

The Singer's Voice Range Profile: Female Professional Opera Soloists

ANICK LAMARCHE^a, STEN TERNSTRÖM^a, PETER PABON^b

^a Kungliga Tekniska Högskolan, School of Computer Science and Communication,
Department of Speech, Music and Hearing

^b Royal Conservatory, The Hague / University Utrecht / Voice Quality Systems

Abstract

This work concerns the collection of 30 Voice Range Profiles (VRPs) of female operatic voice. **Objectives:** We address the questions: Is there a need for a singer's protocol in VRP acquisition? Are physiological measurements sufficient or should the measurement of performance capabilities also be included? Can we address the female singing voice in general or is there a case for categorizing voices when studying phonetographic data? **Method:** Subjects performed a series of structured tasks involving both standard speech voice protocols and additional singing tasks. Singers also completed an extensive questionnaire. **Results:** Physiological VRPs differ from performance VRPs. Two new VRP metrics: the voice area above a defined level threshold, and the dynamic range independent from F_0 , were found to be useful in the analysis of singer VRP's. Task design had no effect on performance VRP outcomes. Voice category differences were mainly attributable to phonation frequency based information. **Conclusion:** Results support the clinical importance of addressing the vocal instrument as it is used in performance. Equally important is the elaboration of a protocol suitable for the singing voice. The given context and instructions can be more important than task design for performance VRPs. Yet, for physiological VRP recordings, task design remains critical. Both types of VRPs are suggested for a singer's voice evaluation.

Introduction

The Voice Range Profile (VRP) or phonetogram, is an increasingly popular clinical tool that produces a two-dimensional image of the range of a voice in frequency and in amplitude. The appeal of such a tool lies in its capacity to depict subtleties of voice function and provide both quantitative and qualitative data. Sulter, in a study on

differences in phonetogram features between male and female subjects with and without vocal training, commented on the scarcity of reliable VRP data studies [1].

Many more VRP data have since been collected [2-10] but only a handful of studies have focused on VRP recordings of the singing voice [11-14]. These studies are often based on subject groups that consist mostly of students in training populations, amateurs, or a mix of choristers and soloists.

The VRP is known to be sensitive to gender, age, as well as vowels and other individual characteristics [1, 4, 6, 15-17]. It would follow that the VRP could also be dependent on training and/or profession [1]. In the case of the singer, the VRP could ideally be sensitive enough to distinguish subtleties of the professional singer's voice. Although a few university music programs in Europe have performed systematic VRP recordings of their students, few detailed analyses of singer VRPs have been published. Most VRP studies seem to focus on groups of speakers, and use the singer or trained group as a comparison point. The VRP seems to hold great potential for describing the singing voice, but in order for the VRP to become more clinically relevant, a frame of reference is needed to account for singer-specific issues, the possible impact of task design, and the possible need for additional or alternative VRP-derived singer specific metrics. This study's aim was to investigate whether VRP recording practice needs to be modified in order to be relevant to the singing voice.

Three main research questions were formulated.

Question 1. Is there a need to subclassify voices by singer category in a subject/patient VRP group?

Question 2. What tasks should be included in the protocol when the subject or patient is a singer? More specifically, should the tasks be musically designed to be as representative as possible of singing or singing exercises?

Question 3. Are there significant differences between the physiological VRP (i.e., the standard VRP) and the performance VRP (a VRP entailing singing voice quality with dynamics appropriate for the stage)? In the affirmative, where do these differences lie?

Method

Data Acquisition

The method for data acquisition was the same as in an earlier study [18]. For the reader's convenience, it is briefly restated here. Recordings were performed with a computerized, 16 bit linear acquisition, phonetograph (*Phog*, version 2.00.10, Saven Hitech AB, Sweden). This system accumulates phonation time in 2-D bins, or cells, 1 semitone (ST) wide and 1 dB high. Cells are plotted according to the UEP standard 2/1 (dB/ST) aspect ratio.

Since *Phog* is based on a peak-picking F_0 extraction, inevitably there was some degree of F_0 tracking latching onto higher harmonics. The recorded material was

inspected manually and the few instances of mistracking were removed. The recordings took place in a sound-treated and isolated recording studio (volume 45 m³, ceiling height 3 m, reverberation time, T30= 0.1 s, reverberation radius >1.2 m across the spectrum, and 0.5 m deep absorbents). Singers were asked to adopt a singing stance. Head and body movements were restricted as much as possible without impeding the freedom of the artist. The microphone-to-mouth distance (30 cm) was measured at the beginning of each task.

A condenser microphone (Brüel & Kjaer, model 4003, Denmark) was used with a pre-amplifier (Brüel & Kjaer, model 2812) and a line amplifier (Nyvalla-DSP Audio Interface Box). Singers were given a single piece earphone (Bassonic-Champion 4939, USA) to hear prompting tones during one of the tasks. For details concerning the voicing detection thresholds, the reader is referred to Lamarche *et al.*[18].

Subjects

Group criteria for this study were strict. The group included three voice categories: 6 contraltos, 8 mezzo-sopranos, 16 sopranos. Inclusion criteria included female opera soloist, non-smoking, more than 4 years of training, no ear-nose throat medical history, no respiratory problems and no actual voice complaints. No laryngoscopic examinations were performed. At the time of the recordings, all subjects were actively performing on classical/opera stages.

30 female opera singers with a mean age of 33.7 ± 8.8 years were recorded. The project was ethically vetted by the *Regional etikprövningsnämnden i Stockholm* (certificate 1358-31). Subjects were remunerated for their participation. Subjects had on average a training experience of 13.4 ± 5.9 years. Table 1 lists information and taxonomy pertinent to the subject group.

Procedure and tasks

The data collection took place from December 2006 to May 2008. In order to document the subject group thoroughly, each singer filled in a questionnaire addressing general health and vocal practice.

They also participated in five different types of recordings: one habitual speech range profile (SRP) one physiological VRP (VRP_{phys}) and three versions of a performance VRP (VRP_{perf}). Tone duration for the sustained tone tasks was roughly 2 seconds on the vowel [a]. The completion of all tasks took approximately 50-55 minutes. No specific instructions related to type of phonation and/or vocal strategies were given. Rather, subjects were asked to sing in a way representative of their performance voice use.

| Subjects | Age | Self-reported Voice Classification | Years of Training | Taxonomy | |
|----------|-----|------------------------------------|-------------------|----------|--------|
| 1 | 28 | Lyric soprano | 6 | 4.1b | R/T: m |
| 2 | 37 | Coloratura soprano | 9 | 3.1a | N: M |
| 3 | 43 | Lyric soprano | 6 | 2.1 | I: P |
| 4 | 26 | Lyric mezzo | 11 | 4.1b | R/T: m |
| 5 | 55 | Dramatic mezzo | 25 | 3.1a | N: M |
| 6 | 43 | Lyric soprano | 22 | 3.1a | N: M |
| 7 | 28 | Coloratura mezzo | 8 | 4.1b | R/T: m |
| 8 | 26 | Lyric soprano | 11 | 4.1b | R/T: m |
| 9 | 25 | Lyric soprano | 9 | 4.1b | R/T: m |
| 10 | 26 | Lyric mezzo | 8½ | 4.1b | R/T: m |
| 11 | 29 | Lyric soprano | 13 | 4.1b | R/T: m |
| 12 | 41 | Lyric mezzo | 17 | 3.1b | N: m |
| 13 | 39 | Lyric-dramatic mezzo | 15 | 2.1 | I: P |
| 14 | 25 | Lyric soprano | 8 | 4.1b | R/T: m |
| 15 | 32 | Lyric soprano | 17 | 4.1b | R/T: m |
| 16 | 20 | Lyric color. soprano | 9 | 4.1b | R/T: m |
| 17 | 25 | Lyric contralto | 8 | 4.1b | R/T: m |
| 18 | 28 | Lyric soprano | 9 | 4.1b | R/T: m |
| 19 | 20 | Lyric soprano | 6 | 4.1b | R/T: m |
| 20 | 46 | Light lyric soprano | 20 | 2.1 | I: P |
| 21 | 33 | Dramatic mezzo | 11 | 2.1 | I: P |
| 22 | 31 | Lyric soprano | 12 | 4.1b | R/T: m |
| 23 | 33 | Lyric soprano | 13 | 3.1a | N: M |
| 24 | 33 | Coloratura contralto | 11 | 4.1a | R/T: M |
| 25 | 33 | Dramatic soprano | 16 | 3.1a | N: M |
| 26 | 40 | Contralto | 10 | 2.1 | I: P |
| 27 | 33 | Contralto lyric | 23 | 4.1b | R/T: m |
| 28 | 48 | Contralto | 20 | 4.1b | R/T: m |
| 29 | 35 | Contralto | 17 | 3.1a | N: M |
| 30 | 49 | Mezzo lyric-dramatic | 27 | 4.1b | R/T: m |

Table 1 Participants' age, self-reported voice type, years of singing training, and taxonomy [37]. The following abbreviations are employed: Regional/Touring (R/T), National (N), International (I) and Major principal (M) and minor principal (m).

For the VRP_{phys}, the objective was the recording of minimum and maximum productions regardless of phonation type or laryngeal mechanism while for the VRP_{perf}, we wanted to capture the voice as it is used on stage. All five tasks were recorded in one session. The subjects could communicate with the investigator by intercom and visual contact through a window was possible. They could however not see the phonetogram display to avoid interference with a parallel task studied in Lamarche *et al.*[18].

Task 1a: A thematic spontaneous speech task was performed. Subjects were asked to make a 1 minute description of their warm-up routine.

Task 1b: A counting exercise in which the subject used soft (but no whisper), regular and loud public speaking voice. Separate SRPs were saved for each task. Subjects spoke in their native tongue (Swedish, French or German). Henceforth, the SRPs will be referred to as SRPs (1a and 1b).

Task 2: The VRP_{phys}. The aim was to register explicitly the subject's vocal extremes in pitch and in level. This was done with a descending *glissando* (a slow frequency sweep) and ascending *glissando* exercise on the vowel [a]. The *glissandi* were repeated and modified to acquire the best possible achievement (as deemed by the subject and the investigator).

