Title:
Voice source variation between vowels in male opera singers

Running title:
Voice source and vowels in male singers

Authors:
Johan Sundberg
Department of Speech, Music and Hearing, School of Computer Science and Communication, KTH, Stockholm, Sweden and
University College of Music Education Stockholm

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

Corresponding author:
Johan Sundberg, Professor, Department of Speech, Music and Hearing, School of Computer Science and Communication, KTH, SE-10044 Stockholm, Sweden
Phone: +468 790 7873
Email: pjohan@speech.kth.se

Abstract

Objectives: The theory of non-linear source-filter interaction predicts that the glottal voice source should be affected by the frequency relationship between formants and partials. An attempt to experimentally verify this theory is presented.

Study design: Glottal voice source and electrolaryngograph (ELG) signal differences between vowels were analyzed in vowel sequences, sung at four pitches with the same degree of vocal loudness by professional opera singers. Also, the relationships between such differences and the frequency distance between the first formant (F1) and its closest partial were examined.

Methods: A digital Laryngograph microprocessor was used to simultaneously record audio and ELG signals. The former was inverse-filtered and voice source parameters and formant frequencies were extracted. The amplitude quotient of the derivative of the ELG signal (AQ_{dELG}) and the contact quotient were also compared.

Results: A one-way repeated measures ANOVA revealed significant differences between vowels, for contact quotient at four pitches and for MFDR at three pitches. For other voice source parameters, differences were found at one or two pitches only. No consistent correlation was found between MFDR and the distance between F1 and its closest partial.

Conclusions: The glottal voice source tends to vary between vowels, presumably because of non-linear source-filter interaction, but the variation does not seem to be dependent upon the frequency distance between F1 and its closest partial.

Key words: Sung vowels; Inverse-filtering; Voice source; Formant frequencies; Electrolaryngograph

Introduction

According to classical singing pedagogy some vowels can be produced more easily than others at a given pitch[1]. This seems to contradict the classical source-filter theory of voice production, predicting that the glottal airflow is independent of vocal tract resonances, i.e., formants[2]. Rather, it supports the assumption that the glottal airflow is affected by the formants due to non-linear source-filter interaction[3, 4].

The theory of non-linear source-filter interaction in voice production has been developed over the last decades[5]. It predicts that, when the first formant (F1) coincides with or crosses over a lower spectrum partial, voice instabilities may occur, e.g., fundamental frequency (F0) jumps, subharmonic frequencies and changes in the amplitude of the voice source fundamental[5, 6]. Under certain conditions such feedback may facilitate vocal fold oscillation, i.e., elicit a more efficient conversion of aerodynamic to acoustic energy[3]. More specifically, the sound pressure level (SPL) of a vowel may increase by as much as 10 dB if one of the lowest harmonics is just below the first formant frequency. On the other hand, it may be weakened if one of those partials is located just above the first formant frequency[5]. This means that the interaction should be milder for male speech and greater for female and child voices. In male singing, however, an interaction should be likely to occur in and above the *passaggio*, i.e. E4 (± 330 Hz) to G4 (±400Hz)[7].

The theory of non-linear source-filter interaction has been tested and confirmed in experiments using physical models, computer simulation[8], excised larynges[9, 10] and voice source analysis in a single speaker[11]. For example, in model experiments with a simplified two-mass model connected to a straight tube, sub-harmonic vibrations and deterministic chaos were observed when F0 and F1 coincided[12]. The theory has also been tested in experiments. For example, Titze and associates (2008) had 18 subjects, none of whom had extensive vocal training, perform vocal exercise where F1 was passed by a partial. In many cases various types of F0 disturbances, such as pitch jumps, and bifurcations, were observed when a partial was close to the first formant5. Moreover, using electrolaryngography (ELG), a non-invasive tool for documenting vocal fold contact[13], differences have been observed in contacting and de-contacting events between different spoken vowels; both the open quotient and the speed quotient differed[14].

In singing, control of the vocal output is important, so uncontrolled pitch jumps and other instabilities would be totally unacceptable. One way to circumvent them would be to avoid the situation that a partial is just above F1. However, the effects of the frequency relations between F1 and its closest partial have not been measured in singers, neither with respect to the flow glottogram of different vowels, nor with respect to the ELG waveform. Hence, it seemed worthwhile (1) to compare voice source parameters between vowels and (2) to investigate whether vowel differences between such parameters could be explained by source-filter non-linear interaction. In particular, we tested if the vocal tract excitation, i.e. the maximum flow declination rate (MFDR), was greater when the frequency distance between F1 and its closest partial, henceforth *Minimum(F1-n*F0),* was positive, i.e., when F1 was just above the closest partial, and smaller when it was negative, i.e., when F1 was just below its closest partial.

