Putting the Singing Voice on the Map

Towards Improving the Quantitative Evaluation of Voice Status
in Professional Female Singers

ANICK M-J LAMARCHE

Doctoral Thesis
Stockholm, Sweden 2009

Abstract

Diagnostic and evaluative methods used in voice care are mostly designed for the speaking voice, and are not necessarily directly applicable to the singing voice. This thesis investigated the possibilities of fine tuning, improving and quantifying the voice status assessment of the singer, focusing especially on the Western operatic female voice.

In Paper I, possible singer-specific Voice Range Profile (VRP) characteristics and tasks were explored and VRP data for 30 professional female Western opera singers was collected. Vocal productions were controlled for a physiological VRP (VRP_{phys}) and for a stage performance context (VRP_{perf}) and outcome differences were identified. Task design was critical for the VRP_{phys} but had very little effect on the VRP_{perf}. Significant voice category differences (between soprano,mezzo-soprano and contralto) were limited to frequency-related metrics. Two new VRP metrics, the area above 90 dB ($Perc_{\geq 90dB}$) and the sound pressure level extent (SPL_{ext}), were found to be key metrics to the study of VRPs for singers.

Paper II investigated, in conjunction with the VRP, whether the sound pressure level (SPL) or the skin acceleration level (SAL) was more correlated to the subglottal pressure (P_s). SAL was much less F_0 dependent than SPL and facilitated the interpretation of VRP data. However, the correlation between SAL and P_s was found to be weaker than that between SPL and P_s.

Papers III and IV explored the mapping of self-perceived impairment-related difficulties into the VRP. A modified phonetograph was tested first with a healthy singer population and then with a singer-patient group. Subjects used a button device to communicate their self-perceptions while singing, and were consistent in task replications as well as across different tasks. Healthy singers pressed mostly at the extreme limits of the VRP, where loss of vocal control could be expected and their presses were mostly concentrated on the periphery of the VRP area. Singer patient button- press patterns were distinct from patterns observed in healthy singers. Singer patients pressed mainly inside the VRP boundaries, in the higher range and at intermediate intensities.

In Paper V, the Voice Handicap Index for singers was translated and adapted to Swedish (Röst Handikap Index för sångare or RHI-s). The questionnaire was found to be a reliable and a valid instrument. High correlations between general perceptual patient VAS ratings and the questionnaire scores underscored the instrument's internal coherence. Overall, patient scores (including subscales) were significantly higher than healthy singer scores. The results showed implicitly the necessity and usefulness of adapting clinical procedures to specific patient populations.

Together, the results of these five papers can ultimately be of value to voice clinicians who are treating singers. The results obtained also contribute to the understanding of the singing voice and underline the importance of properly documenting the singing voice.

Keywords: Voice Range Profile, Phonetogram, Singing voice, Performance, Clinical assessment, Health, Voice disorder, Self-perception, Proprioceptive feedback

Acknowledgments

During the course of my four doctoral years in Sweden, I have had the privilege to work aside a very unique and inspiring man with a great vision, a funny funny sense of humour, the widest of horizons, versatility, astute musicianship, positivism and accessibility. From day one to day thesis-final, I was met with continuous support and wisdom. There are few people that one encounters in a lifetime who are versed in so many things, open and knowledgeable whilst down to earth and humble. I have tried to do my best to learn as much as possible from this precious collaboration that has taught me everything, from (the importance and the role of parentheses), microphone clipping and troubleshooting, the various ways of looking at data and asking questions, to cultural capsules in the form of rhymes and bizarre Broadway lines. *Sten Ternström* is no doubt one of those amazing people -and I am truly thankful for the chance I have had to meet him this early in my career! The journey in an engineering school and context was a challenge for a singer, yet, *Sten* was particularly skilled in smoothing out the transition from the performance world to the scientific one.

Others have also directly impacted the work presented here. I was especially lucky to collaborate with knowledgeable as well as fun-loving individuals!

I am indebted to *Peter Pabon* for his genuine interest and passion for all things VRP and many entertaining back and forth emails and telephone conversations. Naturally, our collaboration is precious to me and his experience in the craft of VRP recording and signal processing has been an enormous help in my endeavours.

I am grateful for *Dominique Morsomme*'s belief and sincere interest in the button-VRP project. It was so exciting (and somewhat relieving) to meet someone who had similar questions and thoughts regarding the singing voice! Her many

I am also blessed to have two Swedish "families": the *Irminger-Street* and the *Markstad*. They have provided me a home away from home, helped me out in various ways (including the impossible: Stockholm lodging). I have been extremely "spoiled" and have been included in their circles as one of their own. Thank you for caring so much!

My parents and family (the biological version): I save keep the most tender appreciation for you. Your unlimited patience, your tolerance of the soprano diva temperament that can surface at times and your encouragement and words of wisdom throughout the many hops and turns in my life (especially during the course of the last four years), have helped me tremendously. Votre soutien et votre amour inconditionel contribue énormément à qui je suis aujourd'hui et m'ont permis de realiser ce projet d'envergure gigantesque et souvent envahissante. Merci!

My grand-parents: malgré qu'il leur soit souvent difficile de suivre mes maintes péripéties, mes choix et mes cheminements de vie (souvent lointains), vous me surprenez et m'impressionez tant par votre capacité de compréhension. Je me sens réellement choyée de vous avoir à mes côtés...malgré l'outre-mer.

Friends that believe in me and have tolerated all the whining and the talking about thesis writing. My music friends that might not really understand but have accepted and respected my choice to leave music for science. The friends that I have found during the course of my Swedish experience: many a times they have proved that, finding a friend in Sweden is finding a friend for life. Childhood friends with whom I can still laugh and cry have shown that long distance relationships can work. This one is a solid one; merci *Julie, Lynne, Denise* et *Marie-Hélène*!

Petko, my dearest friend and partner in life who made sure I would make it through alive and in one piece. So much of this work has been possible because of you: crash course on basic signal processing, computer loans, logarithmic exercises, debugging the Anick version of Matlab code, breakfast making (that is worth half the thesis right there) and lending an attentive ear to my daily doses of frustration, stress, fear, throws of voice passion and technological downs. Голям Благодаря мили мой!

Abbreviations and Definitions

ANOVA: Analysis of variance

A-weighting: A curve used in sound pressure level measurement which has filter characteristics that modify the frequency response so that it approximately follows the inverse of the equal loudness contour in low level sounds (about 40 phon).

Bonferroni test: A post hoc statistical test used to examine multi-comparison when an analysis of variance shows significant results.

BMI: Body Mass Index. This index is generally based on the ratio of weight (in kg) to height squared(in meters). $BMI = kg/m^2$. A BMI between 18.5 and 24.9 designates a healthy weight, overweight is usually situated between 25 and 29.9 while obesity is defined by BMI > 30.

CCM: Contemporary Commercial Music. A referential term for all non-classical singing genres.

Closed-ended: A two-pole question structure which restricts subject responses to stated alternatives or to "yes/no". Such questions are also known as dichotomous or saturated type questions

C-weighting: The C-weighting curve approximately follows the inverse of the 100 phon curve. It is often used as an equivalent to linear weighting.

DC: Direct Current, used here to denote data acquisition with a frequency response down to 0 Hz (DC).

DSI: Dysphonia Severity Index. Established in 2000 as an objective and quantitative correlate to perceived voice quality [183]

F_0: Fundamental frequency. The repetition rate of vocal fold oscillations, in cycles per second (Hz).

Fd: Functional dysphonia.

***ff*:** Very loud. Musical symbol for a high dynamic level.

F$_0$-F1: The tuning of the fundamental frequency to the first vocal tract resonance. An acoustic strategy that is present mostly in high female singing.

HNR: Harmonic-to-noise ratio.

I: Intensity. The acoustic power impinging in a given direction on a unit area (a vector value). Its magnitude is given in RMS watts per square meter in the SI-metric system. The term "voice intensity" is often used in the voice literature, even though it is usually the sound *pressure* (a scalar value) that has been measured, in RMS pascal. Ideally, the total radiated vocal *power* should be measured, but this is technically difficult to do. For most voice recordings, the distinction between intensity and pressure is of little consequence, since the standard level references for intensity and pressure have been chosen so as to give the same magnitudes on a decibel scale. The SPL measure can only be partially representative of total radiated power, because of the directivity of the voice The standard reference intensity, I$_0$, is 10^{-12} $Watt/m^2$ (see also SPL below).

Laryngeal Vibratory mechanism: a term which designates a specific glottal configuration characterised by the shape of the vocal folds and by the muscular tension at play and which has been suggested in lieu of register [132]. Typically, the most frequent mechanisms in VRP recording are M1 and M2 (corresponding to the quality registers of "chest/modal" and "falsetto/head" voice). This aspect of singing continues to be a debate matter in both pedagogical-performance and scientific circles.

LPR: Laryngeal Pharyngeal Reflux. An acid back flow in the oesophagus that enters the throat and voice box due to upper and lower esophageal sphincter malfunction.

MANOVA: Multivariate analysis of variance.

Messa di Voce: Italian term meaning "placing the voice". This has become a universal vocal exercise which originates from the old Italian schooling tradition. A note is sung very quietly and is gradually and smoothly made louder and then similarly made quiet again.

mf: Medium loud. Musical symbol for an intermediate dynamic level.

MPT: Maximum phonation time.

Open-ended: Question formulation which requires the respondent to formulate an answer in his/her own words (usually entails a descriptive answer).

Proprioception: In latin, proprius, meaning "one's own," is combined to perception — to refer to the human senses. Proprioception is a sensory modality that provides feedback solely on the status of the bodys' internal events. It

is the sense that indicates whether the body is moving with required effort, as well as where the various parts of the body are located in relation to each other.

Passagio: An Italian term used in the Western Opera Singing tradition to designate the transition area between laryngeal mechanisms. This term along with the area it designates remains a subject of debate in scientific forums.

P_{thresh}: Phonation Threshold Pressure, mathematically defined by Titze[169] as $P_{\text{thresh}} = 0.14 + 0.06(F/F_{M0})2$.

pp: Very soft. Musical symbol for a low dynamic level.

Phonetograph: The instrument (either software and/or hardware) that is used to record a VRP.

P_s: Subglottal pressure, herein estimated by the intraoral pressure during p-occlusions, and measured in centimeters of water column relative to atmospheric pressure.

Register: Musical notes which are sung with the same quality.

R.E.G.W.R.: Ryan-Einot-Gabriel-Welsch Range comparison test.

