels sung by males above the *passaggio*. Neumann and associates (11) found that in register transition between 'chest' and 'head' on open back vowels "*F2 starts to resonate H3 when F1 cannot follow the rising H2 anymore. When singing through passaggio a professional singer anticipates and avoids a register violation by allowing F1 to fall below H2 before both become locked together by an elevated larynx*" (p 321 and 324). Our findings agree with these studies to the extent that we found the singers tuned their F1 to a frequency below H2 for the highest note or notes. However, Neumann and associates (11) noted that F2 was tuned to H4 in the high range of the 'chest' register. We found that this situation happened at single pitches (e.g. singer 3 around 270 Hz) when F0 was in the appropriate range for this to happen by coincidence, i.e., it occurred without any marked changes of F2 between scale tones (as shown in Figure 3). The main reason for the coincidence seemed to be that F0 changed while F2 remained basically the same as in the neighboring scale tones.

According to Titze (15) instability of phonation can be expected when F1 coincides with H2 because of a non-linear source-filter interaction. Several cases of such coincidence were found (singer 1 at 350 Hz; singer 2 at 340 Hz; singer 3 at 320 Hz; and singer 4 at 300 Hz). Titze further predicts that instability can be expected also when F1 is placed below H2, a situation observed in several of the highest notes analyzed (singer 1 for the highest pitch, F0 = 390 Hz; singer 2 at F0>320 Hz; singer

3 at F0 = 370 Hz; and singer 4 at F0>340 Hz). However, no cases of instability were observed in our recordings.

Non-linear source-filter interaction is likely to boost certain partials in the radiated spectrum. Inverse filtering only implies that the effects of the vocal tract transfer function on the radiated sound are eliminated. If non-linear source-filter interaction is present it would appear as an irregularity of the source spectrum envelope emerging from inverse filtering analysis. In the classical versions of the exercise no clear irregularities were observed in the source spectrum envelope, not even in the region of the singer's formant, as illustrated in Figure 6. Thus, an appropriate clustering of F3, F4, and F5 yielded a smoothly falling source spectrum envelope. Such a formant clustering has been observed to occur also in an acoustical model of the vocal tract, which contained a representation of the larynx tube and the pyriform sinuses (17). Thus, it seems that the singers must have managed to avoid a non-linear source filter interaction. The method they used for achieving this is an open question.

As shown in Figure 3, there were several instances where F1 and or F2 coincided with the frequencies of one of the lowest spectrum partials. The frequency response of a tube resonator rises and falls sharply just around the formant frequency. The slope of the frequency response curve nearest the formant depends on the bandwidth. The bandwidth of F1 can be about 50 Hz in a non-nasalized /a/ vowel. If so, a 25 Hz F0 change may cause an amplitude modulation of nearly 6 dB, just because of resonance. Such an F0 change corresponds to 1.2 semitones at about 320 Hz, which is a typical peak-to-peak modulation amplitude in vibrato. In other words, the singers are likely to produce a great amplitude modulation in cases where they tuned F1 near the second spectrum partial. This may or may not contribute to the voice timbre. In any event, it implies that the audio spectrum will vary substantially over the vibrato cycle. For this reason, the spectra in Figure 5 were taken at the upper turning point of the vibrato cycle.

In the same Figure it can be seen that the singers tuned their F1 and F2 differently relative to H2 and H3 for the highest note in the classical version. Yet, they produced the same audio spectrum shape, a rising spectrum envelop over the three lowest partials. This variation of formant tuning strategies for achieving a specific spectrum property might have been a response to anatomical differences between the singers. These inter-individual differences in formant tuning strategies appear to demonstrate the potential relevance of undertaking case studies to assess formant strategies in singers. The consistent spectrum characteristics for the highest notes of the exercise support the assumption of usefulness of real-time spectrum displays in the singing studio. It would help students to direct their attention to a relevant dimension of sung tones. This would belong to the background of the increasing use of VoceVista and other software displaying the audio signal spectrum in real-time.

The results obtained in the present investigation invite various future studies. First, our analyses have yielded a large set of flow glottograms that merit detailed analysis with regard to the customary parameters. Another tempting study is the assessment of formant strategies that these singers applied for other vowels. This study could be realized using the analysis method tested in the present investigation. Also, the observed formant strategies and audio spectrum characteristics invite a listening test with synthesized singing.

CONCLUSIONS

The main findings of the present investigation of professional baritone and tenor singers' renderings of ascending/descending scales were that the classical and non-classical formant tuning strategies differed in a consistent way. First, the strategies could be clearly perceived as different. Second, they were characterized by differing formant frequency patterns: F1 and F2 tended to be lower in the classical versions and F1 was between 1.5 and 5 semitones lower than H2 for the highest tone. Third, a rising spectrum envelope over the three lowest partials in the highest notes was a common denominator for the classical versions of the highest tones. Fourth, F1 coincided with H2 at some scale tones in all singers' classical versions. Also in the non-classical versions, F1 and F2 apparently happened to coincide with a lower spectrum partial for some tones in the scale, without being preceded or followed by a clear change in their frequencies. Fifth, inverse filtering analysis of the tones produced with the classical formant tuning strategy showed no clear signs of a non-linear source filter interaction, neither when F1 or F2 coincided with a spectrum partial, nor when F1 was slightly lower than a partial. Thus, the major characteristics of the spectra produced by these singers with the classical formant tuning strategy could be accurately explained by the classical linear source-filter theory of voice production. On the other hand, in examples produced with a non-classical formant tuning strategy, some evidence of a non-linear source filter interaction was found in terms of source spectrum irregularities.

