Spectrum effects of a velopharyngeal opening in singing

Brian P. Gill
Jacob School of Music, University of Indiana Bloomington, USA

Jessica Lee
School of Music, New York University, USA

Filipa M.B. Lã
Center for Social Sciences, University of Coimbra, Portugal and Institute of Interdisciplinary Research, University of Coimbra, Portugal

Johan Sundberg
Department of Speech Music Hearing, School of Electrical Engineering and Computer Science, KTH Stockholm Sweden and University College of Music Education Stockholm, Sweden

Corresponding author: Johan Sundberg, Department of Speech Music Hearing, School of Electrical Engineering and Computer Science, KTH Stockholm Sweden, email jsu@kth.se

Biography
Johan Sundberg, PhD in Musicology 1966, awarded a personal chair in Music Acoustics at KTH Stockholm in 1979, member of the Swedish Royal Academy of Music. His main research areas are the function, acoustics, and expressivity of the singing voice and the theory of music performance. His book Röstlära (The science of the Singing Voice) has been translated into English, German, Japanese, and Portuguese.

'Declarations of interest: none'.

Abstract

The question whether or not velo-pharyngeal opening is advantageous in singing has been discussed for a very long time among teachers of singing. The present investigation analyzes the acoustic consequences of a large, a narrow, and a non-existent VPO. A divided flow mask (nasal and oral) connected to flow transducers recorded the nasal and oral DC flows in four female and four male classically trained singers while they sang vowel sequences at different pitches under these three experimental conditions. Acoustic effects were analyzed in three long-term-average spectra parameters: (i) the sound level at the fundamental frequency, (ii) the level of the highest peak below 1 kHz, and (iii) the level of the highest peak in the 2-4 kHz region. For a narrow VPO, an increase of the level of the highest peak in the 2-4 kHz region was observed. As this peak is an essential voice component in the classical singing tradition, a narrow VPO seems beneficial in this type of singing.

Keywords
Nasalization, long-term-average spectrum, nasal airflow, vowel sequence, nasal sound amplitude

1. 1. Introduction

Nasality, along with the inclusion or avoidance of a velopharyngeal opening (VPO) during singing, has long been a contentious subject. Some teachers of singing strongly recommend a singer avoid singing with a VPO [1, 2, 3], while other pedagogues argue that a narrow VPO is advantageous for the voice timbre [4, 5].

Attempts to measure the VPO in singers have been reported earlier. Pershall and Boone [6], as well as Yanagisawa and associates [7], using endoscopy above and below the velum, found no VPO in their male and female singer subjects. Some other studies have used a combination of nasal fiberscope and nasal and oral airflow measurements, and the results have shown that many singers sing with a VPO [8, 9]. Similar findings have been reported in other investigations [5, 10].

Several experiments have been carried out to investigate perceptual effects of a VPO as well. Vennard (1967) [11] had the noses and nasopharynges stuffed with gauze and the maxillary sinuses filled with water in five male singers. The resultant spectra were not different than under normal singing conditions and expert listeners were unable to distinguish between the different conditions. These results thus suggest that these subjects sang with a closed velopharyngeal port. Rousselot (1924) [12] reported an experiment carried out by Passavant in which tubes of different sizes were inserted into the velo-pharyngeal port; a 12 mm^2 VPO had no effect on the quality of the vowel, whereas a VPO of $28mm^2$ resulted in a nasal sound.

With regard to the acoustic effects, a nasal vowel quality has been found to result from both a decrease in amplitude of the first vowel formant and a widening of its bandwidth [13]. In addition, a spectrum peak near 250 Hz is typically observed in nasalized vowels [14]. Fant (1960) [15] studied the effects of a VPO using an electric analogue of the vocal tract. He showed that nasalization produced a pole-zero pair in the transfer function. This pole-zero pair appeared in the vicinity of the first formant, which could attenuate and even split this formant. Also Feng and Castelli [10] used an electrical analogue with a pharyngeal port connected to two branches, one for the oral and one for the nasal cavity [16]. They varied the coupling between the branches from complete to zero, thus modeling a velo-pharyngeal port ranging from wide open to completely closed. The dip in the transfer functions varied in frequency but disappeared when the port was very narrow.