RHI-s: "RöstHandikappIndex för sångare", the Swedish adapted version of the Voice Handicap Index for singers

RMS: Root mean square

SAL: Skin acceleration level, mainly a measure of tissue vibrations. In Paper I, it is recorded at the jugular notch and at the sternum bone.

ST: Semitone. A logarithmic measure of frequency ratios; a semitone is 1/12 of an octave and represents the frequency ratio of $\sqrt[12]{2:1}$.

SPL: Sound Pressure Level relative to $20\mu Pa$. All measurements of SPL were performed at 30 cm microphone-to-mouth distance, unless otherwise specified.

SRP: Speech Range Profile. This type of recording is based on running speech

SD: Standard deviation

SVS: Singing voice specialist. A professional expert of the singing voice who is qualified to retrain singers recovering from illness or injury.

UEP: Union of European Phoniatricians

VAS: Visual Analogue Scale. This psychometric scale is used to measure subjective responses. It is a horizontal line, 100 mm in length, anchored by word descriptors at each end

VC: Vital capacity.

VRP: Voice Range Profile. In this work the VRP was always recorded with a computerised phonetograph.

VRP$_{phys}$ Physiological voice range profile. This type of VRP charts the physiological vocal limits of an individual. Soft phonations (the lower contour) correspond to the physiologic minimum intensities for each frequency. In turn, these minimal intensities can be related to phonation threshold pressure. Subjects are also encouraged to visit their loudest phonations. Voice quality is normally disregarded.

VRP$_{perf}$: Performance voice range profile. This type of VRP recording, similar to the "musical range profile" comprises not only voice quality but also performance relevant use of the singing voice. The singers determine the pitch and the dynamic limits with respect to what is musically acceptable to them in a performance context.

Western Opera: A musical dramatic work developed in the 17th century Italy in which the actors sing some or all of their parts and where many art forms are united; music plays a dominant role. The Western opera stylistically follows the classical music traditions of Europe and North America. Western has become misleading in that the notion of the Western world has changed appreciably with globalisation.

White noise: A noise with a constant sound energy within equally wide frequency bands.

List of Publications

The papers will be referred to by their Roman numbers.

Paper I:

Lamarche A, Ternström S & Pabon P. (2008) The Singer's Voice Range Profile: Professional Western Opera Female Soloists. *Journal of Voice.*
In press, Copyright Elsevier December 2008. (Printed with permission).

Paper II:

Lamarche A & Ternström S. (2006) An Exploration of Skin Acceleration Levels as a Measure of Phonatory Function in Singing Voice. *Journal of Voice,* 22(1):10-22.
Url: http://dx.doi.org/10.1016/j.jvoice.2006.08.005. (Printed with permission).

Paper III:

Lamarche A, Ternström S & Hertegård S. (2008) Not Just Sound: Supplementing the Voice Range Profile with the Singer's Own Perceptions of Vocal Challenges. *Logopedics Phoniatrics and Vocology,* 13:1-8.
Url: http://www.informaworld.com/10.1080/14015430802239759. (Printed with permission).

Paper IV:

Lamarche A, Morsomme D & Ternström S. (2008) Not Just Sound II: an Investigation of Singer-Patient Self-Perceptions Mapped into the Voice Range Profile. *Journal of Speech, Language and Hearing Research* (in review).

Paper V:

Lamarche A, Westerlund J, Verduyckt I & Ternström, S. (2009) The VHI Adapted to Singers: a Swedish Translation and Adaptation. *Logopedics, Phoniatrics and Vocology* (in review).

Other related published work:

Pabon P, Lamarche A & Ternström S. (2009) Fourier Descriptor Analysis and Unification of Voice Range Profile Contours: Method and Applications. *Journal of Speech, Language and Hearing Research* (in review).

Thesis Contributions

Paper I (VRP singer): Author AL carried out the major part of the work (investigation protocol, recording, analysis and write-up). PP helped with the analysis methodology. ST assisted in the editing of the manuscript.

Paper II (SAL): Author ST proposed the project. Authors ST and AL planned the investigation. Author AL carried out all measurements and recordings. Authors AL and ST conducted a joint analysis and co-authored the report.

Paper III (Button-healthy): Author ST instigated the investigation idea. Author ST developed the testing device. Author AL planned the investigation, carried out the measurements and recordings and was responsible for the major part of the analysis and write-up process. SH assisted in the investigation. ST assisted in the editing of the manuscript.

Paper IV (Button-patients): Author ST supplied technological support. Author AL performed the major part of the work. Author DM helped with the recordings, providing both the recording location and participants. DM also contributed in a section of the paper's discussion. Author ST assisted in the editing of the manuscript.

Paper V (RHI-s): This project was initiated and realised by author AL. Author IV assisted in the translation process and author JW carried out the statistical analysis. Author AL was responsible for writing the report, JW helped with revisions. Author ST helped with the editing of the manuscript.

Contents

List of Figures

Part I

The Singing Voice

Chapter 1

Introduction

"It's got to be perfection"
Dame Nellie Melba

See yourself for a moment, as a well-known opera diva with a fully booked agenda for the upcoming five years. Tired and experiencing insurmountable levels of stress, your voice suddenly begins to show some signs of instability, fatigue and even perhaps injury. Who do you turn to? What steps do you take to solve the problem? In this case, the diva most likely turns to one of the world's top voice experts, who has accumulated years of experience and has developed a unique set of tools in dealing with the singing voice.

But what happens in the case of the opera debutant or even the voice student? Most risks for voice disorders exist in the training and early stages of a career. These younger singers might not have the practical tools or even the knowledge to access the world expert sought by our opera diva and rather, might need to rely on the help of a general voice clinician. In such cases, it would be a great asset for this voice clinician to have access to singing-voice clinical resources. This thesis contributes to establishing such resources.

1.1 Background and Problem

When singers have voice-related problems, the experience tends to take on dramatic proportions. In *Professional Voice: The Science and Art of Clinical Care*, 2nd edition, 1997 [133], Sataloff cautions that announcing a voice pathology to a singer is comparable to announcing a life-threatening illness. Openly disclosing a

vocal disorder seems to remain taboo in the singing world. Indeed, misconceptions about vocal disorders abound in the performing arts community, and are often accompanied by the unfounded idea that one's career could be at stake. Evidently, this account of singers and vocal disorders attests to the difference between vocal concerns of a singer and that of non-singers. A singer's vocal concern can be further understood when a parallel is drawn between elite athletes and professional singers [9, 24, 97, 134, 94, 125]. Both these populations rely on performance abilities that must approach perfection. Since the performance and function dichotomy is a matter of accuracy, control and flexibility, athletes and singers have very low tolerance thresholds for subtle changes of state. These changes entail great consequences for performance aptitude. This ties in well with the notion of skill, where achievement is maximised and random variation is minimised.

The field of voice science is now rapidly expanding, as is the understanding of speech and voice mechanisms. As early as 1994, Cleveland [26] surveyed the then previous 25 years of singing-voice research which he referred to as most productive and demystifying (non-linear source filter theory, the concept of spectrum resonant peak cluster, voice register/mechanism understanding, flow measurements, vocal fold vibration modeling, laryngeal musculature histology and dissection, singing synthesis, Voice Range Profile to name a few).

Surprisingly, despite such a tremendous gain in voice knowledge, a discrepancy seems to persist between theoretical knowledge and its application in the voice clinic. Although the prevalence of vocal disorders is highest for vocal performing artists, and despite the fact that singers are recognised as high priority patients, the clinical voice assessment of these patients is still heavily defined according to speech voice function. Growing awareness of the need for specific treatment of the singing voice, however, has brought light to terms like the "singing-voice specialist" (SVS) coined in the 1980s [125]. Encouragingly, the voice research community continues to strive for more uniformity and credence in singing voice support systems. The importance of voice behaviour and the type of voice use has gained much more attention in the scheme of voice evaluation and treatment. For instance, recent literature is increasingly concerned with fulfilling the vocal performer's needs. *Vocal Arts Medicine* (Benninger, Jacobson & Johnson; 1994), *Professional Voice : The Science And Art of Clinical Care* (Sataloff; 1997), *The Singer's Voice* (Benninger & Murry; 2008), *Care of the Professional Voice* (Davies & Jahn; 2004), *Care of the Professional Voice* (Irving, Epstein & Harries; 1997) are examples of textbooks specifically geared to help clinicians caring for high-performance voice users, often singers.

This growing body of voice performance literature has a broad scope and is often quite general. Few reports exist in which clinical measures and procedures are reviewed specifically in relation to the singing voice. The work of Carroll *et al.* [24] is an example of a rare attempt to achieve representative respiratory and glottal-efficiency normative measures of the singer population. That study demonstrated considerable differences between singer and non-singer measures, and led to the advocacy of separate normative data collection for the evaluation of singers.

Elias *et al.*, [41] also point in a similar direction, concluding that normative baseline data, in this case strobovideolaryngoscopy, are needed for the proper evaluation of professional singers. In 1992, Klingholz clearly commented the need, in phoniatrics, to establish distinctions amongst three groups: the vocally trained (singers), normal (healthy) and pathological voices [83]. Many have concentrated on examining differences between singers and speakers without necessarily addressing their clinical implications. A tailored assessment of the singing voice is crucial to the design of effective rehabilitation. Despite the advocacy of this last statement, there is nonetheless a paucity of *quantitative* singing-voice data, leaving few evidence-based resources for singing-voice therapy or treatment programmes.

1.2 The Singing Voice in the Clinic

When Baken in 1987 first published a textbook of clinical measurement of speech and voice [6], the world of voice was given a great reference tool for the objective assessment of voice. This book's innovation was that it included thorough overviews of various equipment and test methods together with examples of results and norms to better enable the comprehension as well as the comparison of evaluation procedures. At present, a surge of interest for the voice and its disorders has led to an explosion of the literature and an increased documentation of the voice. Thanks to the progress of technology, many more possibilities exist in objectifying and evaluating the vocal instrument. Procedures that once were considered inaccessible are now simplified, automatic and more reliable.

Nevertheless, it is the author's general impression that many of these resources are not fully exploited in the voice clinician's work. Even when resources are in place, some clinicians seem ambivalent or perhaps even intimidated with respect to the evaluation of voice. The considerable amount of available voice and clinical care documentation seems unsuccessful in demystifying the details of clinical practice and in assigning more deserving room to voice and voice disorder study in pertinent educational programmes. Belhau & Oates attest to the above in their response to the lack of gold standards and uniformity in voice care practice[8].