For the VRP_{perf}, singers were instructed to sing as they deemed *musically acceptable for the stage*. Singing voice quality and vibrato were *obligatory* and the aim was to adhere to one's stage singing ideals at all times, *both* in pitch and in vocal dynamics.

At the start of each VRP_{perf} task, subjects were asked to sing a *messa di voce* on a comfortable tone in order to exercise and explore their full performance-mode dynamic range.

Task 3: A first VRP_{perf} was recorded with prompted frequencies equivalent to the musical notes *C-E-G-A* in several octaves across the singer's range. Prompted tones were augmented by semitones at the extremes [28]. Tones were sung on the vowel [a] in a *messa di voce* exercise (sustained pitches performed with increasing and decreasing vocal dynamic).

Task 4: This performance VRP task consisted of an ascending-descending order vocalise (triad carrier) on the vowel [a] in *pianissimo*, *mezzo-forte* and *fortissimo* (medium, soft and loud). Subjects were reminded to keep their task performance true to their vocal use on stage.

Task 5: For the third VRP_{perf}, subjects performed their best audition aria with lyrics. This task served to obtain a minimum of 1 minute of the voice in its most representative context. This was the only sung task that involved several different vowels. In a previous study, [20] the authors concluded that vowel variation in the high female opera singing voice VRP was negligible due to formant tuning.

Metrics of Importance

Here enumerated follow the metrics considered to be of interest for VRP analysis of the singing voice.

Minimum and Maximum Frequency (f_{\min}/f_{\max}): These values denote the minimum and maximum values of F_0 occurring in a given VRP.

Frequency Range (Rge): The F_0 range is simply $f_{\max} - f_{\min}$. It is expressed in octaves or semitones. An extended range in the physiological VRP is often assumed to be a logical consequence of voice training [1, 21]. This expectation could even possibly extend to the SRP [22]. Little information on the frequency range of singer subjects has however been reported in studies of VRP recording. Here, range will be reported for the SRPs, the physiological and the performance VRPs according to voice category.

Minimum SPL (SPL_{\min}): Minimum SPL values in the VRP_{perf} can be expected to be much higher than those expected for SRPs and for the VRP_{phys}. Schultz-Coulon estimated up to a 10-20 dB difference between a singer's *pianissimo* and a speaker's soft tone [24]. The main reason is simply that on stage even the quiet tones must be heard at the back of the hall, where phonation at the physiological threshold would be inaudible. Another reason is that control of the tone is poor at the threshold.

Maximum SPL (SPL_{\max}): According to previous reports, this metric would also be expected to vary with the type of VRP recording. However, the direction of this variation remains unclear. Certain studies claim that physiological VRPs show higher maximal intensities. Singers might however be inhibited in a laboratory setting, but more easily draw on their full resources when given the proper context.

SPL Range (SPL_{rge}): Western opera and lyrical vocal music require a substantial dynamic range. We recall here that SPL covaries strongly with F_0 [15,20,25]. It is acoustically inevitable that low SPL values will be difficult or impossible to produce at high frequencies, and vice versa for the lower range. Hence a large F_0 range will tend to be associated with a large range in SPL. Therefore the overall SPL range does not directly reflect the singer's ability to modify her output power.

Average SPL Extent (SPL_{ext}): For a given F_0 , we define the SPL_{ext} as the level difference between the upper and lower bounds of the contour; in other words as the height of the phonation area at any given F_0 . This extent is then averaged from lowest to highest F_0 , giving a metric for how much the singer can modify SPL at constant F_0 . In this way the dependency of SPL on F_0 is compensated for. Since the voices studied here are trained in maintaining dynamic stability across the frequency range of the voice, we can expect the SPL extent of a singer to be larger and more consistent than that found for untrained voices.

Area: This VRP metric quantifies the 2-dimensional range. It is calculated by counting all visited cells; or sometimes all cells contained within the grand contour, be they visited or not. This metric is widely used in VRP analysis and in comparisons between studies. According to Awan and especially Sulter's reports of a logistic regression, the enclosed area metric was best at differentiating female untrained voices from female singers [1, 21].

Area above 90 dB: Singers need to be heard when they stand on a stage and are accompanied. Indeed, classical singing technique develops the ability to produce loud sounds and also to maintain higher energy in the 2.5-3 kHz region of the spectrum (the singer's formant cluster, or spectrum resonance peak). Without amplification, a certain minimum power is needed to make oneself heard in a given performance situation. Although the voice spectrum would also be relevant, it is plausible that a rough criterion for a useable stage voice could be the VRP area above some minimum SPL (corresponding to a minimum singer power). The question is then how to select a suitable threshold level. In an earlier unpublished study, data was collected that could be applied for this purpose. 3 sopranos and 2 mezzo-sopranos were asked to phonate on a series of different pitches on a /papapa/ exercise. They phonated in *piano*, *mezzo forte* and *forte*. The SPL range obtained for these five singers measured at 30 cm from the mouth and for the midi pitches 60, 65, 69, 74, 79 (C^4 - G^5) was 66-112 dB. The mean SPL for a *piano* across all singers was 83 dB. The level increment was 6.5 dB between *piano* and *mezzoforte* and 3.6 dB from *mezzoforte* to *forte*. A *mezzoforte* was equivalent to roughly 90 dB. This agrees well with data from Nawka [26]. The exact value of the chosen threshold level is not critical, as it is unlikely to have a large effect on the conclusions arising from comparing VRPs; but in order to be normative, the choice must be well informed.

For analysis purposes, this area will be related to the total area and a percentage of vocal presence in the 90 or more dB area will be reported ($\text{Percent}_{\geq 90\text{dB}}$).

VRP slope: Slope metrics can be defined in many ways and are not readily compared from one study to another. Not only do slopes depend on many factors such as mouth radiation, voice source parameters (mean flow declination rate, pulse rate) and possibly acoustic strategies (F_0 - $F1$ tuning) [25], but they are also very dependent on the actual VRP shape. Some earlier studies have reported slope values for partial contour segments [8, 21]; however, such slope values would reflect the total effect of several underlying mechanisms that would need to be accounted for separately. In producing group data, many different shapes are averaged to give a group contour, and so a slope value in this instance becomes less informative. Furthermore, VRP shapes tend to be rounded and make it difficult to systematically define a tangent. It is also debatable what the slope value actually represents, when the phonatory modes are not accounted for separately. For these reasons, slopes will not be reported in this paper.

The SRP recordings (1a and 1b) were analysed with the SRP metrics: minimum, maximum, range and average in frequency and in SPL. The total area of phonation was also reported.

Analysis

The normality of the distribution was assessed by examining closely the kurtosis and skewness levels. Comparative statistical tests were selected to assess SRP and VRP data. The probability alpha was set to 0.01. A general linear multivariate analysis was performed for the dependent variables: R_{ge} , f_{min} , f_{max} , SPL_{rge} , SPL_{min} , SPL_{max} , SPL_{ext} , Area, Percent $_{>90dB}$. Fixed factors were Task (4 levels-here we excluded continuous speech tasks) and Voice Category (3 levels). In the event that the F test resulted in significant differences, the Ryan-Einot-Gabriel-Welsch Range test was conducted to assess the difference among the factors and dependents. The non-parametric Wilcoxon Signed Rank test for paired samples, was performed for SRP data. All analysis was performed with SPSS 15.0 for Windows, SPSS Inc.

The Fourier transform (FT) is often used in image processing to detect and assess shapes. A novel Fourier Descriptor (FD) approach to contour averaging was used here to compare and depict the collected data. The Fourier descriptor method has several useful features, including the ability to deal with translation, scale changes and even rotation. A *contour spectrum* is calculated, filtered and inverse transformed to yield a smooth curve that connects each point of the VRP contour. New data points can then be interpolated over this contour.

This technique allows for the creation of average contours regardless of their original sampling (F_0 /SPL range or area size), and can also depict the co-variation across the averaged contours (see figure 1). This enables the comparison of multi-source data in one graph. A methodological paper concerning the detailed description of the FD technique is currently in review (Pabon, Lamarche & Ternström, in review).

Results

Questionnaire results are tabulated in Table 2. This group of subjects was overall healthy with moderate physical training habits, healthy weight and very low intake of medicine. Vocal habits were rated “moderate,” yet extensive voice use and training experience were noted.

Descriptive statistics for the VRP metrics are reported in a series of tables. Table 3 gives the group means and standard deviations for SRP metrics. Table 4 reports the statistics per voice category for the sung tasks. The VRP_{phys} was only introduced later in the experiment, and so the number of subjects for which the VRP_{phys} is available is smaller (Sopranos=8, Mezzosopranos=2, Contraltos=6).

SRP metrics did not vary substantially from task 1a to 1b, and standard deviations (SD) were quite small, indicating good agreement within the group. For the other tasks, differences were more noticeable from one task to another. From the physiological to the performance VRP, the frequency range was reduced from 3.3 to 2.8 octaves (38.6 to 33.3 semitones). Naturally the Aria performance VRP has a much more reduced range (constrained by the composition chosen by the singer). In fact, the results in all metrics but one were constrained when moving from the physiological task to the aria. The exception was the percentage of the voice use at 90 dB and above, which increased (from 30% in the physiological profile to 51 % in the aria).

Averaged VRPs depict the results for each task while differentiating the voice categories. Figure 1 illustrates the contour averages and covariation for the counting task (1b). Clearly, mezzosopranos and contraltos, even in speech, exercise their low range more than the sopranos. In figure 2 the averages and covariation are displayed for the speaking task(1a), for which the same observation can be made.

The significant differences between speaking (1a) and counting (1b) tasks were observed for SFF, SPL_{min} , SPL_{rge} and Area. Table 5 gives the test results. These results can also be assessed in Figure 3 where the SRP (1a) for the complete group ($N=30$) is superimposed onto the SRP (1b). Figure 4 a) displays the SRP(1b) within the physiological contour of the group. The speech area covers roughly 37% of the physiological VRP area. Figure 4 b) shows the corresponding comparison for the performance VRP. Figures 5 and 6 (a-c) illustrate the results for the contour averaging of the physiological and performance tasks.