Tanner and associates [17] combined measures of VPO size with recordings of nasal and oral flow in sopranos. They observed a VPO in these singers and assumed that such an opening might reduce vocal loading, thus being advantageous for the classically trained singing voice.

Fujimura and Lindquist [18] analyzed sweep-tone measurements of the vocal tract of speakers who silently articulated vowels with a closed glottis. They found that the frequency of the transfer function dip caused by nasalization depended on the vowel. Their data were reanalyzed by Båvegård and associates [19] who also used a computerized model for studying the effect of VPOs of different sizes on the transfer function of the vowel /a/. They noted that even a narrow VPO produced a dip near 600Hz, which attenuated the peak level near the first formant, but failed to affect the levels of the third and higher formants.

Summarizing, much research has been spent on the presence of a VPO in singers. However, despite evidence of VPO of different shapes and sizes, it is still controversial whether singers benefit from it. For example, the timbral effects of singing with a VPO are unclear. The present study was designed to examine the acoustic effects of using three degrees of VPO: none, narrow, and wide. The degrees of a VPO were documented in terms of the amplitudes of nasal and oral DC and AC airflow. The associated acoustic effects **were** examined in long-term average spectra (LTAS) of vowel sequences.

2. Method

The recording setup is shown in Figure 1. The audio signal was picked up by a head-mounted omnidirectional electret microphone (Knowles EK3132, Knowles Headquarters, Itasca, IL), which was fastened to the subject's shirt at a measured distance from the mouth opening. Nasal and oral flows were captured by means of a flow mask (Glottal Enterprises, Syracuse, NY) that the singers held well sealed to their faces. A plastic divider plate at the level of the upper lip separated the oral and the nasal airflow. The audio and the airflow signals were all digitized with a sampling frequency 48000Hz and recorded in separate channels by the SpeechStudio software (Laryngograph Ltd, London UK). The mask produced a slight dampening of the formants. This effect, however, was negligible, as it was present in all experimental conditions. In the recordings, also an electroglottograph signal was recorded, which, was not used in the present study.

< *Please insert Figure 1 about here* >

Eight singers, between the ages of 27 and 39 years old, all graduate students at New York University, volunteered to participate in this experiment: four males - two tenors (T1 and T2) and two baritones (B1 and B2) - and four females - all sopranos (S1 – S4). Their task was to sing sequences of six vowels [i-e-æ-a-o-u] on pitches in the lower and middle parts of their ranges as well as in their upper *passaggio* (see Table 1). The sequences were sung with three different VPO conditions: *None*, *Slight* (narrow) and *Much* (wide). During the recordings, nasal flow was displayed on the PC screen and monitored by co-author BG. The singers repeated each version of the vowel sequence at least once, and more, if the nasal flow revealed that the experimental condition was not appropriately realized.

< *Please insert Table 1 about here*>

The audio signal was calibrated by recording a 1000Hz sine wave with sound pressure level (SPL) measured at the recording microphone by means of a sound level meter. For each participant, this SPL value and the microphone distance were announced in the recording. The flow signals were calibrated by injecting 140ml of air into the pressure transducer. This was obtained by emptying a calibrated syringe into the respective flow recording system. Calibration of the flow signals was performed using the calibration tool of the **custom** made *Sopran* software (by Svante Granqvist, www.tolvan.com). It measured the average of the recorded deflections in the airflow tracks and multiplied them with the duration of the airflow injections.