It is not the scope of this dissertation to elaborate on the details concerning clinical voice procedures, instruments and measures entailed in the complete assessment of the speaker's voice. This information can be conveniently found in current textbooks which thematically focus on voice, voice diagnostics and voice disorders, as well as the overall aims and evaluative procedure of voice care [6, 34, 156].

Rather, this work is especially concerned with the voice-care situation with respect to the singing voice. If some uncertainty is found regarding the speaker's evaluation, the situation is even more precarious in respect to the singing voice. The clinical evaluation and management (medical, behavioural, and environmental) of singer patients share many aspects of the typical speaker evaluation. Nevertheless, there are some areas of consideration that could make a difference in one's understanding of the vocal complaint, and consequently one's choice of rehabilitation

and treatment. In their publication of best-practice guidelines, Belhau & Oates [8] clearly point out the need to develop specialised protocols suitable for specific populations, namely singers. Since there are established (or at least documented) evaluation procedures in addressing speech, it could be of interest to achieve a similar result for the singing voice evaluation.

Because the knowledge and experience of singing-voice care is relatively new and reserved to field experts, surveying these experts' opinions and clinical approaches was deemed an interesting exercise. A questionnaire, containing seven open-ended questions and room for further commentary, was distributed to 12 established SVS's throughout the world (United States, Australia, Sweden and Belgium). Nine respondents provided their professional opinions and insights. The questionnaire surveyed different aspects of the SVS' work with singers. Furthermore, opinions were solicited on existing resources and possible existing shortfalls in the current voice care system. In what follows, the questions and the corresponding answers are summarised.

1.   What is the main difference between a speech patient and a singer patient?

This question was posed to generally assess the clinical distinctions that are currently made between a speech patient and a singer patient. The question was kept intentionally broad in order to allow various kinds of differences to arise. As expected, answers were diverse. Despite a unanimous affirmation of a difference, none of the responses agreed on the nature of this patient difference. Some responses could be thematically categorised according to vocal differences, while the remaining answers were focused on the psychological-/career-based needs of the patient. In fact, the responses were equally divided in this regard. In the vocal differences group, answers touched on breath management, sound level, vocal control, frequency and voice quality. The degree of proprioception was also mentioned. Some of the responses did not specify differences per se, but rather described the need for more in-depth approaches to pitch and power ranges as well as vocal/laryngeal flexibility and vibrato. The answers of the second group dealt with the performance aspects of voice (stress, nervousness, stage conditions), the career realities (time press), affective sensitivity, motivation and goals and overall vocal understanding.

2.   Do you follow a specific protocol/routine in your assessments? Is it the same for singer patients? What does it entail?

All of the respondents reported an adherence to a specific and, for the most parts, unchanging patient assessment protocol. In the eight responses obtained (one individual did not work with non-singers), there was no perfect agreement among protocols (with the exception of two singing voice specialists (SVS) working in the same clinic). Only one formalised protocol was mentioned: the Estill Voice Training Protocol. Overall, the given protocol details included: patient history (medical,

case and social); maximum phonation time; vital capacity; the quotient of these two former variables; P_s; F_0, as well as extreme range values in frequency and in intensity (in reading, and in continuous and sustained tone contexts); jitter; shimmer; DSI and LPR related questions; VHI; scale passages and songs (according to ability); palpation of neck musculature; larynx position identification. One can thus conclude that most protocols encompassed laryngeal, aerodynamic, acoustic and perceptual, and in a few instances, biomechanical measures of voice.

If this questionnaire item had been limited to the two first sub-questions, important information might have been overlooked. Indeed, even though the initial responses did not differentiate protocols between non-singer and singer patients, some differences were noted in the protocol details listed by the respondents. Some SVS's (3/9) mentioned the additional inclusion of the Voice Range Profile (VRP) in the case of a singer patient while others (2/9) specified a particular attention to the singer's VRP recording (filling the complete area instead of contours only). Vibrato analysis was also listed (3/9) as well as the singing of repertoire to examine technique (4/9). Finally, the participating Belgian clinicians included the singer-adapted VHI as an integral part of the protocol (in the case of a singer patient).

These last protocol variations are perhaps not as formalised as the protocol in place in the working environment of the clinician and thus, no "formal" protocol differences between singers and non-singers are elucidated in initial responses to question 2.

3. *According to you, what is important in the assessment of a singer patient?*

Given that certain protocols in place are respected (dictated by the work place), yet might not fully correspond to the clinician's own opinions of what is instrumental in the singer patient's evaluation, this particular question was formulated to further investigate the important considerations involved in the singer patient's assessment.

Responses were somewhat redundant in that they mostly elaborated the aspects mentioned in question 2. This was interpreted as a confirmation of protocol suitability and a positive outcome.

One novel detail weighed heavily in all of the responses. Environmental and profession related economical factors as well as the singer's opinion appeared also to be important in the assessment of the singer. One respondent specifically identified the importance of compensatory behaviour examination. Only one respondent underscored the necessity of singing technique knowledge to understand the vocal loss at stake. Finally, two respondents also mentioned the singer-specific issue of emergency contingencies.

4. *How do singer patients typically respond to clinical measures?*

The aim of this question was to investigate the informal response of singer patients regarding submitted tests and measurements. While question 3 brought forth the clinician's own opinions concerning the singer's evaluation process, question 4 aimed

at exploring the clinician's perceptions of the singer patient's reaction to the evaluation.

All nine answers were categorised in four clinician-perceived singer-patient responses: curiousity, open mindedness, sensitivity and concrete expectations. All in all, the clinicians' responses mostly indicated positive singer-patient reactions. The majority of the respondents also denoted a certain fragility of the singer patient (nervousness and intimate relation with the vocal instrument) .

5.   *Are singer-patient measurement results comparable to those of non-singers? (ex: max. phonation time, subglottal pressure at low-med-high frequencies, vital capacity, etc.) How do they differ?*

Since the available literature concentrates mainly on speech, the normative data referred to in the practical appraisal of voice disorders is also often based on the vocal abilities of non-singers. For example, Baken included certain sporadic data concerning the trained voice and very little explicit normative singing voice information [6]. For the present questionnaire, it was interesting to establish whether clinicians, in their daily work with singer patients, encountered differences in measurement outcomes. If so, these differences would need to be identified. Two respondents could not answer this question. Six of the remaining respondents did confirm that singer results typically exceeded that of non-singers. Singer patients were noted to have increased VC, increased P_s, greater VRP area, longer MPT (MPT>30 seconds), the marked effect of vibrato, affected tremor indices, and different H/N ratio (in the case of baritones). One respondent explicitly listed singer-patient differences encountered with the use of the CSL MDVP system. Another respondent clearly noted the inapplicability of existing normative data. Finally, one respondent reported equivalent results for singers and non-singer in matters of the speaking voice.

6.   *Do you notice some gaps (shortfalls) in the evaluation of singer patients? If so, describe your thoughts here.*

This question was perhaps misformulated as it elicited somewhat defensive responses that largely motivated the completeness of personal approaches. However, it was not the question's intent to incite the clinician to ponder on the suitability or even the credibility of their chosen patient approach. Rather, the question served to evoke perceived general weaknesses of the current voice care available for singers. Some of the responses obtained provided a broader analytical view of the voice care system. As opinions varied from one respondent to the other, the comments are listed in what follows:

- too much attention is devoted to the vocal folds and the dynamic aspects of the voice are often neglected

- the visualisation of the larynx remains quite limited (a 3D real-time action image would be optimal)

- it would be useful to obtain real-time patient feedback during tasking and in combination to laryngeal visualisation

- Voice categories and singing genres are often not well accounted for in protocols

- the necessity of both rigid and flexible scope examination (in the case of the singer patient) is not well understood by some ENTs

- a persistent lack of understanding and coordination of the potential devastating consequences of certain non-vocal issues (reflux, allergies, dental work, jaw issues and non-vocal fold surgery)

- the tendency to overlook the performance context and to consider sustained phonation only

- too little focus on the speaking voice of the singer

7.   *Are you aware of any clinical adaptations of protocol, methodology and/or equipment to the singing voice? If so, briefly describe them here.*

This last question aimed at surveying the existing singing-voice specific and clinical relevant tools pertaining to singer-patient care. SVS's were deemed to be inclined in being best informed on the singing voice relevance of various tools. The Australian respondent was not aware of any such adaptations in use in Australia. Another respondent simply did not answer this question. The remaining responses (7/9) addressed a collection of singer normative data used with the CSL MDVP system, the adaptation of the Estill Voice Training Protocol and the development of voice evaluation and treatment software that take the singing voice into account (Voice Evaluation Suite & Virtual Voice trainer), the work on learning approaches by D. Roth and K. Verdolini, the VHI adapted for singers, the patient history questionnaire elaborated by Sataloff, and the sophisticated equipment found in certain voice laboratories.

In summary, a difference between singer patients and non-singer patients was confirmed by all respondents. Many aspects, related not exclusively to the voice, have bearing on the difference between singer and non-singer patients. In view of this definitive difference, it might seem surprising that evaluation protocols remain essentially the same. Indeed, respondents all confirmed the good suitability of current protocols in evaluating the singer's voice status. However, in defining protocol procedures, many indicated that they expanded or slightly modified certain aspects of the protocols in their evaluation of the singer patient. As question 3 illustrated well, existing protocols do suitably evaluate the voice status. The main difference lies in the interpretation of the evaluation results. This interpretation needs to be supported by the right resources and there seems to be a shortage of singer-specific

information and adaptations. The holistic and dynamic aspects of the singing voice are seemingly too often neglected and deserve more recognition. Encouragingly, some initial work has been performed in fine-tuning the singer's voice care, yet, the present responses indicate the need for further such developments.

All in all, the feedback gained with this brief questionnaire revealed many valuable aspects of the clinical realities in respect to the singing voice. The singing voice specialists' standpoints confirmed the importance of differentiating the singer patient's needs from those of the non-singer patient in order to improve singer patient care. Interestingly, there was no emphasis on expertise levels inherent to quality singer-patient care. The requirement of personal singing expertise (although it might offer a great advantage), is perhaps an unfounded belief and most likely a consequence of ambiguous standards and lack of referential data (as suggested by Belhau & Oates [8]). The latter deficiency was also addressed in some of the respondents' comments. This doctoral project finds inspiration, motivation and support in such findings.

1.3   Thesis Objectives

The work presented here is largely focused on the improvement of the evaluation, and thereby, the rehabilitation and care of the singer patient. The overall objective was to improve support for subsequent evidence-based studies. Scientific studies of the singing voice are often met with scepticism, because they usually fail to use subjects of the highest proficiency. In voice experiments, terminology such as "trained voice" and "professional voice" has been employed to designate quite a variety of voices, not always including singers. Voice classification systems such as the taxonomy of singers [22] or the clearly defined level scheme that Koufman proposes [90] are helpful contributions in clarifying subject groups and hence the pertinence of results.