Method

Eight male classically trained singers, 23–42 years old (mean 31.1, SD 6.9) with varying levels of professional expertise, volunteered as subjects (Table 1). They were asked to sing a sequence of the vowels /i, e, a, o, u/ on each of the pitches E3, G3, A3 and C4, keeping vocal loudness constant. Each task was repeated once. The pitches were given to the subjects by means of the custom made MADDE software (by Svante Granqvist, KTH, Stockholm, Sweden).

<Please insert Table 1 about here>

All recordings were made in a sound-treated studio in the Steinhardt School of Culture, Education and Human Development at New York University. A Laryngograph microprocessor (Laryngograph Ltd, London, UK) was used to record audio and ELG simultaneously. The former was picked up by a head-mounted omnidirectional electret microphone (Knowles EK3132, Knowles Corporation, Itasca, IL) placed at a mouth-to-microphone distance of approximately 15 cm. The sound level was calibrated by means of a 1 kHz sine wave, the SPL of which was measured next to the recording microphone by means of a sound level meter. The value observed was announced on the recording. Both signals were recorded using Laryngograph Speech Studio (Laryngograph Ltd, London, UK) software and stored as wav files.

The voice source was analyzed in terms of flow glottograms derived from the audio signal after integration and inverse filtering. Inverse filtering is a classical method in voice analysis[15, 16].The strategy is to eliminate the influence of the vocal tract resonance characteristics on the radiated sound. This is realized by filtering the signal by a set of filters representing the inverse of the transfer function of the vocal tract. The method offers information on both the glottal airflow waveform (flow glottogram) and on the formant frequencies and bandwidths. The accuracy is particularly high in cases where a partial is close to a formant, as illustrated in Figure 1, showing the effect of setting the F2 filter 4% above and below the correct value.

<Please insert Figure 1 about here>

Samples of the different vowels were analyzed using the custom made Decap software for inverse filtering (Svante Granqvist, KTH, Stockholm, Sweden). This program can be set to display waveform and spectrum in separate windows, as described in detail elsewhere[17]. The frequencies and bandwidths of the inverse filters are set manually and the classical equations are applied for calculating the transfer function that corresponds to the chosen combination of formant frequencies and bandwidths. The software can display the filtered voice source waveform and the spectrum in quasi-real time. The formant frequencies and bandwidths can be saved in a formant data file. Provided that the filters are correctly set, the output displays the waveform and spectrum of the transglottal airflow, also including effects of nonlinear source-filter interaction, if any. It should be noted that inverse filtering yields a representation of glottal flow, but not of glottal area, since glottal area is non-linearly related to glottal flow[11, 18]. Thus, only the effects of the vocal tract transfer function are eliminated from the input signal. The program can also display an additional signal, such as ELG or its derivative (dELG), which can be delayed so as to compensate for the time lag between the audio and the ELG signals.

For the inverse filtering, the formant frequencies and bandwidths were adjusted according to three criteria: (1) ripple-free closed phase; (2) voice source spectrum envelope as void of peaks and valleys near the formant frequencies as possible; and (3) synchrony between the positive peak of the dELG and the MFDR. The last mentioned criterion is based on the fact that vocal fold contact must cause a sudden decrease of glottal airflow[19].

As mentioned, there are reasons to expect that the sound level produced is affected by the frequency distance between F1 and the harmonic lying closest to it, the *Minimum(F1-n*F0)*. This sound level depends on the strength by which the vocal tract is excited by the voice source, and that strength is determined by MFDR. Therefore, this voice source parameter and F0 were measured using the *Glottal flow parameter measurement* tool contained in the custom made *Sopran* software (available at www.tolvan.com, last inspected September 2014). Other flow glottogram parameters measured by the same software were pulse amplitude, normalized amplitude quotient (NAQ), level difference between the first and the second harmonic of the source spectrum (H1-H2), and the closed quotient (Q_{Closed}). In addition, the contact quotient ($Q_{Contact}$) and the amplitude quotient of the derivative of ELG (AQ_{dELG}) were determined from the ELG signal. The last mentioned measure was calculated by a novel tool, developed as a script in the Sopran software by Svante Granqvist. It is defined as the smoothed ratio between the amplitude of the positive peak of the dELG signal and the amplitude of the negative peak of the dELG signal multiplied by -1. Thus, it reflects how much steeper the maximum contacting speed is than the maximum de-contacting speed. This measure is closely related to the EGG speed quotient used by Marasek (1996), which, however, was based on a linear approximation of the EGG waveform[14].