3. Analysis

The degree of nasalization can be expected to affect the nasal DC flow and the amplitude of the acoustic signal radiated from the nostrils; therefore, both these parameters were analyzed. Measurements were made in the final take of the vowel sequences for each of the three conditions. Using the Soundswell signal workstation (Neovius Data, Lidingö Sweden) , the corresponding audio signal was analyzed with respect to the long-term-average spectrum (LTAS), using a 128 points FFT. This resulted in an analysis bandwidth equivalent to 375 Hz and smooth LTAS curves showing level values at multiples of 93.75Hz yielded. Three measures were taken from the LTAS curves: (i) the level at the fundamental frequency f_o; (ii) the maximum LTAS level between 0 and 1 KHz; and (iii) the maximum LTAS level between 2 and 4 KHz. These measures were selected on the basis of previous observations that showed effects of different degrees of nasalization in these three frequency regions [4].

Using the Statistical Package for Social Sciences (SPSS), version 21 (IBM, 2011), a Shapiro-Wilk test was applied to test the normality of data for all variables and pitches. In addition, a one-way between-subjects ANOVA was carried out to test whether significant differences existed between the three experimental conditions for each of the independent variables that showed a normal distribution: oral and nasal DC flow amplitudes, root mean square (RMS) amplitudes of the nasal and oral flow signals, and the maximum LTAS level between 0 and1 kHz. A non-parametric Kruskal-Wallis test was carried out for the two remaining variables that showed a non-normal distribution: the LTAS level at F0 and the maximum LTAS level between 2 and 4 kHz.

4. Results

As mentioned, a VPO should generate a nasal airstream and also radiate sound from the nostrils. Basically, the amplitudes of both these signals should increase with increasing VPO. Figure 2 compares the difference between these signals for the three conditions: *None*, *Slight* and *Much*. A strong correlation would not be expected, as the flow depends on the resistance in the nasal tract while the amplitude of the radiated sound depends both on this resistance and the nasal resonance characteristics. Nevertheless, a clear correlation between the signals was observed, with the coefficient of determination amounting to 0.421. Figure 2 also shows that the three conditions were associated with the expected systematic variation of the two nasal airflow parameters.

< Please insert Figure 2 about here >

Figure 3 shows, for each condition, the amplitude of the nasal DC flow signal averaged across all subjects and pitches. The average differed significantly between the conditions, being smallest for the *None* condition and greatest in the *Much* condition, thus verifying that the subjects followed the instruction.

< Please insert Figure 3 about here >

Figure 4 compares the LTAS level at the f_0 for the three conditions. The left and right panels show the effects observed in the *None* and *Slight* conditions and in the *None* and *Much* conditions, respectively. The Kruskal-Wallis test revealed that the level at the fundamental was not significantly different between the *None* (median = - 23.00 dB; IQ = 10.78) and the *Slight* conditions (median = - 24.49 dB; IQ = 12.45), while it was significantly weaker in the *Much* condition (median = - 30. 02; IQ = 8.95) [χ^2 (2) = 15.701; p < 0.001].

< Please insert Figure 4 about here >

The corresponding results for the maximum LTAS level below 1 kHz are shown in Figure 5. Typically, this level is almost identical with the equivalent sound level**.** The left and right panels show the effects of the *Much* and the *Slight* conditions, both versus the *None* condition. *Slight* did not show any marked difference from *None* (mean differences, - 6.21 dB for *Slight* and - 6.29 dB for *None*). However, *Much* showed a clear effect. According to the trendline, the difference was marginal for peaks of low amplitudes but great for peaks of higher amplitudes; a -10 dB in the *None* condition corresponded to a peak of less than -20 dB in the *Much* condition. These differences were confirmed by the results of the one-way between-subjects ANOVA [χ^2 (2) = 16.232 dB; p < 0.000].