Table 6 (a-b) is an adapted SPSS table of the multivariate analysis results for the sung tasks. The fixed factors Task and Voice Category both had a significant effect on VRP metrics. There was no interaction between the factors. In table 6 a) results for Pillai's Trace are reported. With the exception of SPL_{max} , all metrics varied significantly with the Task (Table 6 b). Conversely, Voice Category seems to have had a limited effect, with significant levels of difference obtained for the f_{min}/f_{max} and range metrics only.

| Characteristics | Soprano | Mezzo-Soprano | Contralto | Group Values |
|------------------------------------|----------------|----------------|----------------|----------------|
| Dispersion | 16 | 8 | 6 | 30 |
| Age group | 20-46 | 25-55 | 33-48 | 20-55 |
| Age mean | 30,9 | 36,8 | 36,3 | 33,7 |
| Age Stdev | 8,0 | 11,3 | 6,7 | 8,8 |
| Voice Training/yr mean | 12,1 | 15,6 | 14,8 | 13,4 |
| Voice Training /week | | | | |
| A-Daily or more | 10 / 16 | 5 / 8 | 5 / 6 | 20 / 30 |
| B-4 _x to 6 _x | 6 / 10 | 3 / 8 | | 9 / 30 |
| C-Less than 4x | | | 1 / 6 | 1 / 30 |
| Training mean length | 1:20 hr | 1:05 hr | 1:15 hr | 1:12 hr |
| Use of spoken voice | moderate | moderate | moderate | moderate |
| Use of singing voice | mod-great | moderate | moderate | moderate |
| Tobacco intake | 1 | 0 | 0 | 1 |
| Body Mass Index | | | | |
| A-healthy | 12 / 16 | 5 / 8 | 5 / 6 | 22 / 30 |
| B-overweight | 3 / 16 | 2 / 8 | 1 / 6 | 6 / 30 |
| C-obese | 1 / 16 | 1 / 8 | 0 | 2 / 30 |
| Physical | | | | |
| Training/week | 2 _x | 2 _x | 4 _x | 3 _x |
| Medicine intake | | | | |
| A-prescribed | | | | |
| B-contraceptives | 2 | 2 | 3 | 7 |
| /hormones | | | | |
| C-over the counter | 3 | 2 | | 5 |
| /homeopathic | | | | |
| D-allergies/asthma | 2 | 2 | 1 | 5 |
| E-none | 9 | 2 | 2 | 13 |

Table 2. Physical and vocal health questionnaire results for a group of 30 singers. We denote frequency of training with ‘x’ (The one case of tobacco intake which is here reported is not associated to smoking but rather to “snuff”).

| | Spontaneous Speech | | Counting | |
|--------------------|--------------------|-------|----------|-------|
| | Mean | SD | Mean | SD |
| f_{\max} | 362.55 | 43.00 | 355.67 | 41.43 |
| f_{\min} | 146.62 | 21.60 | 152.12 | 21.25 |
| Rge (octave) | 1.31 | 0.27 | 1.24 | 0.17 |
| Rge (ST) | 15.73 | 3.18 | 14.76 | 2.12 |
| SFF | 229.86 | 25.00 | 238.18 | 26.62 |
| SPL_{\max} | 84.63 | 5.20 | 84.79 | 3.52 |
| SPL_{\min} | 60.17 | 3.88 | 57.14 | 3.37 |
| SPL_{rge} | 24.47 | 5.34 | 27.66 | 3.99 |
| SPL_{avg} | 71.80 | 3.60 | 71.06 | 2.50 |
| SPL_{ext} | 13.77 | 3.25 | 17.44 | 2.83 |
| Area | 224.67 | 60.30 | 280.69 | 61.02 |

Table 3 Descriptive statistics for the SRP of 30 subjects. Frequencies are reported in Hz and sound pressure levels in dB relative to 30 cm. Rge statistics are in octaves but a semitone (ST) conversion is provided for convenience. Area is determined by the number of visited cells. SD is short for standard deviation.

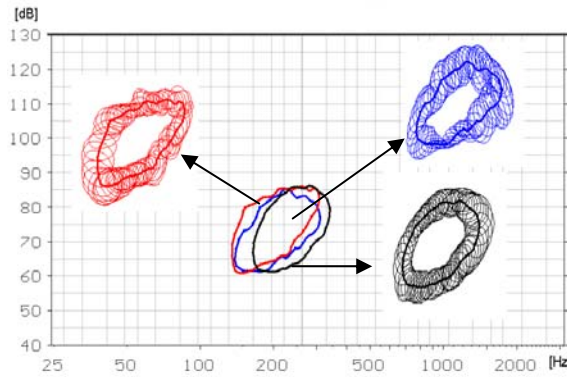


Figure 1 Average SRP contours for the counting speech task 1b (N= 30, soprano in black (16), mezzosopranos in red (8) and contralto in blue (6). The insets show the two-dimensional standard deviation as ellipses, whose orientation also suggests the local covariation of F_0 with SPL.

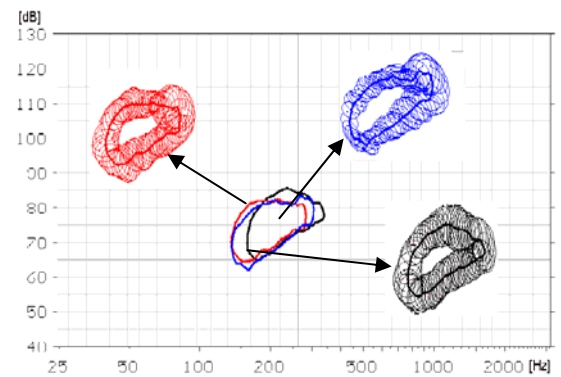


Figure 2 Average SRP contours for the spontaneous speech task 1a (N= 30, soprano in black (N=16), mezzosopranos in red (8) and contralto in blue (6). Insets show the standard deviations as for Figure 1.

| Task | | Category | | | | | | |
|-------------------------------------|----------|-----------------|--------|----------------------|--------|-------------------|---------------|--------|
| | | Soprano Mean | SD | Mezzosoprano Mean | SD | Contralto Mean | Total Mean | SD |
| f_{\max} | Fysio | 1315.23 | 223.86 | 1176.66 | 96.01 | 1186.53 | 1226.14 | 80.42 |
| | Pitch | 1295.13 | 115.89 | 1147.25 | 181.50 | 982.23 | 1141.54 | 39.70 |
| | Vocalise | 1173.02 | 79.70 | 1061.70 | 195.46 | 988.00 | 1074.24 | 58.66 |
| | Excerpt | 970.61 | 128.42 | 942.07 | 150.36 | 800.97 | 904.55 | 27.67 |
| f_{\min} | Fysio | 151.53 | 11.71 | 113.66 | 13.89 | 128.59 | 131.26 | 2.47 |
| | Pitch | 174.46 | 19.71 | 131.71 | 16.74 | 160.37 | 155.51 | 2.73 |
| | Vocalise | 169.72 | 24.23 | 144.93 | 35.39 | 158.20 | 157.61 | 7.46 |
| | Excerpt | 258.22 | 36.90 | 241.41 | 41.46 | 215.05 | 238.23 | 7.61 |
| Rge | Fysio | 3.14 | 0.27 | 3.45 | 0.07 | 3.20 | 3.26 | 0.16 |
| | Pitch | 2.92 | 0.20 | 3.14 | 0.25 | 2.64 | 2.90 | 0.15 |
| | Vocalise | 2.83 | 0.25 | 2.91 | 0.42 | 2.67 | 2.80 | 0.08 |
| | Excerpt | 1.93 | 0.18 | 1.99 | 0.26 | 1.90 | 1.94 | 0.05 |
| SPL_{\max} | Fysio | 112.75 | 8.00 | 115.50 | 2.12 | 113.17 | 113.81 | 3.24 |
| | Pitch | 114.00 | 3.58 | 112.38 | 6.09 | 114.60 | 113.66 | 1.27 |
| | Vocalise | 113.50 | 2.92 | 114.00 | 5.61 | 108.83 | 112.11 | 1.75 |
| | Excerpt | 111.50 | 4.27 | 110.50 | 5.98 | 108.67 | 110.22 | 0.90 |
| SPL_{\min} | Fysio | 56.38 | 3.46 | 55.50 | 0.71 | 58.00 | 56.63 | 1.38 |
| | Pitch | 63.75 | 3.75 | 64.13 | 5.03 | 60.00 | 62.63 | 0.64 |
| | Vocalise | 66.38 | 4.49 | 65.00 | 4.11 | 65.83 | 65.74 | 0.22 |
| | Excerpt | 71.44 | 6.25 | 74.00 | 3.63 | 67.50 | 70.98 | 1.47 |
| SPL_{rge} | Fysio | 56.38 | 6.82 | 60.00 | 1.41 | 55.17 | 57.18 | 3.96 |
| | Pitch | 50.25 | 3.47 | 48.25 | 8.99 | 54.60 | 51.03 | 2.80 |
| | Vocalise | 47.13 | 3.83 | 49.00 | 8.75 | 43.00 | 46.38 | 2.99 |
| | Excerpt | 40.06 | 5.28 | 36.50 | 5.15 | 41.17 | 39.24 | 0.17 |
| SPL_{ext} | Fysio | 26.86 | 4.79 | 29.07 | 4.41 | 27.48 | 27.80 | 1.02 |
| | Pitch | 17.34 | 2.63 | 15.94 | 3.83 | 19.64 | 17.64 | 0.60 |
| | Vocalise | 16.16 | 4.00 | 18.24 | 6.11 | 13.69 | 16.03 | 1.90 |
| | Excerpt | 14.24 | 2.75 | 12.79 | 3.58 | 13.86 | 13.63 | 0.70 |
| Area | Fysio | 762.75 | 275.58 | 725.00 | 4.24 | 734.50 | 740.75 | 140.66 |
| | Pitch | 528.06 | 104.94 | 534.63 | 137.60 | 522.00 | 528.23 | 17.33 |
| | Vocalise | 541.94 | 144.60 | 642.50 | 215.72 | 478.33 | 554.26 | 57.34 |
| | Excerpt | 338.44 | 78.20 | 312.50 | 76.90 | 324.17 | 325.03 | 1.39 |
| $\text{Percent}_{\geq 90\text{dB}}$ | Fysio | 30.66 | 9.57 | 28.13 | 1.20 | 30.87 | 29.89 | 4.39 |
| | Pitch | 46.59 | 7.32 | 41.50 | 8.99 | 39.66 | 42.58 | 1.07 |
| | Vocalise | 48.74 | 13.32 | 41.43 | 8.66 | 37.95 | 42.70 | 3.51 |
| | Excerpt | 50.38 | 15.25 | 53.99 | 12.95 | 46.50 | 50.29 | 1.17 |

Table 4 Means for VRP metrics per Voice category and per sung Task. The standard deviation is referred to as SD. Range is indicated in octaves.