Flow glottogram and ELG measures may differ considerably between individuals; for example, the same subglottal pressure would obviously produce larger pulse amplitudes in individuals with long vocal folds than with short vocal folds, since longer vocal folds produce a larger glottal area. Therefore, all glottal parameters were converted to z-scores.

To test if the voice source parameters and the ELG measures varied significantly between vowels, a one-way within-subjects repeated-measures ANOVA was run on these z-score values, which is calculated as the difference between the value and the average of all values, divided by the standard deviation.

Results

Many flow glottograms deviated considerably from the classical form with skewed triangle-shaped pulses separated by horizontal portions that represent the closed phase. Mostly, the deviations differed between the vowels in the same sequence, as can be seen in the examples shown in Figure 2. The tilt of the closed phase varied considerably and sometimes contained a ripple or a bump that was impossible to cancel with realistic formant frequencies. In some cases, a bump appeared in the pulse, e.g., as in the vowel /o/ in the figure.

<Please insert Figure 2 about here>

The result of the ANOVA is shown in Table 2. $Q_{Contact}$ showed a significant variation for all four pitches analyzed and MFDR for three pitches. Pulse amplitude, H1-H2 and AQ_{dELG} showed a significant variation between vowels for two pitches, while NAQ and Q_{Closed} differed significantly only for one pitch.

<Please insert Table 2 about here>

The frequencies used for the inverse filters for F1 and F2 are shown in Figure 3. The subjects showed a common pattern varying with vowels in the expected manner. No systematic variation with pitch can be observed, except for F2, which was higher for the pitch C4 than for the lower pitches in some cases.

<Please insert Figure 3 about here>

According to the theory of non-linear source-filter interaction, the frequency distance between F1 and its closest partial affects the intensity that this partial has in the voice source; a formant just above the partial will boost its intensity, whereas a formant just below a partial will attenuate it[6]. The partial closest to F1 mostly has the highest amplitude in the spectrum and the SPL tends to be entirely determined by the intensity of that partial[20], [21]. The SPL, in turn, is strongly dependent on MFDR, which represents the strength with which the vocal tract is being excited by the voice source. In other words, according to the theory of non-linear source-filter interaction for voice, MFDR should vary systematically depending on *Minimum(F1-n*F0*). When Minimum(F1-n*F0) is small and positive, i.e., when F1 is just above the closest partial (inertive reactance), the amplitude of this partial should be increased. Conversely, when Minimum(F1-n*F0) is small and negative, i.e., when F1 is just below the closest partial (compliant reactance), its amplitude should be attenuated[4].

The above suggests that singers would prefer to tune F1 to a frequency just above its nearest partial. Figure 4 shows *Minimum(F1-n*F0)* for the different vowels and pitches. Positive values refer to cases where the F1 was higher than its closest partial, and vice versa. If the singers preferred to place F1 just above its closest partial, there would be a greater number of positive than negative values. Moreover, as the effect would be greatest when *Minimum(F1-n*F0)* is both positive and small, there should be a great number of low positive values. Such effects were not found; the number of positive values was greater than the number of negative values in five singers, and lower in three singers. Furthermore, few values were positive and small. and there was no clear difference between cases where the second and the third partial was closest to F1. Thus, there seemed to be no preference among these singers to tune F1 to a frequency just above a partial.

<Please insert Figure 4 about here>

Figure 5 shows each singer's MFDR, in l/s^2, for the different vowels and pitches. For all singers, MFDR differed considerably between vowels sung on a given pitch. The variation was particularly great for singer 3 and small for singer 8. In many cases MFDR was lower for /i/ and /u/ than for /a/.

<Please insert Figure 5 about here>

As mentioned, the theory of non-linear source-filter interaction predicts that vowels with a low positive *Minimum(F1-n*F0)* will be associated with higher MFDR values than vowels with a low negative *Minimum(F1-n*F0).* In both cases, the effect should decrease when the frequency separation between F1 and the partial increases. This was tested by plotting MFDR as function of *Minimum(F1-n*F0)*, see Figure 6. In each panel of the figure the correlation coefficients between MFDR and the *Minimum(F1-n*F0)* are listed for negative values and positive values of *Minimum(F1-n*F0)* separately. Thus, in each panel the midline represents the case when the formant coincides with a partial.