< Please insert Figure 5 about here >

Figure 6 shows the corresponding results for the maximum LTAS in the high frequency range, between 2 and 4 kHz. This part of the spectrum is crucial for allowing a voice to be heard over a loud orchestral accompaniment [20]. As compared with the *None* condition, this LTAS level increased significantly both in the *Much* and in the *Slight* conditions [χ^2 (2) = 16.467 dB; p < 0.000]. However, the results differed between the female and male subjects. For the female subjects, the increase exceeded 7dB in 17 out of the 24 vowel sequences for the *Slight* condition, and only 10 for the *Much* condition. For the male subjects, no corresponding difference could be seen, except for one singer, who produced a 7dB stronger amplitude for the *Much* as compared to the *None* condition.

< Please insert Figure 6 about here >

5. Discussion

The success of this experiment was dependent upon the participants' ability to systematically vary their VPO. The statistical analysis of the nasal DC flows confirmed that they were able to do so in accordance with the instructions.

The results were based on LTAS averaged across different pitches and vowels. Given the chosen analysis bandwith (375 Hz), such LTAS curves will have very different appearances, depending on the pitch; for low pitches, the curve will be continuous while for high pitches, each harmonic produces a separate peak. However, this did not affect the results, since only the maximum LTAS levels in the three frequency bands were measured. As the chosen FFT analysis yielded LTAS level values at multiples of 93.75Hz, the level at f_0 could be reliably measured even at the lowest pitches.

The main findings of this investigation can be summarized as follows. As compared with vowel sequences produced with no VPO, a wide VPO was associated with a weaker fundamental, a lower level of the highest LTAS peak below 1 kHz, and, for some female singers, a higher peak in the 2-4 kHz range. The effect of a narrow VPO was almost nil in the level of the fundamental and in the level of the highest peak below 1 kHz. By contrast, the highest peak in the 2-4 kHz range increased, particularly for the female singers, as observed by Tanner and associates [17].

As the *Much* condition would correspond to nasalization, the results are in agreement with earlier research with regard to the decrease of LTAS levels below 1 kH, as was mentioned in the introduction. This effect was not observed in the *Slight* condition, which seems in accordance with the findings of Feng and Castelli [10, 16] who observed no effect on the transfer function of their model for a low degree of coupling between the nasal and vocal tracts.

Also the level increase of the highest LTAS peak in the 2-4 kHz range associated in the *Slight* condition is in accordance with the earlier observation that the spectrum of a "slightly nasalized" vowel /a/ showed a considerably increased level in this frequency range[4].

It is commonly assumed that, even a slight VPO results in a nasal quality of vowel sounds and should therefore be avoided in training singers[1, 3]. However, Birch and associates showed that professional opera singers, who had been found to sing the vowel /a/ with a VPO, did not produce a voice timbre that was perceived as nasal by a panel of singing teachers [21].

The boost in the 2 to 4 kHz part of the spectrum is somewhat similar to the effect of an increase in vocal loudness, i.e. increase of subglottal pressure P_{Sub} [22, 23]. A higher P_{Sub} can be expected to produce a greater total airflow. However, the average airflow was nearly the same for both the female and the male singers in the *None* and *Slight* conditions, while it was mostly greater in the *Much* condition. Another possibility is a neural connection between the velum and the larynx; a frequent observation in laryngoscopic exams is that the space between the ventricular folds typically widens during the production of the consonant /m/ (Per-Åke Lindestad, personal communication). It may also be relevant that the auditory feedback of one's own voice, caused by the VPO, is affected. It would be interesting to analyze the voice source, e.g. in terms of the EGG signal for the *None*, *Slight*, and *Much* conditions. In any event, the reason underlying why and under what conditions a narrow VPO is associated with a boost of high frequency partials merits further research.

As mentioned, an increase in vocal loudness is normally produced by increasing P_{Sub} [24]. A 10 dB increase is typically associated with a 15 dB increase of the LTAS level in the 2-4 kHz frequency range. This has been found in both untrained and trained voices [22, 23]. The present results showed that also a narrow VPO increases the LTAS level in this frequency range. This suggests that a slight VPO can serve the purpose of optimizing the ratio between sound level produced and P_{Sub}.