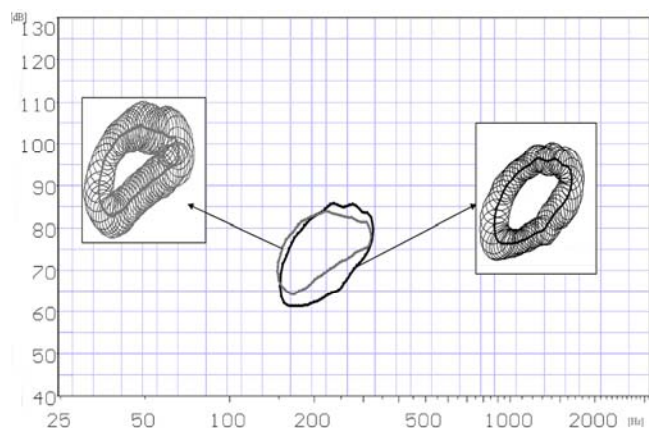


Figure 3 Average SRP contours for two types of speech, counting (1b) and spontaneous (1a), compared. Counting is in black and spontaneous speech is in grey. ($N=30$) Insets show standard deviations as for Figure 1.

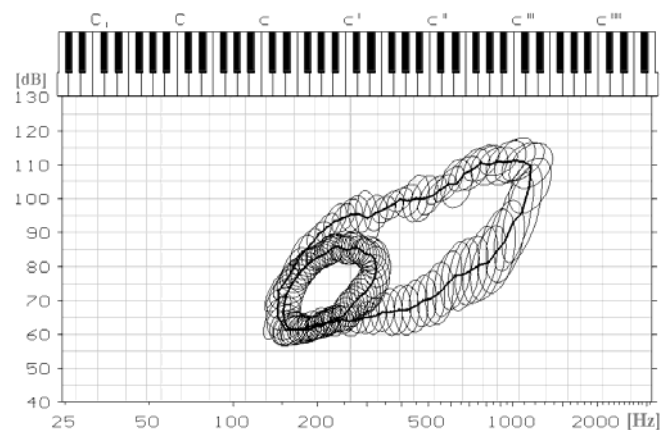


Figure 4a. Average SRP countours for speech (counting, inner contour) and for the physiological task (outer contour). Covariation ellipses are included. The lower minimal curves align closely, as do the lower portions of the upper contours. (counting, $N=30$, physiological, $N=16$).

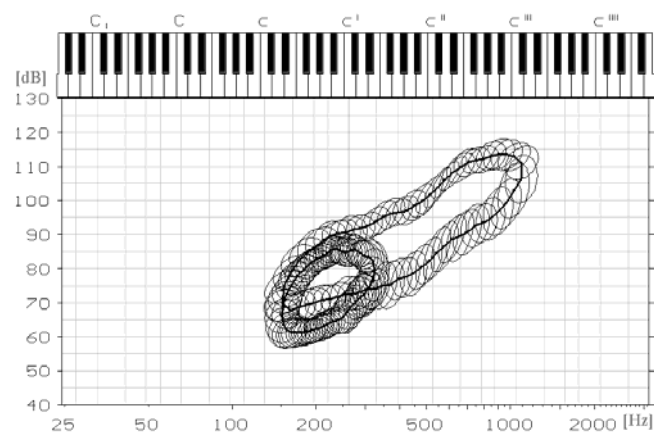


Figure 4b. Average SRP countours for speech-counting (1b) and for the sung performance task , including covariation ellipses ($N=30$).

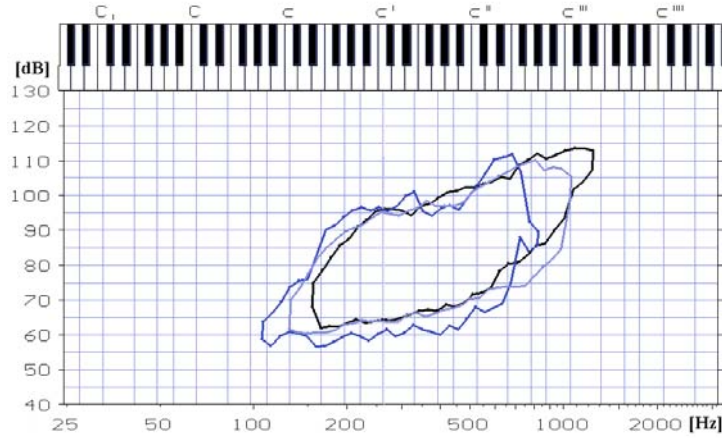


Figure 5 Average VRP contours for the physiological task (Task 2); for sopranos in black (8), for mezzosopranos in blue (2) and for contraltos in light grey (6), superimposed for comparison.

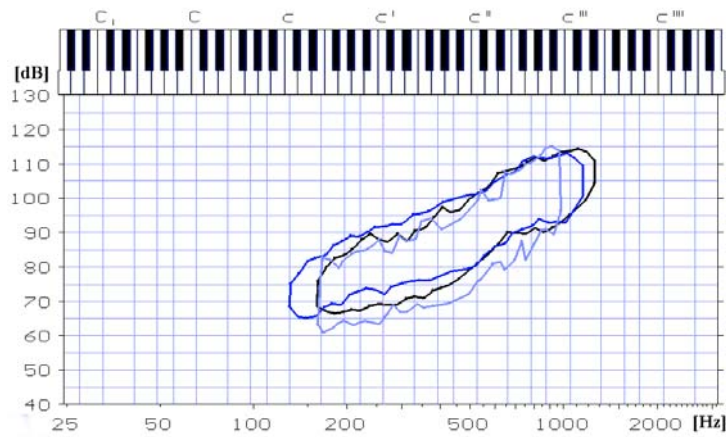


Figure 6 a) Average performance VRP contours for the discrete pitch task (Task 3). Sopranos are in black (16), mezzosopranos in blue (8) and contraltos in light grey (6).

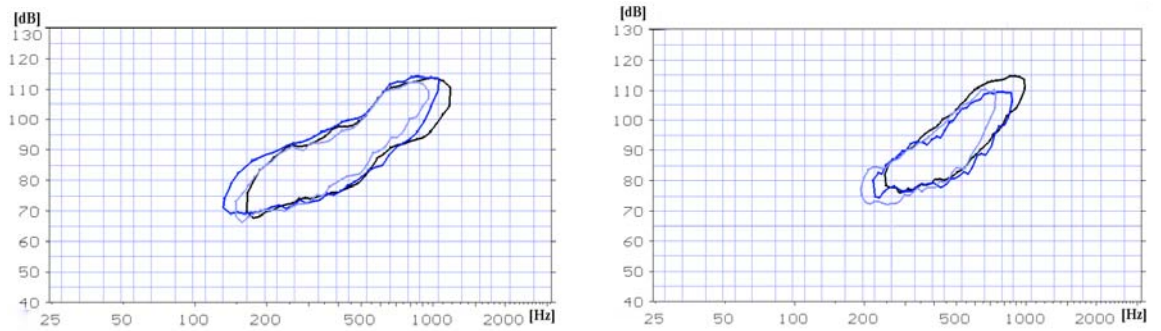


Figure 6 b) Average performance VRP contours for the vocalise task (Task 4) and **c)** Average performance VRP contours for the aria excerpt (Task 5). Sopranos are in black (16), mezzosopranos in blue (8) and contraltos are in grey (6).

| | Z | Asymp. Sig. (2-tailed) |
|-----------------------------|--------|---------------------------|
| $f_{\max 2} - f_{\max}$ | -0.835 | 0.404 |
| $f_{\min 2} - f_{\min}$ | -1.711 | 0.087 |
| $rge_2 - rge$ | -1.511 | 0.131 |
| $SFF_2 - SFF$ | -2.714 | 0.007 |
| $SPL_{\max 2} - SPL_{\max}$ | -0.364 | 0.716 |
| $SPL_{\min 2} - SPL_{\min}$ | -3.275 | 0.001 |
| $SPL_{rge 2} - SPL_{rge}$ | -2.518 | 0.012 |
| $SPL_{avg 2} - SPL_{avg}$ | -1.330 | 0.184 |
| $SPL_{ext 2} - SPL_{ext}$ | -3.946 | <0.01 |
| $Area_2 - Area$ | -3.482 | <0.01 |

Table 5 Wilcoxon Signed Rank Test results. SRP metrics tagged with a 2 are related to the counting task (1b), SRP features without a number suffix refer to the spontaneous task (1a). The significant differences, $p < 0.01$, are indicated in bold: SFF, SPL_{\min} , SPL_{ext} and Area.

| Effect | | Value | F | Hypothesis df | Error df | Sig. | Partial Eta Squared |
|-----------------|----------------|-------|-------|---------------|----------|-------|---------------------|
| Task | Pillai's Trace | 1.335 | 8.824 | 24 | 264 | <0.01 | 0.445 |
| Category | Pillai's Trace | 0.578 | 4.426 | 16 | 174 | <0.01 | 0.289 |
| Task * Category | Pillai's Trace | 0.679 | 1.452 | 48 | 546 | 0.028 | 0.113 |

Table 6a) Multivariate test results. A significant main effect of both factors, Task and Category, is observed but no interaction between the two is observable. Significance is determined with $p < 0.01$. Degree of freedom is represented by df and significance by Sig.

| Task | |
|------------------------|--------------------------------|
| f_{\max} | (F (3.93)=14.74, $p < 0.01$). |
| f_{\min} | (F (3.93)=64.33, $p < 0.01$). |
| Rge | (F (3.93)=81.35, $p < 0.01$). |
| SPL_{\min} | (F (3.93)=29.52, $p < 0.01$). |
| SPL_{rge} | (F (3.93)=29.69, $p < 0.01$). |
| SPL_{ext} | (F (3.93)=39.46, $p < 0.01$). |
| Area | (F (3.93)=26.58, $p < 0.01$). |
| Percent $_{\geq 90dB}$ | (F (3.93)= 8.72, $p < 0.01$). |
| Category | |
| f_{\max} | (F (2.93)=13.34, $p < 0.01$). |
| f_{\min} | (F (2.93)=11.35, $p < 0.01$). |
| Rge | (F (2.93)=4.84, $p > 0.01$). |

Table 6b) F test. All metrics are significantly different across tasks with the exception of SPL_{\max} (which is not included in the table). Maximum and minimum frequency also significantly differ in terms of voice category. The Rge metric seems also to differ but yields no significant result. Significance is determined by $p < 0.01$.