<Please insert Figure 6 about here>

In some cases, the values to the left of the midline (open circles) are lower than most of those to the right of the midline (filled circles). This means that, in these cases, MFDR tended to be lower when the formant was below its closest partial. This is in accordance with the theory of non-linear source-filter interaction. This theory further predicts that the closer the partial is to the formant, the stronger the effects should be; in other words, when the formant is higher than its closest partial, i.e., for *Minimum(F1-n*F0)*>0, the highest MFDR values should occur close to the midline in the graph. Conversely, when the formant is lower than its closest partial, i.e., for *Minimum(F1-n*F0)*<0, the MFDR values close to the midline should be low. This means the correlations both to the left and to the right of the midline should be negative, so one would expect the MFDR values to the left of the midline in each panel to decrease with decreasing distance to the midline. Moreover, the values to the right of this line should be high close to the midline and decrease with increasing distance from this line. Only singers 3 and 5 showed a clear negative correlation for negative values of *Minimum(F1-n*F0)*. For singers 3, and 7 MFDR increased with increasing positive *Minimum(F1-n*F0)*. For positive values, none of the singers showed a negative correlation. These findings suggest that non-linear source-filter interaction did not explain why MFDR differed significantly between vowels.

One might assume that AQ_{dELG}, i.e., the speed of vocal fold contact, would be greater when F1 is just above its closest partial, i.e., when *Minimum(F1-n*F0)* is small and positive. This assumption was tested by analyzing the correlation between the two. The result is shown in Figure 7. The correlation coefficients for negative and positive values of *Minimum(F1-n*F0)* are shown in the top of the panels. There is an indication of some relationship between AQ_{dELG} and *Minimum(F1-n*F0)* for Singers 4, 5 and 6. However, these correlations differ in sign, and refer to negative values of *Minimum(F1-n*F0)* for Singers 4 and 5, and to positive values for Singer 6. Hence, although AQ_{dELG} differed significantly between vowels at two pitches, it does not seem to be related to the distance between F1 and its closest partial.

<Please insert Figure 7 about here>

Discussion

The present investigation has shown voice source differences between vowels sung by male professional singers. Significant differences were found for all flow glottogram parameters analyzed, at least for one of the four pitches examined. With respect to MFDR and $Q_{Contact}$, significant differences were found for three and four pitches, respectively. However, this variation seemed independent of *Minimum(F1-n*F0)*. A relevant question then is what may have caused the variation?

Nasalization seems like a possible reason. It has been shown that many singers sing with a more or less open velo-pharyngeal port[22]. This complicates the vocal tract transfer function, such that inverse filtering becomes problematic; for such filtering to yield accurate results, the vocal tract transfer function needs to be accurately modelled, which is difficult for nasalized vowels[23]. On the other hand, according to Gobl and Mashie (2013), the effect of a velopharyngeal opening with a cross sectional area ≤1 cm^2 has a minor effect on flow glottogram parameters, except for the very final part of the closing phase, the so-called return phase[24]. This part of the flow glottogram has a main effect on the amplitudes of the higher source spectrum partials; the shorter the return phase, the stronger the high frequency partials in the voice source. However, the return phase cannot affect MFDR appreciably.

Unexpectedly, flow glottograms have recently been found to be quite sensitive to sound reflections in the recording room; depending on the distance to the reflecting object, they may cause disturbances such as a tilt and also ripple in the closed phase or a dent in the flow pulse (Svante Granqvist, personal communication). Many examples of such disturbances were found in the flow glottograms, even though the recordings were made in a sound treated studio and the microphone distance to the mouth was only approximately 15 cm. A much shorter microphone distance would have reduced or eliminated the influence of room reflections. However, this would have caused clipping of the audio signal by the equipment used in the experiment.

The bump in the pulse, illustrated in the case of the vowel /o/ in Figure 2, shows a striking similarity with the flow glottogram shape discussed by Gunnar Fant (1986)[18] and by Martin Rothenberg[25]. Both ascribed this effect to source-filter interaction. In the case shown in the figure, F1 and F2 were 517 Hz and 1009 Hz, respectively, and F0 was 168 Hz. Hence, the third partial was 3*168 = 504 Hz, i.e., just 13 Hz above F1; the sixth partial was 6*168 = 1008 Hz, i.e., just 1Hz from F2. In this case, then, it seems likely that the bump in the pulse was caused by source-filter interaction.