Summarizing the present study has found that a narrow VPO increases the average intensity in the 2- 4 kHz range of a vowel sequence without reducing the overall sound pressure level of the sequence, or the average level of the fundamental. This should be favorable with regard to spectral balance. Hence, a narrow VPO could be regarded as advantageous for singers.

6. Conclusions

The present findings seem to shed some light on the controversy regarding whether a VPO is advantageous or not for training singers in the classical tradition. According to our findings, a wide VPO seems disadvantageous in classical singing, as it decreases the overall sound level and produces a nasalized vowel quality. No VPO also appears to be disadvantageous, as it reduces in the intensity of higher partials in the crucial area of the singers' formant cluster. By contrast, a slight VPO can be advantageous, since it seems to offer singers the possibility to increase the levels of the higher spectrum partials without involving an unnecessary increase of vocal loudness.

Acknowledgements

The authors are indebted to the kind cooperation of the singer subjects.

References

[**1**] Austin, S.F.. Movement of the Velum During Speech and Classically Trained Singers. *Jounal of Voice.* 11 (1997) 212-221.

[**2**] Austin, S.F. From Song to Speech: Nasal Resonance-Fact or Fiction? *Journal of Singing.* 57 (2000) 33-40.

[**3**] McCoy, S. The Seduction of Nasality. *Journal of Singing.* 64:579-582.

[**4**] Sundberg, J., Birch, P., Gümoes, B., Stavad, H., Prytz, S., Karle, A., 2007. Experimental Findings on the Nasal Tract Resonator in Singing. *Journal of Voice.* 21 (2008) 127–137.

[**5**] Perna, N. Nasalance and the Tenor Passaggio. *Journal of Singing.* 70 (2014) 403-410.

[**6**] Pershall, K. E., Boone, D. R. Supraglottic contribution to voice quality. *Journal of Voice.* 1 (1987) 186-190.

[**7**] Yanagisawa, E., Kmucha, S. T., Estill, J. Role of the Soft Palate in Laryngeal Functions and Selected Voice Qualities: simultaneous velolaryngeal videoendoscopy. *Annals of Otology, Rhinology & Laryngology* 99 (1990) 18-28.

[**8**] Gramming, P., Nord, L., Sundberg, J., Elliot, N.Does the nose resonate during singing? *STL-QPSR.* 1993;3&4 (1993) 35-42.

[**9**] Birch, P., Gümoes, B., Stavad, H., Prytz, S., Björkner, E., Sundberg, J. Velum Behavior in Professional Classic Operatic Singing. *Journal of Voice.* 16 (2002) 61-71.

[**10**] McIver, W., Miller, R.J.S. A brief study of nasality in singing. *Journal of Singing.* 52 (1996) 21-26.

[**11**] Vennard, W. *Singing: the mechanism and the technique.* New York: Carl Fischer; 1967.

[**12**] Rousselot, P.J. *Principes de Phonetique Experimentale*. Paris: H. Didier; 1924.

[**13**] Delattre, P. Les attributs acoustiques de la nasalité vocalique et consonantique. *Studia Linguistica.*;8 (1954) 103-109.

[**14**] Maeda, S. Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. In: Krakow MKHaRA, ed. *Phonetics and Phonology.* Vol 5 (1993) Academic Press;:147-167.

[**15**] Fant, G. *Acoustic theory of speech production*. The Hague: Mouton; 1960.

[**16**] Feng, G., Castelli, E. Some acoustic features of nasal and nasalized vowels: A target for vowel nasalization. *The Journal of the Acoustic Society of American.*;99 (1996) 3694-3706.

[**17**] Tanner, K., Roy, N., Merrill, R.M., Power, D. Velopharyngeal port status during classical singing. *Journal of Speech, Language, and Hearing Research*;48 (2005) 1311-1324.