Effects of Voice Category

Table 7 a) contains the statistical details of the Voice Category comparisons.

An overall difference was manifest for the f_{\min} and f_{\max} metrics as well as a borderline result for Rge. The difference lies between the low voices and the high voice with no significant difference for contraltos and mezzos in both f_{\min} and f_{\max} metrics. Furthermore, there are no significant differences between mezzos and sopranos for the f_{\max} . The class averages for the Rge metric were very close to each other (2.6 octaves for contraltos; 2.7 for mezzosopranos; 2.6 for sopranos).

| Category | | f_{\max} Subset | | f_{\min} Subset | | Rge Subset |
|----------|--------------|----------------------|---------|----------------------|--------|------------|
| | | 1 | 2 | 1 | 2 | |
| REGWR | Contralto | 989.75 | | 165.78 | | 2.60 |
| Test | Mezzosoprano | 1060.06 | 1060.06 | 168.14 | | 2.74 |
| | Soprano | | 1170.39 | | 193.76 | 2.64 |

Table 7a) Table of the R-E-G-W-R multiple comparisons test for Voice category. Means that appear in the same homogeneous subset are not significantly different from each other ($p < 0.01$).

Effect of Task

Post hoc comparisons revealed no significant differences between the discrete pitch task (Task 3) and the vocalise exercise (Task 4). These observations are corroborated in Table 7b). As expected, the sung aria was significantly different from all other tasks, except in SPL_{ext} where it could not be differentiated from the vocalise task, and in $\text{Percent}_{\geq 90\text{dB}}$ where all performance tasks were not distinct from one another. Figure 7 illustrates the contour averages for the three performance tasks. Both the discrete pitch and vocalise tasks yielded rather similar vocal outputs. The differences that could be noted were related mostly to lower VRP contour details.

In the frequency metrics, the VRP_{phys} did not differ significantly from the VRP_{perf} . Rather, a marked distinction between both types of VRPs was associated to intensity metrics. In the SPL_{\min} , SPL_{ext} and the Area metrics, there was a clear distinction between the VRP_{phys} and the VRP_{perf} . No significant differences were found between the VRP_{phys} and Task 3 data with respect to SPL_{rge} and $\text{Percent}_{\geq 90\text{dB}}$. Figure 8 shows the two contour averages. Since greater statistical difference was found between the VRP_{phys} and Task 4, the vocalise contour was used to represent the VRP_{perf} .

| R-E-G-W-R Test | | Subset 1 | Subset 2 | Subset 3 |
|--------------------------------|-------------------|-------------|-------------|-------------|
| f_{max} | Aria | 929.07 | | |
| | Vocalise | | 1200.39 | |
| | Discrete Pitch | | 1106.33 | |
| | Physiological | | 1249.65 | |
| f_{min} | Aria | | 245.11 | |
| | Vocalise | 160.24 | | |
| | Discrete Pitch | 160.80 | | |
| | Physiological | 138.19 | | |
| R_{ge} | Aria | 1.94 | | |
| | Vocalise | | 2.93 | 2.93 |
| | Discrete Pitch | | 2.82 | |
| | Physiological | | | 3.20 |
| SPL_{min} | Aria | | | 71.33 |
| | Vocalise | | 65.90 | |
| | Discrete Pitch | | 63.21 | |
| | Physiological | 56.88 | | |
| SPL_{rge} | Aria | 39.33 | | |
| | Vocalise | | 46.80 | |
| | Discrete Pitch | | 50.45 | 50.45 |
| | Physiological | | | 56.38 |
| SPL_{ext} | Aria | 13.77 | | |
| | Vocalise | 16.22 | 16.22 | |
| | Discrete Pitch | | 17.35 | |
| | Physiological | | | 27.36 |
| Area | Aria | 328.67 | | |
| | Vocalise | | 556.03 | |
| | Discrete Pitch | | 528.83 | |
| | Physiological | | | 747.44 |
| Percent_{≥90dB} | Aria | | 50.56 | |
| | Vocalise | | 44.63 | |
| | Discrete Pitch | | 43.99 | |
| | Physiological | 30.42 | | |

Table 7b) Table of the R-E-G-W-R multiple comparison test for Task. Means obtained for each metric are tabulated, and means that appear in the same homogeneous subset are not significantly different from each other ($p < 0.01$).

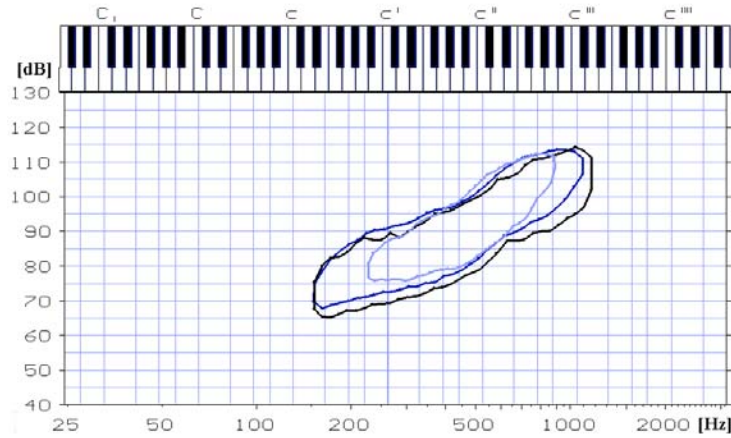


Figure 7. Average VRP contours for the discrete pitch task (task 3) in black, vocalise task (task 4) in blue; aria excerpt (task 5) in light grey (N=30).

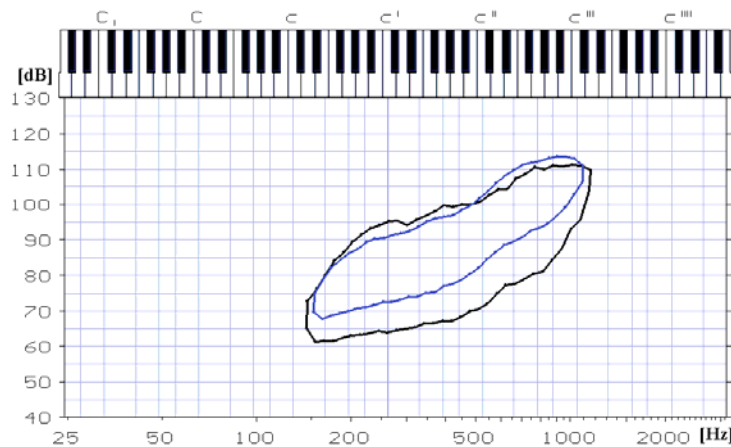


Figure 8. Comparison of averaged contours for the physiological (Task 2-black) and vocalise (Task 4-blue) tasks, for a group of 16 singers.

Discussion

The results reported in this study help elucidate the performance aspects of the singing voice and how they might impact the VRP. A professional Western opera soloist has different requirements for his/her instrument than does a speaker [26]. As seen earlier, some have demonstrated range differences between the physiological VRPs of untrained and trained voices however, physiological ranges might not necessarily greatly differ in practice. Rather, voice control, is often considered the greatest differentiating aspect between trained and untrained. The contour of the VRP_{phys} does not readily lend itself to the interpretation of such a vocal feature. The VRP_{phys} strives rather to capture the minimum threshold of phonation as well as unrefined vocal transitions. On the other hand, the VRP_{perf} might enable us to understand subtleties of what can be considered functional for a singer (considerably

different from the speaker's need for vocal function). Just as the SRP enables the clinician to obtain a behavioral type of VRP acquisition, the VRP_{phys} seems to demonstrate interesting behavioral aspects of the singing voice that are akin to singing and not necessarily present in non-singing voice use.

Speech

SRP data was here included since it seldom accompanies VRP reports in other studies but is an important part of the total voice evaluation. Our result for the SRPs (Task 1a and 1b) agreed well with the speech range data of Drew and Sapir [28]. They reported an increase of SFF in reading when comparing spontaneous speech and reading tasks. In our study, the reading task was substituted by the counting task (Task 1b). Drew and Sapir reported a mean of 219 Hz for speech and an increased mean of 230 Hz for reading. We found averages of 242 Hz (1a) and 251 Hz (1b) respectively. (Only our soprano data is commented, since the Drew and Sapir study was conducted with 10 healthy soprano subjects). When compared to healthy female native Swedish speakers, the SFFs obtained here (both for 1a and 1b) are quite high. Kitzing reported a SFF of 193 Hz with a standard deviation of 2.7 semitones for a group of 141 Swedish female speakers [30]. Yet, when observed per voice category, the SFF averages obtained (soprano=242 Hz, mezzosopranos=212 Hz, contraltos=220 Hz) relate somewhat better to Nadoleszny's results as reported by Drew and Sapir (soprano=262 Hz, mezzosopranos=230 Hz, contraltos=212 Hz). Awan also reports a higher SFF for a group of trained voices as opposed to untrained voices [22]. See Table 4 a) for detailed SRP results.