Lim and associates (2006) studied the relationship between jaw opening and the EGG speed quotient in different vowels pronounced by speakers. They found that this quotient was significantly lower in the vowel /u/ than in the vowels /a, e, i, o/[26]. The speed quotient should be closely related to the AQ_{dELG}. We found significant variation of this parameter between vowels for two pitches. Thus, like the results reported by Lim & al, our findings indicate that the contacting speed of the vocal folds is not the same between vowels.

The subjects used in the present study were classically trained singers. It is often advantageous to investigate voice physiology and acoustics in singer subjects, since they have particularly good control over their voices. This would imply that they minimize the influence of random factors on phonation. According to the theory of non-linear source-filter interaction, a negative *Minimum(F1-n*F0)* that is close to zero should expose the singer to the risk of instabilities and uncontrolled variation of MFDR, and hence vocal loudness. For example, when performing unfamiliar vocal tasks such as wide glissandos, untrained voices have been found to produce pitch jumps and other uncontrolled vocal events when a partial is passing F1[27]. One would then expect that trained singers avoid the situation that F1 is just below its closest partial. However, in this study the singers did not avoid this situation and yet they did not have pitch instabilities. Apparently, trained singers learn how to circumvent such problems.

According to Titze and associates (2008), the non-linear source-filter interaction is physiologically related to the diameter of the epilaryngeal tube. It has been shown in non-singers that a strong coupling occurs, when the impedances at the glottis and at the epilaryngeal tube are similar[5]. Thus, one way to avoid instabilities may be by “structured practice in the instability region, with the intent of developing muscle patterns that counteract the instabilities.” (in Titze et al., 2008; pp. 1914).

Summarizing, while no pitch instabilities were found, several flow glottogram and ELG differences were observed between vowels. Such differences may be caused by source-filter interaction but, according to these findings, they do not seem to be due to a specific formant/partial relationship. These differences are seemingly much less apparent than pitch instabilities, making them more tolerable, even in professional singing.

Conclusions

Our results have shown that there are substantial flow glottogram and ELG differences between vowels sung in a sequence by professional classically trained singers. While these differences may occur due to non-linear source-filter interaction, our analyses failed to support the assumption that the variation of MFDR was related to the *Minimum(F1-n*F0)*. Thus, much of the cause for the variation of the voice source between vowels remains an open question.

Acknowledgments

The authors gratefully acknowledge the kind cooperation of the singers who volunteered to participate. This work was presented at the Annual Symposium Care of the professional Voice, Philadelphia June 2014. The authors are indebted to the Schering Health Care Ltd. & Bayer Portugal for providing the means to acquire the equipment.

References

1. Coffin B. *Overtones of Bel Canto*. New Brunswick, NJ: Scarecrow Press; 1980.
2. Fant G. *Acoustic Theory of Speech Production*. 2nd ed ed. The Hague, The Netherlands: Mouton; 1960.
3. Titze I. A theoretical study of F0-F1 interaction with application to resonant speaking and singing voice. *J Voice*. 2004;18:292–298.
4. Titze I. Nonlinear source-filter coupling in phonation: theory. *J Acoust Soc Am*. 2008;123:2733–2749.
5. Titze I, Riede T, Popolo P. Nonlinear source–filter coupling in phonation: Vocal exercises. *J Acoust Soc Am*. 2008;123:1902–1915.
6. Titze IR, Worley AS. Modeling source-filter interaction in belting and high-pitched operatic male singing. *Journal of the Acoustical Society of America*. 2009;126:1530–1540.
7. Doscher B. *The Functional Unity of the Singing Voice*. London, UK: The Scarecrow Press, Inc.; 1994.
8. Alipour F, Montequin D, Tayama N. Aerodynamic profiles of a hemilarynx with a vocal tract. *Ann. Otol. Rhinol. Laryngol.* . 2001;110:550–555.
9. Chan R, Titze I. Dependence of phonation threshold pressure on vocal tract acoustics and vocal fold tissue mechanics. *J. Acoust. Soc. Am.* 2006;119:2351–2362.
10. Zhang Z, Neubauer J, Berry D. The influence of subglottal acoustics on laboratory models of phonation. *J Acoust. Soc. Am.* 2006;120:1558–1569.
11. Rothenberg M. Acoustic interaction between the glottal source and the vocal tract. In: Stevens K, Hirano M, eds. *Vocal Fold Physiology*. Tokyo: University of Tokyo Press; 1981:305–328.
12. Haralambos A, Tecumseh F, HP. H. Voice Instabilities due to Source-Tract Interactions. *ACTA ACUSTICA united with ACUSTICA*. 2006;92:468-475.
13. Baken R, Orlikoff R. *Clinical Measurement of Speech and Voice.* 2nd Edition ed. San Diego: Singular Publishing Group Thomson Learning; 2000.
14. Marasek K. Glottal correlates of the word stress and the tense/lax opposition in the German vowels. *Proceedings of ICSLP-96* 1996:1573–1577.
15. Fant G. A new anti-resonance circuit for inverse filtering. *Speech research summary report STL--QPSR*. 1961;4.
16. Alku P. Glottal wave analysis with pitch synchronous iterative adaptive inverse filtering. *Speech Communication*. 1992;11.
17. Sundberg J, Lã F, Gill B. Formant tuning strategies in professional male opera singers. *Journal of Voice*. 2013;27:278-288.
18. Fant G. Glottal flow: Models and interaction. *J. Phonetics* 1986;14:393–399.
19. Lã FMB, Sundberg J. Contact Quotient Versus Closed Quotient: A Comparative Study on Professional Male Singers. *Journal of Voice*. 2014;In Press.
20. Gramming P, Sundberg J. Spectrum factors relevant to phonetogram measurement. *J Acoust Soc Amer* 1988;83:2352-2360.