The collection and the comparison of baseline data for highly skilled singing voices, using tasks and exercises representative of stage voice use, was thus considered essential in achieving well-grounded normative data of the singing voice. The use of controlled conditions to record professional singers is a requirement for creating databases that can serve to establish increased quality understanding of the singing voice.

The studies included in this thesis were all designed to address some of the problems posed by the lack of adaptation of clinical methods and tools to singing voice demands.

Chapter 2

Important Aspects of Voice

"Every art consists of a technical-mechanical part and an aesthetic part.
A singer who cannot overcome the difficulties of the first part can never attain
perfection in the second, not even a genius"
Mathilde Marchesi

Vocal control, a major difference between the speaking and singing voice, is often implied in the very definition of the act of singing. The degree of this control can help identify the position of both vocal acts along the (voice) continuum, singing being on the higher end. The demands in respect to vocal or phonatory control and acoustic output are great in singing and most especially in the Western opera genre. Singers are particularly dependent on high vocal performance and their criteria for a healthy voice are much more stringent than for non-singers. These demands are considered here, according to two fundamental and crucial mechanisms: pitch regulation and intensity. In what follows, these two aspects of the vocal instrument and the understanding of the Voice Range Profile (VRP) are reviewed, with particular attention to the Western operatic female singing voice.

Western Opera Style

The Western operatic genre, which is deeply rooted in the Italian 17th century tradition of singing, is often equated with classical singing. This type of singing is schooled according to a technique and a style that aim a vocal production based on several different vocal aspects. In a recent doctoral disseration, Daffern[36] gave a useful summary of the Western opera singing voice: a high vocal fold contact, vibrato (generally $\pm$ 6 Hz), a low larynx position (although this has yet to be solidly confirmed for female opera singers), timbre, resonance strategies (the spectral energy peak in the area of 3 kHz or F_0- F1 tuning) and intensity (especially the high intensity which is an important factor to both the perception and the singing of opera).

Brief Orientation

The VRP, or what has been referred to in the past as the phonetogram, is a two-dimensional graph in which phonation is mapped as a function of SPL and frequency $(Ph = f(F_0, SPL) = \{0 \mid 1\})$, Figure 2.1. Pabon *et al.* have recently used a cardinal point reference to VRP regions[114]. For example, in Figure 2.1, high frequency and high SPL in the VRP correspond to the northeastern region. This way of orienting oneself with the VRP is practical and is adopted for the purpose of this thesis. The terms "upper contour" and "lower contour" will also be key identifiers in VRP discussions.
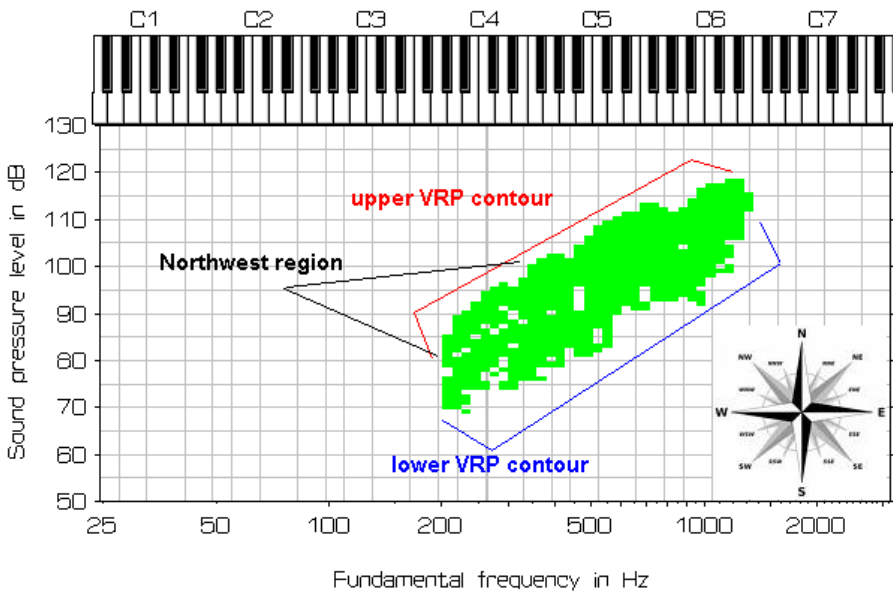
Figure 2.1: The VRP, a performance profile of a healthy lyric soprano, age 27 and active as a regional minor opera role singer. Note that, to accommodate the high SPL levels produced by female singers, the vertical axis was scaled differently than the usual VRP axis (40-120 dB).

2.1 Pitch Regulation

Vibrato aside, the pitch range is most often named as a differentiating marker between speakers and singers. Indeed, in initial vocal training, much attention is devoted to the expansion of the phonation range and often, the promising singer is predisposed with a facility to produce pitches that exceed the speech range. Another great, yet more subtle, control difference between speech and singing is just intonation. After all, singing is not only about words and communication but more

importantly about the proper sequencing of pitch over time (where *proper* refers to the respect of melody and the correct mapping of sound to the notation). In speech, melodic patterns can be varied and manipulated and, in order for communication to be successful, other compensations can be provided when melodic patterns are not respected. In singing, on the other hand, the rigid respect of melody is detrimental.

Although the primary biological function of the larynx concerns an entirely different function than voice production, all healthy voice apparatus are able to produce sound. When we speak of phonation, we address the process in which sound (more correctly labeled as the voice source) is generated by the passage of an airstream through the glottis, which sets the vocal folds into vibration. These vibrations create a harmonic signal that acoustically excites the vocal tract resulting in a radiated vocal output. The perceived pitch of this output is related to the frequency of the voice source (F_0) which in turn corresponds to the vibratory repetition frequency of the vocal folds. F_0 is largely controlled by the laryngeal musculature and by subglottal pressure.

Musculature

As mentioned above, F_0 denotes the vibrating frequency of the vocal folds. The musculature of the larynx (intrinsic directly, extrinsic indirectly) plays a three-fold role in F_0 control; regulating vocal fold tension, mass and elongation. All of the intrinsic muscles of the larynx partake in the adduction/abduction and lengthening/shortening actions that impact the determination of F_0.[1] An adaptation of a Table 2.1, initially created by Hirano and Kakita (1985), and published in the MIT encyclopedia of communication disorders, summarises efficiently the contributions of different intrinsic muscles in the act of phonation.

The *thyroarytenoid* (TA) muscle represents the main portion of the vocal folds and is often referred to as the *vocalis*. Yet, the *vocalis* is in fact only one of two muscle bundles; the other, the *muscularis*, is more laterally located and plays an important role in arytenoid movement [71]. The *muscularis* ensures quick shortening of the vocal folds and the vocalis is used to regulate tension medially. Together, their contractions result in the shortening and the thickening of the vocal folds. Moreover, the *TAs* shortening of the vocal folds increases stiffness.

The *cricothyroid* (CT) muscle divides into a vertical part and an oblique part; attaching at different places. The CT's contractions bring the cricoid arch upwards and thus reduce the space between the larynx's main cartilages and lengthen the vocal folds. This muscle's action is most influential in pitch determination.

The *lateral cricoarytenoid* (LCA) is an adductory muscle, allowing the vocal processes to close by bringing the arytenoids forward and together. The *posterior cricoarytenoid* (PCA), is, in contrast, the chief abductor of the vocal folds and basically reverses the action of the LCA. Finally, the *interarytenoid* muscle (IA) is also subdivided in two parts: a transversal and an oblique part. Together these

---

[1]Extrinsic muscles are also involved in the length adjustments of the vocal folds [150].

Table 2.1: The Phonatory Role of Intrinsic Laryngeal Musculature. Table adapted, with permission, from Kent [76]. The top headers are the abbreviations for cricothyroid, vocalis (or thyroarytenoid), lateral cricoarytenoid, intrarytenoid and the posterior cricoarytenoid, respectively. Italicised text underscores the high degree of the effect listed, parentheses underscore a relatively weak effect.

Vocal fold parameters	CT	VOC	LCA	IA	PCA
Position	Paramedian	*Adduct*	*Anterior adduct*	*Posterior adduct*	*Abduct*
Level	Lower	Lower	*Lower*	-	*Elevate*
Length	*Elongate*	*Shorten*	Elongate	(Shorten)	*Elongate*
Thickness	*Thin*	*Thicken*	Thin	(Thicken)	Thin
Edge	*Sharpen*	*Round*	Sharpen	-	Round
Muscle	*Stiffen*	*Stiffen*	Stiffen	(Slacken)	Stiffen
Mucosa	*Stiffen*	*Slacken*	Stiffen	(Slacken)	Stiffen

two parts see to the proper adduction of the vocal folds, working mainly in the posterior section of the glottis.

According to the theoretical Body-Cover model (a spin-off of the Cover model) [171] both the TA and the CT are importantly involved in regulating the stiffness of the vocal folds, as they are mainly responsible for vocal fold length changes. This model is successful in depicting the differences in muscle recruitment in pitch regulation for both speech and singing. CT and TA both have low activity patterns in speech-like F_0 and they are both involved in raising F_0. This balanced relationship becomes CT dominant in high F_0 phonation and is thus typical in female Western operatic singing.