Acknowledgements

The kind cooperation of the singers is gratefully acknowledged. A preliminary version of this investigation was presented at the *Physiology and Acoustics of Singing* conference at KTH, Stockholm, in August 2010. The authors are indebted to the Schering Health Care Ltd. & Bayer Portugal for providing the means to acquire the equipment, and to the University of Porto for support.

Figure captions

Figure 1. Decap display showing top panel: input signal, middle panel: filtered signal and bottom panel: spectrum before and after filtering. The arrows in the bottom panel show the frequencies and bandwidths of the inverse filters. The curves show the upper and lower limits of typical bandwidth values.

Figure 3. F1 and F2 observed for the vowel /a/ in the classical and non-classical versions of the scale (solid and dotted curves, respectively). Dashed lines refer to the frequencies of the spectrum partials. The mean ratings are given in % of VAS length.

Figure 3. F1 and F2 observed for the vowel /a/ in the classical and non-classical versions of the scale (solid and dotted curves, respectively). Dashed lines refer to the frequencies of the spectrum partials. The mean ratings are given in % of VAS length.

Figure 4. Distance in semitones, between F1 and its nearest partial and between F2 and its nearest partial (two leftmost and rightmost groups of columns,

respectively) observed for the indicated singers in the classical and non-classical versions of the scales (open and hatched columns, respectively). Negative values refer to the case that the formant frequency was lower than the frequency of the nearest partial. The mean ratings are given in % of VAS length.

Figure 5a - e. Decap displays (see Figure 1) of the top tone sung by the indicated singers.

Figure 6. Amplitudes of the voice source spectra as obtained from the inverse filtering of the top tone. Diamonds and triangles refer to the classical and non-classical versions, respectively. The lines represent the trendline computed over partials 2 – 10.

REFERENCES

6. Carlsson G. & Sundberg, J. Formant Frequency Tuning in Singing. *Journal of Voice* 1992; 6: 256-260.

1. Coffin B. Overtones of Bel Canto. The phonetic basis of artistic singing. Metuchen, NJ: Scarecrow Press, 1980.

3. Doscher, B. M. The Functional Unity of the Singing Voice. 2nd Edition. The London: Scarecrow Press Inc. 1994

12. Echternach M. Wissenschaftliche Untersuchungen zu Registerübergängen bei männlichen Stimmen. Habilitation, Albert-Ludwigs-Universität Freiburg, 2010.

16. Fant G. Acoustic Theory of Speech Production, Mouton: The Hague, 1960 (second printing, 1970).

7. Gauffin J, Sundberg J. Spectrum correlates of gottal voice source waveform characteristics. Journal of Speech and Hearing Research 1989; 32: 556-565.

4. Hertegård S, Gauffin J, Sundberg J. A comparison of subglottal and intraoral pressure measurements during phonation J Voice1990;9:149-155.

5. Hirano M, Vennard W, Ohala J. Regulation of register, pitch and intensity of voice. An electromyographic investigation of intrinsic laryngeal muscles. Fol. Phoniat 1970;22:1- 20.

10. Miller DG. Resonance in singing. Voice building through acoustic feedback. Princeton, NJ: Inside View Press, 2008.

2. Miller DG, Schutte HK. Formant Tuning in a Professional Baritone. J Voice, 1990; 4: 231-237.

8. Miller DG, Sulter AM, Schutte HK, Wolf RF. Comparison of vocal tract formants in singing and nonperiodic phonation. J Voice 1997;11:1-11.

11. Neumann K, Schunda P, Hoth S, Euler HA. The interplay between glottis and vocal tract during the male passaggio. Fol. Phoniatr. Logop. 2005;57:308–327.

9. Schutte, HK, Miller DG, Duijnstee M. Resonance strategies revealed in recorded tenor high notes. Folia Phoniatrica et Logopaedica 2005;57:292-307.

17. Sundberg J. Articulatory interpretation of the 'singing formant', J Acoust Soc Amer 1974;55:838-844.

13. Titze IR. A theoretical study of F0-F1 interaction with application to resonant speaking and singing voice. J Voice 2004;18:292-298.

15. Titze IR. Nonlinear source-filter coupling in phonation: Theory. J Acoust Soc Amer 2008;123:2733-2749.

14. Titze IR, Worley AS. Modeling source-filter interaction in belting and high-pitched operatic male singing J. Acoust. Soc. Am 2009;126:1530-1540.