According to Hacki, a speech profile in normal cases should be approximately $\frac{1}{3}$ of the VRP [31]. It is not very clear whether he refers to a VRP_{perf} (like Tasks 3-4 of this study) or a VRP_{phys} (like Task 2 of this study). Data collected in the present study suggested that Hacki's conclusion was most likely based on a VRP_{phys} . Speech and counting contours had a range of 1.3 octaves while the physiological VRP had a 3.3 octave range. In other words, the SRPs recorded in our study occupied the bottom third of the VRP_{phys} , covering 30 to 37% of its total area. When related to the performance profiles, SRPs covered 40-41% of the VRP_{perf} area. Tables 5 a-b) exemplify these observations. On direct juxtaposition, SRPs were not completely enclosed by VRP_{perf} . Although there was a good correspondence in minimum frequency for both the SRPs and the performance profiles, the SRPs displayed lower minimum SPL values than what was found for the performance voice. This falls in line with other reports.

This last observation might correspond to the nature of both types of phonations. Coleman claimed that sustained tones would lead to higher intensities than intermittent phonation such as found in speech [23]. We observed a 7 dB difference in SPL between the soft spoken tones and the sustained performance-like phonations. While there were differences in the lower contour, all profiles followed a similar trajectory for the upper contour. As Pabon has observed (personal communication), the left portion of the upper

contour (the initial rise of the maximum VRP curves) is often a location of convergence when comparing within individuals, within groups and even across groups.

Concerning the maximum SPL in speech, Hacki stated that values of 80-90 dB were normal values for the case of individuals with “good voice capabilities”[31]. In our investigation, we obtained similar results, with maxima of 84 and 85 dB for the speech and counting tasks respectively. Furthermore, Sulter and Awan consider the intensity range of 60-80 dB to be important for normal communication [32, 33]. Subjects in our study had a similar speech intensity range and maintained, on average, a level of 71 dB. Indeed, subjects in this group were quite loud while speaking. This could be a result of the dampened acoustics of the recording studio. Pooling the data for soprano, alto and age group data of Brown *et al* (1993) (corrected for their smaller microphone distance) we obtained a mean of 64 dB [29]. This is a somewhat lower value considering that the subjects were 14 professional singers. A mean level of 62 dB was reported for their nonsinger group. These studies all seem to indicate that in terms of speech power, the differences between speakers and singers are very small.

Effects of Voice Category

A rather weak overall effect of voice category was observed in both the VRP_{phys} and the VRP_{perf} (voice category was not statistically tested for the speech data). Only the minimum and maximum frequency metrics differed significantly between sopranos and contraltos (Table 7a). Mezzosopranos had the largest range with f_{min} approaching that of contraltos and f_{max} near to that of sopranos. This is not unexpected since mezzosopranos are often required to have the same high pitches as sopranos as well as access to lower pitches similarly to contraltos. The distinction between those categories is usually a matter of *tessitura* and timbre. In figure 6b the VRP_{perf} for mezzosopranos and sopranos can be compared. The SPL_{ext} had no statistical weight, yet, upon inspection of the mezzosoprano and soprano performance profiles, it appeared that despite similar ranges for both categories, sopranos demonstrated a greater SPL_{ext} in the higher portion of the voice.

Since this study was not concerned with voice quality metrics or singer self-perceptions, no distinction could be made between what was comfortably or easily executed and that which was not. Such an investigation in combination with the VRP could be interesting. In that case the investigations should include a more even distribution across voice categories.

In VRP recording, grouping of subjects according to voice categories is rarely reported. Hacki (1990) qualitatively explored the differences between ten sopranos and ten contraltos [33]. He found that differences between voice categories were especially clear. He pointed out a smaller SPL extent for the middle frequencies in the soprano voices. He also considered the flat portion of the minimum VRP curve between 131 and 440 Hz to be a characteristic trait of the contralto VRP. These two observations do not agree well with our VRP data for the same voice categories. As is demonstrated

in figure 9a), the SPL extent at middle frequencies for both voices, are quite similar. Rather, this observation seems more relevant for the differences at the F_0 extremes of the physiological VRP. When the lower contours are compared for both voices, we note that the two voice categories converge well with increasing frequency and the slow rise in intensity that usually accompanies them. When the VRP_{perf} are similarly compared (see figure 9b), the voice category differences are manifest in the upper high end of the VRP, where sopranos display a larger SPL extent, consistent with a greater vocal flexibility and control at high pitch. The frequency range difference is again clear and seems to follow voice category definitions.

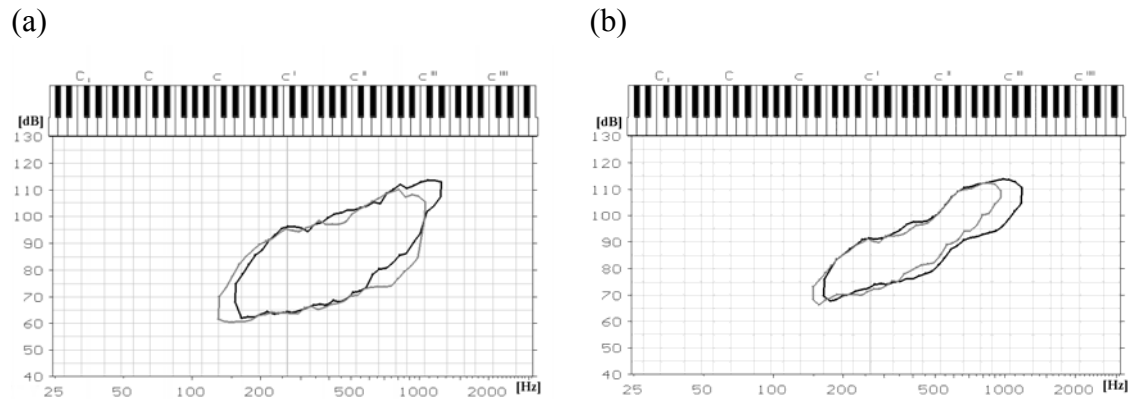


Figure 9 Average VRP contours for the physiological task (a) and the performance tasks (b). Sopranos in black (N=8) and contraltos in light grey (N=6).

Effects of Task

Reich *et al* (1989, 1990) tested thoroughly the effect of different tasks in recording the frequency ranges of children and adults [34]. In those studies it was concluded that continuous tasks such as glissandi or small steps task led to better results in regard to frequency range. For frequency minima, the slower glissando produced lower values than the rapid glissando exercise. Although the authors focused only on frequency, these outcomes can be interestingly related to our results.

According to the earlier stated hypotheses, the tasks for this experiment were designed to test specifically if 1) singers would resort to a more representative use of the voice in a performance task and if 2) in a performance task, a continuous expiratory gesture would lead to higher vocal flexibility (both in frequency and intensity). The inclusion of the aria excerpt served mainly to assess the possible difference between realistic singing and task singing: an approach similar to that used with actors by Emerich *et al* [35].

An overall main effect of tasking was found in the statistical analysis. As expected, the aria excerpt task was significantly different in almost all of the investigated metrics. Similarly to Emerich's study of actor VRPs and Speech Range Profile (SRP), our data confirm that the nature of the task and the performance setting suggested to the singers will impact the results that one obtains [35]. Minimum SPL, for example, was significantly higher in the case of the aria singing as opposed to the discrete pitch and vocalise tasks. Conversely, the aria singing yielded a significant smaller SPL range than the discrete and vocalise tasks. The total area was also significantly smaller than in other tasks.

Contrary to Emerich's results, the singer data did not indicate an increase in maximum intensity values when the context was changed from physiological to a performance setting. In fact, this was the only metric which did not demonstrate any effect of tasking. Maximum intensity levels for singers actually decreased a little when compared to the physiological case. On the other hand, singers in all voice categories increased their VRP area above 90 dB when given a performance context. Emerich concludes that this ability to produce louder phonation in a performance context could cast doubts on the proper voice function strategies of the actors. In the singer's case, the increase of $\text{Percent}_{\geq 90\text{dB}}$ does not evoke concern for the singing strategies of these singers (all professionals with many years of experience) but rather attests to successful training and vocal behavior required in performance. Significant differences for SPL_{ext} were limited to the discrete pitch task and the physiological task. Table 7b) shows this clearly. In fact, the two designs – the discrete pitch task and the vocalise – were not significantly different in any of the nine VRP metrics.

It had been hypothesized that the vocalise task, being a continuous type of task and part of the singer's daily vocal reality, would lead to enlarged singer-specific VRPs. The results obtained here lead us to reject this hypothesis. Differences between aria singing and task singing were not observed for the singing-voice specific metrics. This result speaks to the necessity for introducing two relatively new metrics, the SPL_{ext} and the $\text{Percent}_{\geq 90\text{dB}}$, as well as the importance of including a performance task design when conducting singer VRP recordings. Such findings are clinically relevant. If a patient puts forth a complaint particularly related to his/her singing voice, the clinician could opt for which VRP acquisition to prioritize. In this case, a VRP_{perf} would most likely help elucidate the problem.

These task-related aspects will need consideration for the proper documentation and understanding of the singing voice as it is used regularly by the singer. Performance task design, according to our observations, appears to be less important than the clarity and structure of the instructions. Providing the singer with a realistic voice-use context is also important.

Physiological VRP versus Performance VRP

Observing equivalent physiological and singing phonational range results, Brown *et al.* commented that in a physiological context there might be an unconscious or conscious act of safeguarding the voice and staying well within limits of vocal comfort [29]. They also put forth alternatively that perhaps the biological limits of the vocal mechanisms are similar between singing and speaking and that rather than seeing a range extension in singers, one could anticipate an increase of control throughout the vocal range (Hacki's findings of the shouting voice when superimposed to the VRP point to the same idea [31]). In a similar line of thought, Coleman had earlier postulated that "singers should traverse physiological capabilities with control and artistry" and consequently physiological and performance VRPs should not differ [36]. These contentions do not necessarily fall in line with the present results. Singers need to develop control and artistry regarding certain physiological aspects of the voice such as mechanism transitions, but might never have to employ certain areas of the voice when on stage. The physiological VRPs were significantly different from other tasks in all the metrics investigated in this study. In the case of the f_{rge} , no statistical difference was found between the physiological VRP and task 3 (vocalise). This perhaps ties in with Reich's frequency range investigations. Both of these task designs were based on a continuous gesture, either the glissando or the triad carrier and this kind of task was proven to yield larger ranges.