It is often assumed, as seen above, that the singer's pitch range is greater than the range of the non-singer. However, when measured physiologically, frequency ranges for both groups (notwithstanding voice categorisation) do not vary much. Some authors report no difference at all [35] while others find differences that are statistically significant yet negligible in practice [160]. This is no doubt due to the fact that the anatomical set-up of the larynx is more or less the same for each individual. The important difference in frequency range is uncovered when voice quality and phonatory control are considered: then, a range more comparable to a singer's performable range is obtained. The comparison of the range of the latter kind results in important differences between singers and non-singers. Awan confirmed such differences in his recordings of "musical" VRPs with trained and untrained groups[3]. The results of Paper I also indirectly support this claim, in that the comparison of the singer's physiological and performance VRP indicated negligible frequency-related variations.
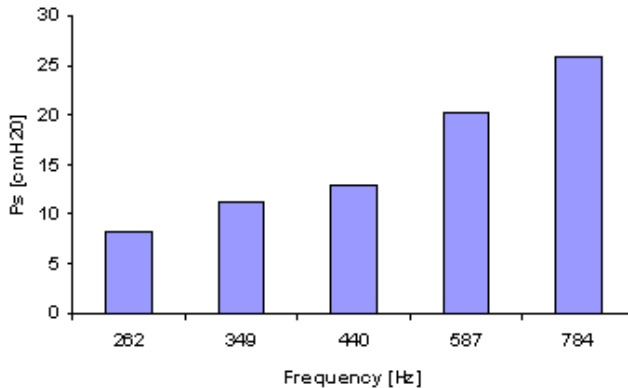
Figure 2.2: P_s as a function of frequency for a constant dynamic (mf). Results are collected for Western opera female singers (N=7), Lamarche 2006 (unpublished)

Subglottal Pressure

In a first approximation, discounting source-filter interactions, the vocal fold vibration is driven by the pressure drop that occurs across the glottis, commonly approximated by P_s. Indeed, in evaluating voice function, it is most interesting to examine the minimum amount of pressure required to initiate and sustain vocal fold vibration. The pressure associated to the threshold between a non-vibratory state and a phonation state is termed the phonation threshold pressure (P_{thresh}). P_{thresh} will vary considerably with the frequency of phonation [149] and also with variables such as the degree of suppleness [147, 177, 176] and of loading [148]. In the same line of thought, P_{thresh} could also vary according to vocal skill. The physiological VRP can be used to monitor P_{thresh} since the lower contour yields phonation threshold levels that can be roughly related to P_{thresh}.

In terms of pitch regulation, P_s generally plays a role in combination with laryngeal muscular recruitment. Indeed, P_s coarsely tunes phonation whereas phonatory motoric actions fine-tune the pitch production. It has been demonstrated that P_s increases with pitch, yet this increase remain fairly low in the low to intermediate range. For example, a P_s increase of 1 cm $H_2 0$ yields roughly 4 Hz in the speech range. When the voice is well trained, as in the case of singers, high P_s is mainly mandated for high pitches when the vocal folds are stretched out and tensed and require higher driving pressures to be set into vibration [162]. Figure 2.2 exemplifies the P_s behaviour that is found for increases of F_0 in the female singing voice.

Nevertheless, the lengthening of the vocal folds is not a purely static event. The vocal fold vibration itself brings some dynamism to the elongation process. The amplitude of vocal fold vibration typically increases with P_s. Titze quantifies this relation's impact on F_0 by measuring the effect of pressure on different vocal fold

lengths [172]. The result is an increase of F_0, more pronounced for shorter folds than for longer folds. Thus, the ratio of vibration amplitude to vocal fold length seems key in F_0 control. At low pitches, where the vocal folds are rather short and slack, only small increments of P_s are needed to increase frequency, whereas at high pitches, much bigger increments will result in relatively little change. This becomes especially interesting in relation to the laryngeal muscle activity. P_s can be used to increase pitch and alleviate muscle action in low to mid ranges, and this economy in turn allows the singer to draw on muscular resources in the higher pitch range.

When vocal problems or disorders occur, a reduction of the pitch range is a typical consequence. For the female Western opera/lyric singer, high pitches are often the first to disappear, as are the fine motor skills in balancing CT, TA and P_s activity. That high pitches are lost is most likely due to a compensatory increase of TA contraction in reaction to a reduced muscle lengthening flexibility. This compensation might be successful in the low and mid-frequency ranges but is practically impossible in the higher range. Moreover, because general vocal fold stiffness is accrued, P_s is also increased, and the increments necessary for vocal vibration in the higher range become quite challenging.

2.2 Vocal Intensity

A certain minimum power is needed for successful speech communication, and this holds also for singing. As with respect to speech, it can be expected that the level in singing voice production will decrease in the event of singing-voice-related problems. Consequently, an attempt to produce adequate sound intensity would then require a heightened level of effort. This vocal dimension is certainly critical for both speech and singing. In relation to the singing voice, Seidner as well as Coleman found that the most distinguishing vocal characteristic between the singing voice of a singer and a non-singer was intense voice [141, 33]. Sundberg [163] also alludes to this difference, specifying that differences are greatest (approximately 20 dB) in the high female voice. In the investigation of the effect of singing voice training, some studies have demonstrated that vocal intensity (either average or minimum SPL) can often mark the difference between the beginner and the advanced singer [106, 94]. Wolf, in motivating the initial VRP concept, also remarked that the ability to voice a high level output on a few pitches did not indicate much about a voice but on the other hand, the ability to sustain a high intensity phonation over a wide frequency range was instrumental to the singer and a manifestation of high efficiency phonation.

In what follows, voice intensity, especially with regard to the singing voice, is briefly reviewed.

Subglottal Pressure

P_s is one of the major determinants of voice intensity. The relationship between P_s and I is rather straightforward when all else is kept equal. One parameter varies
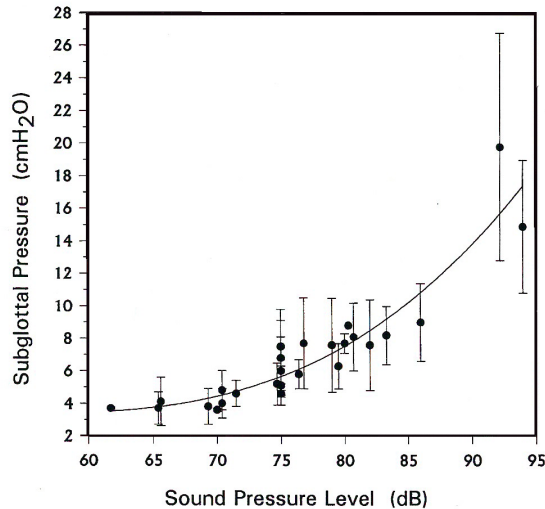
Figure 2.3: Mean P_s as a function of mean SPL summarised from 11 various reports of non-singer investigations (N ranging from 1 to 25). Figure taken from Baken [6]. The curve is a third-order regression fit to the mean P_s values. Error bars indicate the standard deviation for the group averages.

directly with the other. It can be expected that as P_s is increased, the amount of vocal fold adduction must also be increased to sustain vocal fold vibration. The relationship between subglottal pressure and vocal intensity is characteristically non-linear and changes considerably from one individual to another. For normal speakers, Titze and Sundberg [174] established that a doubling of P_s normally yielded a gain of 8 to 9 dB; a result in agreement with Fant's earlier theoretical predictions [45]. In Figure 2.3 mean P_s data is illustrated for different speech voice studies. In singing, Schutte noted a larger gain [137]. In a later study of professional baritones, Sjölander & Sundberg [146] supported Schutte's observation by reporting a gain of approximately 12 dB for a doubling of P_s for the male singing voice. Paper II also confirms a similar gain for the Western operatic soprano and mezzo-soprano singing voice.

Glottal Width

The glottal configuration is also critical in the control of vocal intensity. For the purpose of the next few following lines, we will address the voice source while disregarding the implications of a vocal tract. The level of adduction of the vocal folds will undeniably impact the voice source. More specifically, the level of arytenoid adduction (which is usually described by the open quotient variable) is key in the

production of optimal power. Titze estimated that a power optimum could be obtained when the vocal fold adduction time was equal to the abduction time (an open quotient in the range of 0.5 to 0.6) [171]. In such a state, the vocal fold processes would be closely approximated. From a glottal airflow perspective, Rothenberg stated that a similar open quotient value was ideal for producing strong higher harmonics [130], incidentally, also a determining output power factor. The relation between open quotient and intensity in speech was also investigated by Holmberg *et al.* [68], and similar conclusions were drawn. A decrease of open quotient values could be correlated to an increase in vocal intensity. These observations corroborate with the concept of flow phonation or resonant voice, which is typically defined by minimal vocal fold adduction. Henrich *et al.* [61], further investigated the link between the open quotient and vocal intensity in singing. Similarly to Titze, they found that an increase of 20 dB provoked a decrease of open quotient (from 0.7 to 0.5). However this was not a general observation and rather, was limited to phonations in M1. In M2 (which is mainly employed in female singing), open quotient values were higher ($\pm$ 0.7) and at times even increased with increasing intensity.

The maximum flow declination rate (MFDR), a parameter of the glottal flow, is also known to govern the amplitude of higher harmonics. Then, the MFDR must also be considered as a contributor to vocal power [68] (or at least at speech-like pitches) since loud phonation is best described by higher partial energy, whereas soft phonation is characterised by a strong fundamental [165, 53]. It is useful to note that rising P_s is usually accompanied by a higher MFDR. Other factors, such as the type of phonation (degree of adduction), source-filter interaction and laryngeal mechanisms can also contribute in determining MFDR. That, the airflow of a singer increases in tandem with SPL levels indicates the singer's capacity to quickly adjust the resistance at the level of the vocal folds and maintain good glottal balance.

The VRP becomes a simple yet appealing way to monitor indirectly the glottal efficiency of the singer [163, page 89]. Voice source information can be derived from the systematically recorded minimum and maximum phonation levels for one vowel.

Vocal Tract Transfer Function

The voice source acoustically excites the vocal tract, which means that the total vocal output will be influenced not only by the actual voice source but, also by the process of articulation. Furthermore, this process is non-linear. Events in the vocal tract may also impact the voice source [46, 173]. The transfer aspect of the vocal tract lies in the articulatorily defined vocal tract resonances that selectively amplify the corresponding or nearest voice source harmonics. This phenomenon was theoretically characterised in the source-filter theory [44]. The cross-sectional areas of the vocal tract can be reshaped by articulation to move the vocal tract's resonances. An increased jaw opening will widen the mouth space and reduce the pharyngeal space and thereby increase the frequency of the first resonance. Protrusion of the lips extends the length of the tract and thereby lowers resonance

frequencies. Similar effects are observed for the lowering of the larynx. The shaping of the tongue body mainly influences the second resonance frequency. The tongue tip, combined with lip rounding, impacts the third vocal tract resonance. The first two resonances of the vocal tract are related to vowel definition while the higher resonances (4 and 5) are determinants of voice quality. It goes without saying that this brief revision of the articulatory effects on the resonance of the vocal tract generally presents the amplification function of the vocal tract, and the named effects above are by far independent from each other. For more detailed information, the reader is referred to Ericsdotter's more in depth overview of the articulatory acoustics of the vocal tract[43, pages 135-138].

When vocal tract resonance frequencies appear close together, the sound transfer or voice source amplification is greater. In fact, the singer's resonance cluster (commonly termed as 'singer's formant') is a consequence of such resonance frequency merging. In male Western operatic singing, the vocal tract is shaped to merge the third, fourth and fifth resonance frequencies into a cluster. Due to the transfer function of the vocal tract, the voice-source harmonics in the vicinity of the cluster are amplified. This production of high-frequency energy in the total vocal output happens to coincide with the sensitivity of the human ear to the 2500-3500 Hz frequency region [163, pages 117-124][47, page 315]. From a psycho-acoustic point of view, the voice with energy concentrated in this sensitive auditory region could then be perceived "louder" than a voice deprived of this higher energy, and this despite the quasi-identical overall intensity level of both voices.