[**18**] Fujimura, O., Lindqvist (Gauffin), J. Sweep-Tone Measurements of Vocal-Tract Characteristics. *The Journal of the Acoustical Society of America.*;49 (1971) 541–558.

[**19**] Båvegård, M., Fant, G., Gauffin, J., Liljencrants, J. Vocal tract sweeptone data and model simulations of vowels, laterals and nasals. *STL-QPSR* 1993;3&4 (1993) 43-76.

[**20**] Sundberg, J. *The Science of the Singing Voice*. Dekalb, Illinois: Northern Illinois University Press; 1987.

[**21**] Birch, P., Gümoes, B., Prytz, S., Karle, A., Stavad, H., Sundberg, J. Effects of a velopharyngeal opening on the sound transfer characteristics of the vowel [a]. *TMH-QPSR.* 2002: 43 (2002) 9-15.

[**22**] Sjölander, P, Sundberg, J. Spectrum effects of subglottal pressure variation in professional baritone singers. *The Journal of the Acoustical Society of America.*;115 (2004) 1270–1273.

[**23**] Nordenberg, M., Sundberg, J. Effect on LTAS of vocal loudness variation. *Logopedics Phoniatrics Vocology.*;29 (2004) 183–191.

[**24**] Björklund, S., Sundberg, J. Relationship Between Subglottal Pressure and Sound Pressure Level in Untrained Voices. *Journal of Voice.*;30 (2016) 15–20. https://doi:10.1016/j.jvoice.2015.03.006.

Tables

Table 1. Pitches sung by each participant.

Participants	Pitches
S1 – S4	E4, G4, Bb5, Db5, and E5
T1	G3, Bb3, Db4, E4, and G4
T2	G3, Bb3, Db4, E4, G4, and B4
B1	G2, E3, G3, Bb3, Db4, and E4
B2	D3, F3, Ab3, B3, and D4

Figure captions

Figure 1. Block diagram of the experimental setup: electroglottograph (EGG); electrolaryngograph microprocessor (ELG); MS-110 computer interface for the nasal and oral flow transducers; personal computer (PC).

Figure 2. Comparison of flow and audio signals in terms of the difference between nasal and oral DC flows as function of the difference between the sound levels of these flows. Circles, diamonds and triangles refer to the nasalization conditions *None*, *Slight*, and *Much*, respectively. The dashed line and the equation refer to the trendline.

Figure 3. Average amplitude of the nasal DC flow in the indicated conditions. The circles represent the mean and the bars the standard deviation.

Figure 4. Comparisons of effects on the LTAS level at f_0 of the *None* and *Slight* conditions, and of the *None* and *Much* conditions (left and right panels, respectively). The dashed lines and the equations refer to the trend lines. The dotted lines represent the one-to-one relationship.

Figure 5. Comparisons of effects on the maximum LTAS level below 1 kHz for the *None* and *Slight* conditions, and for the *None* and *Much* conditions (left and right panels, respectively). The dashed lines and the equations refer to the trend lines. The dotted lines represent the one-to-one relationship.

Figure 6. Comparisons of effects on the maximum LTAS level between 2 and 4 kHz for the *None* and *Slight* conditions, and for the *None* and *Much* conditions (left and right panels, respectively). The upper and lower panels show results for the female and the male subjects. Symbols refer to subjects. The dotted lines represent the one-to-one relationship.

Is there an effect on the mask on the recording? Sentence added on page

Which is the rationale for the choice of the acoustical parameters? Reason given on page

Which is the rationale for the choice of broadband LTAS instead of a narrow band LTAS? You are looking for the amplitude of specific harmonic components/ fundamental frequency; I would have used a narrow band filter for the LTAS. This is commented upon in the second paragraph in the Discussion section. Also, we added information on the LTAS analysis on page .

Can you discuss more if there are gender and or voice type effect on your results? The effect of gender is presented in the final paragraph of the ***Results*** section **COMMENTS?**

Could you explain better the figures and place the labels on the axes? Don’t understand