Despite the absence of a main task effect on the SPL_{max} metric, an increase of the percent area equal and above 90 dB could be noted as one moved from the physiological task (task 2) through to Task 3, 4 and 5 (ending with the aria). Unexpectedly, upon visual comparison, all performance tasks exhibit contours that systematically exceed the physiological one in the high rise portion of the maximum curve (note that the highest level is more or less the same for all tasks). Figure 10 depicts this important difference. For most metrics, the performance tasks are contained within the physiological VRP and therefore, this contour detail was important to report. According to statistical analysis and the illustration presented in Figure 10, the physiological VRP might miss completely some vocal capabilities that are present in a performance context.

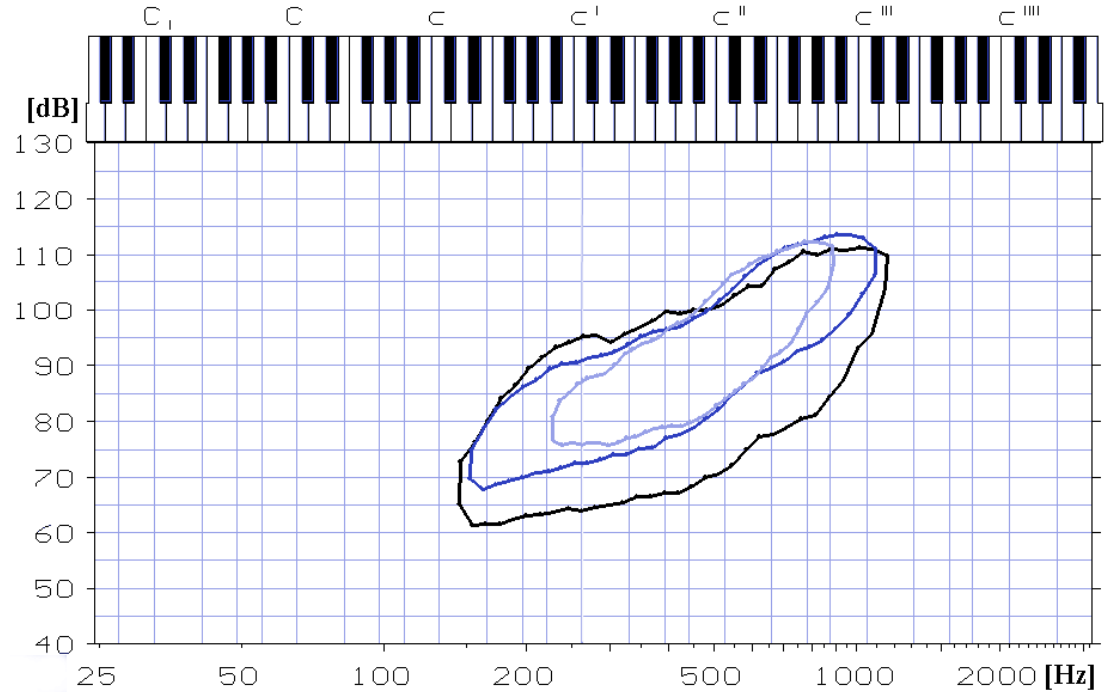


Figure 10. Physiological (in black) and performance contours (aria in light grey and vocalise in blue) for merged soprano, mezzo and contralto groups (N=30). At low levels, the performance contours are well contained by the physiological contour and even align at the low maximum curve rise. At high levels, however, the performance contours exceed the physiological contour, in the uppermost region of the maximum curve. Note however that the maximum SPL values for all VRPs is more or less the same.

The reader may recall that for the performance tasks, the increase of voicing in the higher SPLs was obtained in a studio context which limited the singer's freedom of expression, space and musicality. The context was remote from the realistic setting in which a singer performs. On-stage recordings could well lead to an even greater increase in the area equal and above 90 dB. Emerich's result of actors studied in both on-stage and in-studio monologues seems to support this [35]. The performance VRP might bring us a step closer to a more representative image of the singer's voice, while remaining distinct from the real on-stage vocal behavior.

Physiological VRPs were compared to pre-existing data sources. Figure 11 includes four different normative contours for similar groups. Although all four studies conducted physiological VRPs, there are clear differences in the phonation threshold and/or the minimum curve of the VRP. The data collected in the current experiment have the highest minimum values. When related to our counting speech data, it was found that soft phonations produced in the physiological VRP yielded similar minimum

results (recall figure 4a). For our recording of the counting tasks, subjects were asked to count very softly without whispering. This would indicate that in the physiological VRP, singers stayed in a “respectable phonation” zone instead of dropping to the bare minimum levels possible.

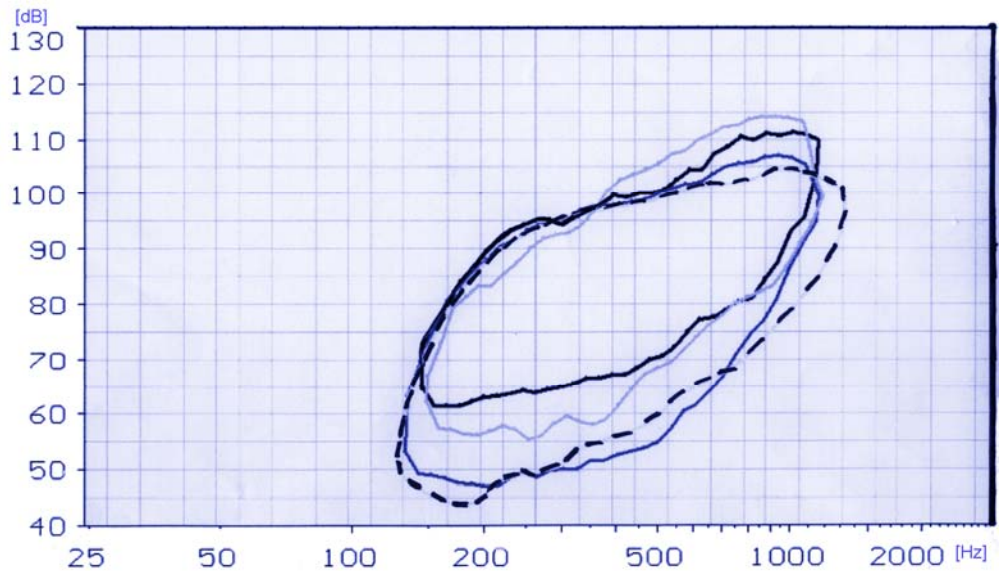


Figure 11. Contour averaging for singer groups. Data is representative of physiological VRPs. Lamarche data in dark blue, N=16, professional classical singers. Sulter data in broken line, N=42, choir singers with ± 2 years experience [1]. Pabon data in blue, N=23, classical singing students (unpublished). Hacki data in light grey, N=10, classical singers, level of skill undefined [32].

There could be two reasonable explanations for this: a procedural effect and/or a control question. Firstly, the *glissando* procedure was selected for its speedy and efficient nature; also, its non-sustained nature was believed to help the singer *not* to sing (instinctively, some subjects reverted to singing quality phonations – especially vibrato – and had to be encouraged by demonstration to abandon it). It could be that in using an ascending continuous pitch gesture, the minimum threshold could not really be obtained in a way representative of the threshold pressure. If a discrete pitch task had been performed instead, a drop of 10-15 dB might be expected. In that event, this study's data would compare better with the other contours (Sulter and Pabon used free phonation in discrete pitch task, except at the higher frequencies where usually glissandi were more easily produced). Reich's results on minimum frequency and tasking could perhaps be generalized here to minimum intensity: a fast continuous vocal gesture automatically raises sound pressure levels.

A second possible explanation for the higher thresholds in the present study might be that singers wanted to keep a certain degree of control as they performed. The

minimum levels for the physiological VRP matched those obtained for SRP (1b) where soft voice was required. Instructions were carefully formulated in regard to voice quality and task approach, but perhaps more attention should have been given to vocal control. It seems like singers might have felt uncomfortable to visit very low levels of phonation due to the instability it could entrain. A similar idea could explain certain differences observed concerning the upper contour as well. Singers tended to be cautious and needed some coaching to freely visit voice transitions. It is believed that higher intensities could be obtained since they are demonstrably present in the performance VRPs.

Group Criteria

The present study is concerned with one particular style of singing. Still, thanks to the VRP's known sensitivity to various aspects of voice and factors such as gender and training [1, 4, 6, 15-17], it could also be of interest when grouping candidates to collect VRP singer data by genres. In the present study, only female professional classical soloists were included. A similar study of female professional musical theater and commercial music could offer useful comparison material.

Technical issues

Automatic phonetographs have spread quickly within the clinical community and their practicality and effectiveness are established. However, in using these devices with the professional operatic singing voice, one needs to attend to certain details that were not necessarily relevant for manual phonetographs nor for the case of the speaker's voice. These include the dynamic range, the phonation occurrence threshold setting (is one going to include *vibrato* or not in the tasking?), the period-time variance threshold, the responsiveness of the F_0 extraction algorithm, and the required duration of phonation.