The classical two-dimensional VRP is somewhat limited in that it only depicts the complete phonation capabilities of the voice. A third dimension can be encoded with information such as the singer's resonance cluster energy or better, the ratio between overall maximum intensity and the singer's resonance cluster intensity. Such spectral additions to the VRP can facilitate the understanding of the singing voice, and better differentiate the singer and the non-singer's voice [141, 20]. Along the same line, recent work with the perceived VRP (PVRP) elucidated the importance of accounting correctly for singer specific spectral events [70].

In the case of the female singing voice, another acoustic strategy is employed to ensure a vocal output fit for the opera house and relatively "cost efficient". Due to the higher tessitura of the female singer, a smaller number of voice harmonics fall in the region of the singer's resonance cluster and so the vocal tract's sound transfer is not as effective as in the male voice. Conversely, by lowering the jaw, the female singer is able to raise the first vocal tract resonance frequency and "tune" it to F_0. This tuning is not necessarily just and most often has somewhat of an overshoot. According to Joliveau, the vocal tract resonance frequency is typically tuned slightly higher than F_0 and incurs a rise of the second vocal tract resonance as well [75]. Such tuning is efficient, both in terms of vocal fold vibration [172] and in vocal output gain [172, 171]. A typical gain up to 30 dB can be observed in the case of such a strategy. Although this type of strategy is part of a natural phenomenon, female singers train it persistently. Differences in the F_0-F1 tuning between untrained and trained females were illustrated by Sundberg

[163]. Nevertheless, until recently, it remained unclear as to how such tuning was in fact used in singing. Joliveau *et al.*, using a novel vocal tract impedance matching technique, were able to confirm the omnipresence of the discussed tuning in high soprano natural singing [75]. Furthermore, it was proposed by Garnier *et al.* [50] that the F_0-F1 tuning strategy might change somewhat with a change of laryngeal mechanism. The results of a pilot study revealed, below 1047 Hz , a possible merging of the two first vocal tract resonances tuned with the second harmonic of the spectrum (2 x F_0), and above 1475 Hz, a tuning of the second vocal tract resonance to F_0. These resonance strategies would seem to be associated with the transition from M2 to M3 ("whistle register").

The type of acoustic strategy mentioned above is however not so useful in securing high intensity at low frequencies. In singing, however, it was suggested that it might be more successful than increasing P_s. SPL at lower frequencies is determined by the harmonic closest to F1. Titze [172] suggested that when singers learned a vocal gesture tuning the harmonics above F_0 to F1, an average gain of 10-20 dB could be possible. However, Joliveau *et al.*, demonstrated that for female singers, there was no evidence for this kind of tuning or gain at lower frequenices [75]. Female singers often report a certain challenge in producing loud pitches in the bottom of their range. Some explanation is provided in Isshiki's work [73] as well in Coleman's [32]. When laryngeal resistance is low (at low pitches) a significant increase in airflow rate creates an unstable condition that leads to glottal cycle aperiodicity, resulting in an unstable/uncomfortable pitch.

2.3   Increase of Intensity with F_0

When these reviewed vocal aspects are considered together, their interaction involves at least three major vocal actions. First, the tension on the vocal folds and on the muscles of the chest area increases with raising F_0. This increase of tensions builds higher lung/subglottal pressure and consequently higher intensity. Secondly, when the vocal tract resonances are adjusted to match the fundamental of a high tone, a higher-level gain results. Thirdly, the filter function of the vocal tract, transferring glottal power to the radiated power has a bias for high harmonics. All of these factors come into play in the dependency of I on F_0.

The role of the singer is to take advantage of this interactive phenomenon while maintaining control over the separate variables. Voice science has often recruited the singing voice due to its utility for studying vocal parameters in isolation. Accomplished vocal artists usually possess exquisite control, accuracy and reproducibility over various vocal parameters, and for this reason, they are expert subjects in whom the variation of vocal parameters can be clearly studied. A basic example is the study of vocal intensity: singers can easily stabilise frequency while manipulating intensity over a wide range).

Part of the VRP's appeal is the mapping of the interaction between I and F_0. Its capacity to depict the singer's skills in controlling and varying these important

vocal variables is also valuable. In this respect, the space obtained between the two VRP curves (named the *area*) is often interesting and particularly relevant in distinguishing the trained from the untrained voice.

Chapter 3

The Voice Range Profile: 1935 to Today

"The great thing in the world is not so much
where we stand, as in what direction we are moving."
-Oliver Wendell Holmes

This literature review will concentrate on the computerised or the automatic VRP, which is increasingly used in the clinical and research realms. No standardisation, however, has been achieved since the introduction of this phonetograph technology. This means that methodology tends to vary from one study to another, as do software settings and even interfaces. The basic VRP recording procedure continues to rely mainly on the 1983 standardisation of the manual VRP [159, 106, 102, 94, 185, 42, 99, 70, 69, 10]

3.1 Search Strategy

With the automatic phonetogram in mind, a review of the literature was performed in order to better comprehend the status and the role of such equipment in the present research and clinical voice fields. More than 115 studies are found with keywords such as: phonetogram, voice profile, voice range profile or frequency and intensity profiles. To the author's knowledge at least seven Ph.D. dissertations have been dedicated to the study of the VRP: Schutte 1980, Stecher 1983, Gramming 1988, Awan 1989, Sulter 1995, Åkerlund 1996 and Heylen 1997 [137, 155, 51, 2, 159, 79, 62]. Only two of the theses used automatic phonetographs. For the purpose of the present study, the selection was limited to reports including the use of a computer or automatic phonetograph or the specific address of the VRP in relation to the singing voice. The year limit was set to 1980 as computerised phonetographs were first introduced at this time. To ensure that the literature review was as complete as possible, the search for relevant studies took recourse to a variety

of sources. PubMED, MEDLINE, Ingenta, ERIC, CINAHL and SCIRUS were queried via the Internet, as well as Internet search engines such as Google Scholar. Attention was given mainly to English and French and, when possible, German language peer-reviewed journal articles. Some studies were found through informal sources (conference proceedings, and other unpublished work).

Prior to the presentation of the overview results and tables, the history and the standardisation process pertaining to the VRP will be briefly visited.

3.2   History

The concept of a VRP was first introduced by Wolf & Sette and Wolf *et al.* in 1935 [182, 181]. Research was conducted to track the maximum SPL phonation over the frequency range of singers. With 50 singers of various training experience, it was demonstrated that SPL increased with F_0. A threshold was reported at a 2-octave range where SPL saturated or slightly decreased, yet this could have been an effect of fewer collected phonations at those extremes. Subjects were found to have an average level range of 51 dB. In continuation to this work, 5 baritone recordings were performed to study the variation of maximum intensity with vowels[182]. Approximately the same results were obtained; SPL increased smoothly with F_0. Levels for [a] and [e] vowels were found to be higher than for [u] and [i]. Stout pursued this line of thought in the study of sung vowels in relation to pitch and intensity[158]. He explored two conditions; one in which frequency was held constant and intensity was manipulated and the second where intensity was held constant and frequency was manipulated. Three male professional singers participated in his investigation. Similar results to what had been previously reported by Wolf *et al.* were obtained, yet, Stout denoted that the amount of SPL increase with F_0 changed as a function of vowel articulation. Stout requested singers to sing musical tones in both soft and loud levels and introduced the concept of space or area by looking systematically at the intensity extent for each sung frequency. He obtained a group level range of 42 dB. From 1952 on, with the work of Frenchmen, Calvet and Malhiac [23], VRP recording and analysis began to account for minimum intensity phonation (measured in phon). Similar work was pursued by Vogelsanger (who began to register intensity measurements logarithmically in dB)[179].

With Waar and Damste's contributions to the literature, a resurgence of interest for the VRP began in the 1970s. Waar and Damste, who proposed the term "phonetogram", expanded the concept of the acoustic measure to include F_0 and SPL covariation. They moved away from an entirely singing voice study focus to examine the applications of the "phonetogram" in the understanding of voice disorders. Until then, frequency and intensity parameters had been studied only in isolation. Damste was also among the first to tackle the topic of graphical display. In 1977, Coleman [32] reported female and male voice *"profiles"*, grounding his construct of the covariation between F_0 and SPL on Damste's work. Particular interest in the VRP shape, given by the upper contour (maximum intensities through-

out the range) as well as the lower contour (the minimum intensities throughout the range), followed suit in work by Komiyama, Schutte, Coleman, Klingholz & Martin[86, 137, 32, 87, 84].

Looking back on the last fifteen years, one notes that the VRP has been used for many purposes, including theoretical analyses of the voice [170, 31, 160, 144, 109, 143, 131, 70, 91], the study of voice from a clinical perspective [4, 119, 101, 11, 160, 25, 62, 64, 167, 94, 63, 185, 139, 5, 95], the course of therapy [38, 154, 69, 180], and the diagnostic characteristics of specific patient groups [7, 65, 1, 78, 72, 145]. For a detailed review, the reader is referred to Heylen [66].

3.3 Standardisation

While some VRP standardisation issues were addressed at a Japanese meeting in 1982 [67], most current VRP investigations refer to a meeting of the Union of European Phoniatricians (UEP) in 1983 [138]. This meeting of voice professionals resulted in guidelines concerning manual VRP measurements:

- The recommendation of a simple sound level meter set with an A-weighting (dB(A)) (this type of weighting is defined in on page ix),

- A tone generator,

- An omnidirectional microphone unfixed to the measuring equipment

- A 30 centimeter microphone-to-mouth distance

- A graphic display window of 15 mm vertical per 10 dB, and 36 mm horizontal per octave

- The measurement of the amplitude of the singer's formant at maximum intensities with a filter system

The UEP meeting also loosely defined a tasking protocol using three vowels (/a/ / i/ /u/). Interestingly, many reports refer to this vowel recommendation to which phonation time is also added. However, the phonation time which seems to be commonly observed, 2 seconds, is rather traced to Coleman's work in 1977. Instrumental VRP studies followed in the aftermath of the UEP recommendations. Certain weaknesses in the chosen standards were identified. These later studies led also to a deeper understanding of certain VRP characteristics. In a thorough investigation of spectrum factors, Gramming (1988)[52] demonstrated the effects of measurements using A-weighted as compared to linear weighted frequency curves. Figure 3.1 illustrates the outcome of those comparisons between linear and A-weighted voice measurements.