21. Titze I. Acoustic interpretation of the voice range profile (phonetogram). *J Speech and hearing Research*. 1991;35:21-34.

22. Birch P, Gümoes B, Stavad H, Prytz S, Björkner E, Sundberg J. Velum behavior in professional classic operatic singing. *J Voice*. 2002;16:61-71.

23. Båvegård M, Fant G, Gauffin J, Liljencrants J. Vocal tract sweeptone data and model simulations of vowels, laterals and nasals. *Quarterly Progress and Status Report, KTH, Stockholm*. 1993;34:43-76.

24. Gobl C, Mahshi J. Inverse Filtering of Nasalized Vowels Using Synthesized Speech. *Journal of Voice*. 2013;27:155-169.

25. Rothenberg M. Cosi fan tutte and what it means or Nonlinear source-tract acoustic interaction in the soprano voice and some implications for the definition of vocal efficiency. In: T.B. B, Sasaki C, Harris KS, eds. *Laryngeal Function in Phonation and Respiration (Proc. Vocal Fold Physiology Conf. 1985)*. San Diego: Singular Publing Group; 1987:254-269.

26. Lim M, Lin E, Bones P. Vowel Effect on Glottal Parameters and the Magnitude of Jaw Opening. *Journal of Voice*. 2004;20:46–54.

27. Henrich N, d'Alessandro C, Doval B, Castellengo M. Glottal open quotient in singing: Measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency,. *Journal of the Acoustical Society of America*. 2005;117:1417-1430.

TABLE CAPTIONS

Table 1. Participating singers' ages, classifications and experiences.

Table 2. Results of a one-way within-subjects repeated-measures ANOVA run on the z-score values of the flow glottogram and ELG parameters for each of the four pitches analyzed. The columns list the median (Mdn) and inter-quartile range (IQR) for the different vowels. Statistically significant differences between vowels (Friedman test, $p < 0.05$) are in bold and marked +.

FIGURE CAPTIONS

Figure 1. Example of the effects on the flow glottogram and source spectrum (upper and lower panels) of mistuning the F1 inverse filter by -20Hz and +20Hz (left and right panels, respectively). The middle panel represents the result of a correct filter setting.

Figure 2. Examples of audio signal, flow glottogram, and dELG (top, middle and bottom curves) of the indicated vowels sung in the same vowel sequence by singer 1 on the pitch E3 ($F0 \approx 165Hz$).

Figure 3. F1 and F2 used in the inverse filtering of the vowels sung at the indicated pitches by the singers.

Figure 4. Frequency distance between F1 and its closest partial, Min(F1-n*F0), for the different vowels sung at the indicated pitches by the eight singers.

Figure 5. MFDR for the different vowels sung by the subjects at the indicated pitches

Figure 6. MFDR as function of the Minimum(F1-n*F0) At the bottom of each panel, the left and right values correspond to the linear correlations for negative and positive values (open and filled circles, respectively) of Minimum (F1-n*F0).

Figure 7. AQdELG as function of Minimum(F1-n*F0). In each panel the left and right values at the top show the correlations for negative and positive values of Minimum(F1-n*F0), open and filled circles, respectively.