Here follows a brief summary of details that would need to be accounted for by the clinician who works with the VRP. Recording the operatic voice at a 30 cm mouth to microphone distance will result in a signal with high decibel values. This is in fact an obstacle which was often met during this data collection and which has seldom been reported. LeBorgne mentioned in passing some student singer phonations of 125 dB in the context of a VRP study using CSL equipment [20]. (She does not report any recording difficulties pertaining to the microphone or the phonetograph and furthermore uses a microphone-to-mouth distance of 15 cm). Most current phonetographs do not have the ability to register higher SPLs than 120 dB. Such high amplitude signals will be clipped. Most commercial phonetographs abide by the conventional display built for speech which ranges from 16 to 4000 Hz and from 40 to 120 dB. This might seem elementary but it nevertheless points to the necessity of creating or adopting a "singing voice interface or mode" in present day phonetographs. (For example, a separate window or interface setting could help mark the differences for the user and have

pre-settings necessary for singing voice recording). For the purpose of this study, an electrical -12 dB pad was used between the microphone and the computer's digital sound card; or alternatively, microphone to mouth distance was increased to 1 m. The signal was thus reduced by 12 dB or 10.5 dB in order to make *Phog* recordings possible and complete. These corrections were later accounted for in post-recording analysis. However, in a clinical context where singer-patients are being evaluated, a VRP program would definitely need to provide immediate proper visual feedback. The SPL limit of the instrument aside, measurement microphones used in VRP recordings of singers may need to tolerate 130 dB, for a 30 cm placement. In a clinical context a headset microphone might be preferred to a fixed microphone. It would then be imperative to select a headset with the proper voice level tolerance for singers (looking not only at saturation but also at distortion thresholds) and calibrate it adequately [38].

Despite the increasing popularity of computerized phonetographs and their capabilities to display additional voice quality information, VRP analysis remains largely focused on contours. Some work [13,14,18,39] has attended more specifically to the interior of the VRP. The VRP might offer much more information than is commonly exploited.

Conclusions

This study investigated the possible importance of recording two types of VRP when addressing the singing voice. Furthermore, the impact of task design was considered and the possible necessity of subdividing subjects into groups according to voice category was explored.

The physiological VRP was found to be different from the performance VRP. It appears important to include both types of VRPs in a singer's voice status analysis as they contribute different kinds of information. While there was no significant difference concerning SPL_{max} , it was observed that the percentage of the voice in the VRP area equal and above 90 dB increased in a performance context. Indeed the $Percent_{\geq 90dB}$ could be a sensitive metric to performance capabilities and would perhaps be more sensitive than the total area metric in the assessment of singer's voices. It is clear that if one records uniquely physiological VRPs of singers, important aspects of voice use might not be represented. The performance context or mindset seems to be key in obtaining a more representative image of the true vocal use of the singer and this seems to apply to other types of professional voice users; actors are a previously reported example.

Different task elicitation methods for the physiological VRP might greatly influence the minimum VRP thresholds. Conversely, no effect of a particular task design could be observed when investigating the performance VRP. Discrete pitch task and a more continuous gesture (vocalise) task led to similar results.

The instructions and the context suggested to the singer are perhaps more important than the particular task design in determining VRP outcomes. The hypothesis that a

vocalise task would yield more representative singing voice VRPs than that obtained with a discrete pitch task is rejected.

Finally, results did not point out any particular need to subdivide a female singer group according to voice category. This suggests that in the case of the singing voice it would be important to also consider other VRP metrics that are based not only on the contour.

All in all, it is expected that this collection of VRP data for a homogenous group of female Western opera singers could be useful and referential in understanding and analysing the female classical singing voice.

Acknowledgments

The authors recognize the generosity of the all the singers who participated in these recordings. The authors are also indebted to the Baxter and Ricard Foundation which partially funded this research. Thank you to Caroline Traube who made sound booths at the University of Montreal accessible. The authors also wish to thank Erwin Schoonderwaldt and Clara Maitre, who helped with Matlab scripts.

References

1. Sulter A M, Schutte H K, Miller D G. Differences in phonetogram features between male and female subjects with and without vocal training. *JVoice*. 1995;9:363–377.
2. Heylen L, Wuyts F L, Mertens F, de Bodt M, PH Van de Heyning. Normative voice range profiles of male and female professional voice users. *JVoice*. 2002;16:1-7.
3. Ma E, Robertson J, Radford C, Vagne S, El-Halabi R, Yiu E. Reliability of Speaking and Maximum Voice Range Measures in Screening for Dysphonia. *JVoice*. 2007;21:397-406.
4. Wuyts F L, Heylen L, Mertens F, du Caju M, Rooman R , Van de Heyning P H, de Bodt M. Effects of age, sex, and disorder on voice range profile characteristics of 230 children. *Ann Otol Rhinol Laryngol*. 2003;112:540–548.
5. Siupsinskiene N. Quantitative analysis of professionally trained versus untrained voices. *Medicina (Kaunas)*. 2003;39:36–46.
6. Hacki T. [Vocal capabilities of nonprofessional singers evaluated by measurement and superimposition of their speaking, shouting and singing voice range profiles]. *HNO*. 1999;47:809–815.
7. Chen S H. Voice range profile of Taiwanese normal young adults: a preliminary study. *Zhonghua Yi Xue Za Zhi (Taipei)*. 1996;58:414–420.
8. Heylen L, Wuyts F L, Mertens F, de Bodt M, Pattyn J, Croux C, Van de Heyning P H . Evaluation of the vocal performance of children using a voice range profile index. *J Speech Lang Hear Res*. 1998;41:232– 238.

9. Ikeda Y, Masuda T, Manako H, Yamashita H, Yamamoto T, Komiya S. Quantitative evaluation of the voice range profile in patients with voice disorder. *Eur Arch Otorhinolaryngol*. 1999;256 Suppl 1:S51–S55.
10. Böhme G, Stuchlik G. Voice profiles and standard voice profile of untrained children. *JVoice*. 1995;9:304–307.
11. LeBorgne WD, Weinrich BD. Phonetogram changes for trained singers over a nine-month period of vocal training. *JVoice*. 2002;16(1):37–43.
12. Hunter E J, Švec J G, Titze I R. Comparison of the produced and perceived voice range profiles in untrained and trained classical singers. *JVoice*. 2006;20:513–526.
13. Roubeau B, Castellengo M, Bodin P, Ragot M. [Laryngeal registers as shown in the voice range profile] *Folia Phoniatr*. 2004;56(5):321–33.
14. Lamesch S, Doval B, Castellengo M. Phonetograms of laryngeal source parameters for different vowels and laryngeal mechanisms. *J Acoust Soc Am*. 2008;123(5):3243.
15. Gramming P, Sundberg J. Spectrum factors relevant to phonetogram measurement. *J Acoust Soc Am*. 1988;83:2352–2360.
16. Gramming P, Åkerlund L. Non-organic dysphonia. II. Phonetograms for normal and pathological voices. *Acta Otolaryngol*. 1988;106:468–476.
17. Pedersen M F. A longitudinal pilot study on phonetograms/voice profiles in pre-pubertal choir boys. *Clin Otolaryngol Allied Sci*. 1993;18:488–491.
18. Lamarche A, Ternström S, Hertegård S. Not just sound: supplementing the Voice Range Profile with the singer's own perceptions of vocal challenges *Logoped Phoniatr Vocol*. 2008. DOI: 10.1080/14015430802239759.
19. Schutte H K, Seidner W. Recommendation by the Union of European Phoniaticians (UEP):Standardizing Voice Area Measurements/Phonetography. *Folia Phoniatr*. 1994;35:286–88.
20. Lamarche A, Ternström S. An Exploration of Skin Acceleration Level as a Measure of Phonatory Function in Singing. *JVoice*. 2008;22(1):10–22.
21. Awan S N. Phonetographic profiles and F0-SPL characteristics of untrained versus trained vocal groups *JVoice*. 1991;5:41–50.
22. Awan S N. Superimposition of speaking voice characteristics and phonetograms in untrained and trained vocal groups. *JVoice*. 1993;7:30–37.
23. Coleman R F. Sources of variation in phonetograms. *JVoice*. 1993;7:1–14.
24. Schultz-Coulon H-J. *Stimmfeldmessung*. Springer-Verlag Berlin, Heidelberg, New York. 1990. 1–63.
25. Titze I R. Acoustic interpretation of the voice range profile (phonetogram). *J Speech Hear Res*. 1992;35:21–34.
26. Nawka T, Seidner W, Cebulla M. Experiments on intentionally influencing the singer's vibrato in *Proceedings of the Stockholm Music Acoustics Conference*. 1993;221–225.
27. Rosen C A, Murry T. Voice handicap index in singers. *JVoice*. 2000;14:370–7.

28. Drew R, Sapir S. Average Speaking Fundamental Frequency in Soprano Singers with and without Symptoms of Vocal Attrition *JVoice*. 1995;9(2):134-141.
29. Brown W S, Morris R J, Hicks D M, Howell E. Phonational Profiles of Female Professional Singers and Nonsingers *JVoice*. 1993;7(3):219-226.
30. Kitzing P. *Glottografisk frekvensindikering. En undersökningsmetod för mätning av röstläge och röstomfång samt framställning av röstfrekvensdistributionen*. PhD thesis, Lund University 1979.
31. Hacki T. Comparative speaking, shouting and singing voice range profile measurement: physiological and pathological aspects *Logoped Phoniatr Vocol*. 1996;21:123-129.
32. Sulter A M, Wit H P, Schutte H K, Miller D G. A structured approach to voice range profile (phonetogram) analysis. *J Speech Hear Res*. 1994;37:1076-1085.
33. Hacki T, Frittrung B, Zywietz Ch, Zupan Ch. Verfahren zur statistischen Ermittlung van Stimmfeldgrenzen - Das Durchschnittsstimfeld *Sprache-Stimme-Gehör*. 1990;14:110-112.
34. Reich A R, Frederickson R R, Mason J A, Schlauch R S. Methodological variables affecting phonational frequency range in adults *Journal of Speech and Hearing Disorders*. 1990;55(1):124-131.
35. Emerich K A, Titze I R, Svec J G, Popolo P S, Logan G. Vocal range and intensity in actors: a studio versus stage comparison. *JVoice*. 2005;19:78-83.
36. Coleman R F, Mabis J H, Hinson J K. Fundamental frequency-sound pressure level profiles of adult male and female voices. *J Speech Hear Res*. 1977;20:197-204.
37. Bunch M, Chapman J. Taxonomy of Singers Used as Subjects in Scientific Research. *JVoice*. 2000;14(3):363-369.
38. Cabrera D, Davis P, Barnes J, Jacobs M, Bell D. Recording the operatic voice for acoustic analysis, *Acoust. Aust*. 2002; 30 (3): 103-108.
39. Pabon J P H. Objective acoustic voice-quality parameters in the computer phonetogram. *JVoice*. 1991;5:203-216.