In the case of A-weighting, the bottom curve, representative of minimum phonation levels, was lowered, particularly for low frequencies. The upper curve, representative of maximum phonation levels, was also lowered, if somewhat less. Gramming
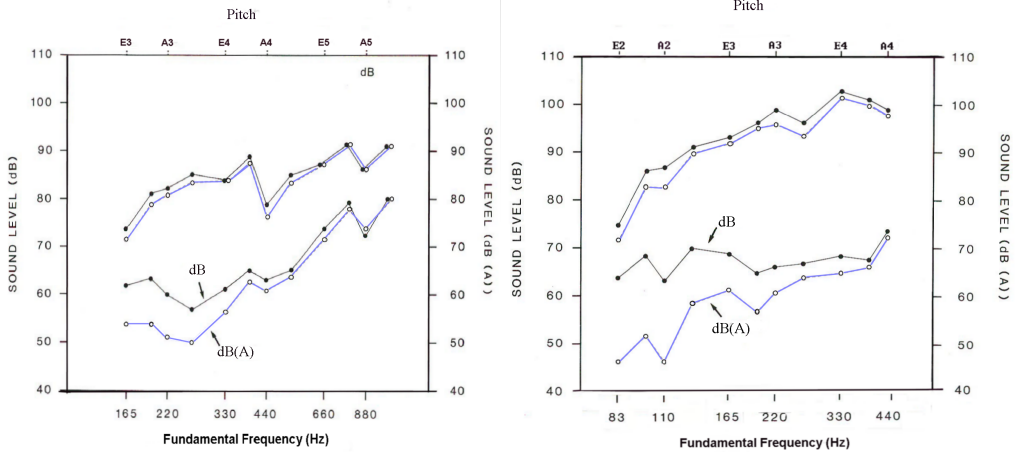
(a) VRP Measurements of a Female Voice   (b) VRP Measurements of a Male Voice

Figure 3.1: The effect of different SPL weighting on the VRP. Both graphs are taken from Gramming. Reprinted with permission from [53]. Copyright 1988, Acoustical Society of America.

studied the SPL variation of several vowels, including those recommended by the UEP, and elucidated reasons for selecting /a/ as a vowel for VRP tasks. Since /a/ has a high first formant, the use of this vowel minimises the chances that the fundamental frequency will conflict with the first formant frequency[52]. Gramming [51] also confirmed an earlier observation noted by Sonninen [152] concerning the cause of VRP contour knees. Sonninen had proposed that such knees or abrupt changes in amplitude, could be attributed mainly to vocal mechanism transitions, agreeing with Klingholz [82], but could also be an acoustic artifact of the crossing of partials with formants.

Titze, in 1992, presented an acoustic interpretation of the VRP shape [170]. His study focused mainly on the co-variation of I and F_0. He explained the difference in the slope of the upper and lower curve by the nature of the spectral distribution in relation to the fundamental frequency. He also dwelt on the strategies for achieving and maintaining a pressure above phonation threshold pressure. A large part of the work differentiated the gain obtained by glottal source manipulations (more efficient in speech and low pitch singing) from the gain related to the interaction between subglottal pressure and F_0.

In the same year, 1992, the International Association of Logopedics and Phoniatrics Voice Committee (IALP) met to discuss assessment topics, one of which focused on the phonetogram[13]. The discussion recorded at this meeting suggested the support of the above-mentioned recommendations by the UEP and motivated

the term "Voice Range Profile" in replacement of phonetogram. At that time, VRP procedures were increasing in popularity in North America and this instance was seen as an opportunity well suited for the implementation of terminological modifications.

Following this discussion, Coleman issued a key paper in 1993 [31], taking to issue meticulous details of methodology and VRP recording set-ups. This paper touched on several aspects of the variability found in VRPs. Gramming had previously examined the long-term and short-term variability involved in recording a given subject. Coleman followed up on this theme, including many further aspects of variability related to set-up, methodology, instructions and physiological and what he called the "musical range of phonation" VRP.

In 1994, Titze produced a standardisation paper addressing utterances used in research and clinical investigations [172]. A portion of this paper focused specifically on VRP utterances. In that paper, the VRP is considered as a basic reference in defining test utterances. In other words, the VRP maps out the boundaries for the testing. Interestingly, glides, or what Titze names dynamic tasks, on sustained vowels, are suggested for the VRP recording. The VRP procedures mentioned in this paper include both speed and accuracy. Finally, a normalised low-medium-high range sampling is also mentioned. That paper further addresses other issues that can be generalised to overall voice testing but that relate well to VRP recording. F_0 extraction aspects that would require consensus in the voice community are brought forth: the meaning of F_0 in chaotic, highly random signals, the definition of perturbation upper limits, the selection of appropriate microphones and the external effects of recording (noise, room acoustics or source-receiver stationarity) on measures like jitter and shimmer.

In 1995, the topic of VRP standardisation was revisited by the IALP during the XXIII Congress. This discussion was intended to revise and update the 1983 standards in view of the then-current technical progress [120]. For a decade, new automatic phonetographs had been used and so this discussion focused mainly on certain phonetograph technological details. The VRP display was established: a 40-120 dB vertical axis versus a 50-2000 Hz horizontal axis (to this day, this display frame seems to be the standard). A and C-weightings were once again compared and tested. A-weighting was found acceptable and rather advantageous in the event that the influence of background noise needed to be minimised . A-weighting was recommended even though the signals obtained with both types of weightings were negligibly different ($\pm$ 3 dB) only from approximately 500 Hz on. At lower frequencies a maximum difference of 10 dB was found. It was claimed that, since the strongest energy of a voice signal at low frequencies lies mostly in the first formant (given [a:] is used for tasking), the noted 10 dB discrepancy between the two weighthing was practically negligible. The IALP paper also suggested a flexible frequency window to fit different matching phonation capabilities of subjects (and seemed to take into account differences between singers and what they referred to as "tone-deaf" individuals). A semitone resolution of $\approx$ 6 % ($\pm$ 3 % maximum quantisation error) in phonetograph measurements was determined as a

necessary standard. This semitone resolution for bin definition is indeed practical and is used in most current phonetographs. Here it is understood that the VRP is a gross total vocal output measurement which cannot provide the fine frequency detail that might be sought in specific investigations of the disordered voice. In this case, other measurements might be better suited for the frequency analysis of certain behaviours. A minimum of 0.5 seconds up to a maximum of 60 seconds were suggested recording times; 1 minute was deemed the most suitable for running speech tasks. This recommendation did not necessarily account for measurement replicability and is rather loosely defined in that the recording time for a VRP recording is highly dependent on the nature of the task and the investigation question. An audio frequency range of 40-15 000 Hz was suggested as a standard to properly record the energy of the voice signal. It may be noted that, in the vast majority of cases, the SPL would be adequately represented by the energy below 2000 Hz. These recommendations were put forth in 1995, when computers did not have the capacities and speed that have become a given in present day computers. The paper's contribution, made on the basis of the equipment capabilities at the time is no longer entirely relevant to the modern automatic phonetograph (for example, a response time based on a threshold of accumulated occurrence per cell is at present a common feature of phonetographs). The paper, not only looked at the important considerations of the automatic phoentograph but also suggested VRP default metrics: area calculation (dB * ST), lowest/highest frequencies and SPL as well as respective ranges.

Heylen, in 1996, commented on the need for standardisation in his review of Coleman's work regarding VRP sources of variability [65]. Still, in 2000, in spite of many discussions and papers, Baken reported an absence of standardisation for the VRP in a second edition of his manual on Clinical Measurement of Speech and Voice [6]. At present, numerous hospital clinics in Europe (or at least in northern Europe and Scandinavia, the Netherlands and Belgium) are equipped with computerised phonetographs, with a choice of different commercially available systems. The current VRP reality seems to be disjunct from its manual past and some of the above-mentioned recommendations. Criteria for the set-up of automatic phonetography have yet to be formally established.

3.4   Computerised Voice Range Profiles

The automatic VRP was first mentioned in the early 1980s. Inspired by the work of Rauhut *et al.* that made use of an automatic X-Y plotter [126], Gross [55], developed a "half automatic VRP". However, this often-cited article does not present any data nor technical specifications of the equipment. The focus is given to the improved objectivity obtained with such a phonetograph as well as the freedom gained by the clinician, namely frequency extraction/judgment. The paper is a landmark in that it created a new avenue for voice phonetography. Work from Bloothooft further elaborated the concept of an automatic phonetograph and the

mapping of contours on a X (logarithmic scale-Hz), Y (linear scale-dB) coordinate graphical system. With these new recording possibilities, the concept of the VRP also evolved from a set of curves to an area or space.

In 1981, the first fully-computerised VRPs were recorded with 14 singers [14]. Bloothooft particularly demonstrated the interest that lay in the VRP area by mapping the singer's voice mechanisms ("register"). Moreover, the possibilities of integrating a third, spectral dimension to the VRP were elaborated. Even then, a tentative measure of the relative strength of the harmonics in the 2 to 3 kHz region (the difference between total intensity and the singing resonance cluster) could illustrate the spectral behaviours pertinent to loud and soft phonation. This new proposed recording system had the capacity to measure voice in a 70-1300 Hz and 20-120 dB range and extract frequency and intensity 20 times per second.

A fully automatic phonetograph was also developed in Finland [152]. The F_0 / I Analysis, Phoniatric Application I, Version 1.0 by Raimo Toivonen worked in connection to a Speech Processing System. This system was able to sample speech at 10 kHz. Individual pitch periods were identified (based on a time domain analysis) and the amplitude of each pitch period was computed. This information was then plotted on a X,Y graph as mentioned above. In 1988 [151] the micro-computer application could run in real-time, and had a commercialised user interface. The "voice field" or the equivalent to the VRP, worked on the principle of a two-dimensional histogram. Each VRP cell had a two-dimensional bin and the number of phonation occurrences per cell yielded a third dimension. Cells were able to store up to 65 635 occurrences. A cell contour threshold could be manipulated to eliminate artifactual variation. However, the system was limited in its pitch detection of chaotic or irregular phonation and could only measure voice in a 40-500 Hz range.

Table 3.1 presents a summarising comparison of the general differences found between the manual or "classic" VRP method and the computerised VRP. In Table 3.2, investigations that have given particular attention to the automatic phonetograph and its development are listed, while Table 3.3 lists the main phonetographs that are currently available on the market.

Table 3.1: Features of Manual and Automatic VRPs Compared

Features Compared	Manual VRP	Computerised VRP
Voice Range	Maximum voice range (F_0 and I) capacity	More comprehensive/realistic voice range (due to faster sampling)
Phonation Time	Phonation is required to be 2-3 seconds	Regulation of phonation time threshold according to needs
F_0- SPL match	F_0 and I are matched manually in post-recording processing	F_0 and I are measured synchronously
Processing	Static	Dynamic and real-time
Feedback	Investigator feedback	Visual concurrent biofeedback
Phonation Type	Limited to vowels and sustained sounds	Records connected speech, reading, vowels, consonants and singing (up to 4000 Hz)
Support Ranges	Range $\leqq$ 3 octaves, $\leqq$ 50 dB	Range of 16-4000 Hz, 40-125 dB [a]
Requirements	Musicality is a prerequisite (both for the subject and the investigator)	Musicality is not required (neither for the subject nor the investigator)
Display	2-dimensional display (F_0 and I)	3-dimensional display (F_0 and I and number of occurrences per cell or other parameters such as the crest factor, jitter, singing resonance peak energy)

[a]These ranges are however not standardised and screen displays vary from one phonetograph to the other (e.g., Phog, VoiceProfiler, lingWAVES and KayElementrics have all different displays).

Table 3.1: (continued)

Features Compared	Manual VRP	Computerised VRP
Boundaries	The boundaries of the VRP are clear	VRP boundaries are not always clear (due to cycle-cycle analysis)
Recording Time	Lengthy acquisition time	Time efficient for contour recordings (somewhat lengthier for full area sampling)
Pitch Extraction	Pitch extraction is not possible with chaotic voices	Pitch extraction algorithms have difficulties detecting irregular phonation
Data Comparison	Cross study comparisons have to be done manually	Data collection and comparison are facilitated due to storage and norm building options

Table 3.3: An Overview List of Current Phonetographs

Name	Company	Country
Phog 2.50	Saven Hitech	Sweden
LingWAVES	Wevosys	Germany
Voice Profiler 4.0	Alphatron	Netherlands
Dr. SPEECH, Phonetogram v.4	Tiger DRS Inc.	USA
Sesane v.3.2	S.Q. Lab	France
Voice Range Profile, Model 4325	KayPENTAX	USA
Protrain	Avaaz Innovations	Canada
Phonomat 84	Homoth	Germany

As previously stated, the standardisation of the VRP is still pending. A small step towards this goal would be a standardisation at the level of the measurement equipment. In this light, a comparative, simultaneous parallel recording with as many as the above listed phonetographs could be exciting and quite informative. This kind of analysis was however out of the scope of this thesis. The suggestion

Table 3.2: A Summary of Studies Focusing on the Use of an Automatic Phonetograph. Title square brackets indicate papers originally written in another language than English.

Year	Subject/Theme	Author(s)	Purpose
1980	[Voice Field, a more objective voice diagnostic method]	Gross, [55]	Introduction of a "half automatic" phonetograph
1981	A computer-controlled device for voice-profile registration	Bloothoof, [14]	14 singer recordings
1983	Quantitative evaluation of the voice field	Klingholz & Martin, [84]	A computer program that evaluates the VRP with 2nd-order curves
1984	A phonetograph for use in Clinical Praxis	Pedersen *et al.,* [122]	Manual VRP compared to automatic VRP
1988	[Evaluation of the quantitative speaking voice production: the phonetogram of the speaking voice in relation to that of the singing voice]	Hacki, [56]	First mention of "*Phonomat*" Homoth, a fully automised phonetograph developed by Hacki and colleagues
1985	Computer voice fields of connected speech	Sonninen *et al.,* [152]	The development of a automatic voice field for clinical purposes
1986	Observation on voice production by means of computer fields	Vilkman *et al.,* [178]	Assess the capacity of the computer voice field to display basic voice production features.
1986	Computerized Phonetograms for Clinical Use	Pedersen *et al.,* [121]	Unseen
1987	Computer voice fields in basic phonation research: rotation vs. gliding in cricothyroid articulation	Sonninen *et al*,[153]	The description of the computer voice field and the method concerning it
1988	Automatic phonetogram recording supplemented with acoustical voice-quality parameters.	Pabon & Plomp, [117]	Jitter, sharpness and breathiness

is put forth in the hope that such comparative assessments of equipment might be made in the future.

A recent investigation [114] demonstrated a strong partial convergence across multi-source computerised and manual data. As mentioned in the paper, different data sets of similar subject populations seemed to align consistently with the ascending lower part of the upper VRP curve. Hence in the middle-range and

Table 3.2: (Continued)

Year	Subject/Theme	Author(s)	Purpose
1991	Objective acoustic voice-quality parameters in the computer phonetogram	Pabon, [111]	Comparison of manual VRP results to automatic VRP equivalents
1991	Computed phonetograms in adult patients with benign voice disorders before and after treatment	Pedersen, [118]	Unseen
1992	Computer aided evaluation of phonetograms	Klingholz, [83]	Elliptical analysis of VRP shape
1993	The phonetogram in singing voice analysis and synthesis	Pabon, [112]	Presents possibilities for synthesising voice by aid of the VRP
1994	Automatic phonetographic recording of normal voice	Kotby &Orabi, [89]	A collection of normative VRP data of non-singers
1994	A structured approach to voice range profile (phonetogram) analysis.	Sulter *et al.,* [161]	Fourier Descriptors and automatic analysis
1998	Dynamics and voice quality information in the computer phonetograms of children'svoices	Pabon *et al.,* [116]	A comparison between manual and automatic VRP results, an averaging technique
1999	On the comparison of computerised phonetogram	Bloothooft *et al.,* [16]	Hidden Markov Model suggestion for VRP categorisation
2006	[Comparison of the results obtained through manual and automatic phonetogram]	Montojo et al., [103]	Manual VRP compared to Dr. Speech

the speech-range conditions, variability was negligible. On the other hand, when extreme conditions were considered, phonetograph differences emerged. Such variation might not be of concern for speech and the typical speaker's range, which does not often visit extreme vocal conditions, as far as the phonetograph's recording capability is concerned. In singing, however, these differences could matter.

Most of the overall divergence between studies can be expected to originate from the protocol and other non-coding VRP variables, such as the calibration procedures and peripheral equipment. With the automisation of the VRP, a step was taken towards obtaining improved objectivity in regards to VRP voice evaluations. However, many more aspects involved in recording a VRP remain as possible sources of bias. Studies which have examined the effects of time of day [102], the potential role of the investigator [175], warm-up [31], instruction formulation [151],

cognitive aptitude and motivation and emotions of the subject [151, 117, 31, 19, 172] all point to variables that are difficult to control and which thus might jeopardise VRP reliability and validity.

Phonetography is often compared to audiometry, yet VRP acquisition is much more than the testing of a relatively passive phenomenon such as the human hearing. The voice is complex and this makes any attempt to measure it quite challenging and fascinating. Internal factors such as personality traits, emotional and pathological states as well as different habits of voice production and external factors such as the environment, all lead to very individual measures that are far from fixed. Indeed, the variation that accompanies voicing is often unconscious, especially in speech where phonation is mainly automatically managed. The VRP, when compared to the audiometer, has to contend with more sources of variation.

This motivates further work in attempting to objectify the VRP process as much as possible, in order to gain control over the variables that can be regulated.

3.5   Metrics of Importance

Although numerous studies include the VRP as a measure, very little has been written on VRP interpretation. Moreover, the lack of established VRP metrics hinders VRP comparison, as well as the establishment of standard reference values for the VRP[160, 65, 154, 10]. No formal recommendations are in place with concern to the use of VRP metrics. In reaction to this, an inventory of VRP investigations was taken , in order to assess recurring metrics and observe the tendencies and the interest in VRP measurements. Table 3.4 tabulates the information obtained from the current VRP literature.

As exemplified in Table 3.4, the traditional and most basic metrics, those which could be assessed by the manual VRP method, remain those most often reported. Heylen had earlier demonstrated that the general interest in VRP analysis was directed to frequency aspects of voice, and very little attention was given to the power aspect [62]. On the contrary, here it is found that there seems to be an overall interest in frequency and intensity-related metrics, including their interaction (unlike Heylen's, only studies of the VRP are accounted for in this review). Some authors have introduced new metrics, but all too often, those metrics are investigated by the initiators alone and their application does not seem to generalise to the voice community. All in all, an implicit consensus seems to exist concerning metrics specifically related to frequency and SPL ranges and maximum values as well as area. Still, the methodological aspects of measurement and report vary extensively from one study to another. Some might measure the VRP area in cm^2 while others will measure in dB x semitone. Furthermore, in some studies, the VRP area is divided into subareas that refer either to speech, voice mechanisms or singing power. Some investigate the overall VRP slope while others choose to look separately at the slopes of each contour curve. Reports of maximum and minimum frequency and SPL for the overall VRP are common, yet sometimes it is unclear

if the metrics refer to absolute extremes or rather designate F_0 related maxima and minima. The metric of "comfortable" intensity is also occasionally reported. Additionally, there is a dichotomy in reports of mean SPL: some studies report the SPL average while others report the Leq.

3.6 VRP Analysis

The extraction of metrics from VRPs is a common practice in the analysis of voice function. However, most of the information obtained with these metrics can indeed be extracted with other types of voice measures and do not take advantage of the two-dimensionality of the VRP representation. Qualitative judgments of VRPs are certainly desirable and can be of great clinical assistance. However, this type of analysis is difficult to extend to data comparisons, and relies heavily on the clinician or the researcher's subjective experience. Some attempts have been made to objectify and code VRP evaluation. Those efforts can be summarised in seven types of approaches.

- Rescaling/Normalisation

- Ellipses

- Indexing

- Shape Descriptors

- 50 % Overlap Method

- 95 % Prediction Intervals and Mathematical Transformations

- Contour Averaging by DFT

Rescaling/Normalisation

This VRP approach was borne out of the need to quantitatively compare data across subjects. Since intensity is dependent on frequency, a form of normalisation of the VRP frequency axis was suggested. The normalisation consisted of dividing a subject's full frequency range by 10 % increments for a total of 11 data points. The division of the range thus defines the tones that are prompted to the subject. This approach was first defined by Coleman *et al.* [32] and reappears slightly modified in Gramming's dissertation work [51]. Instead of computing a normalisation to define the VRP recording exercise, the full frequency range was explored with the subject and only in a post-processing stage were frequencies converted to semitones in relation to the lowest phonation of the individual tested. The F_0 for each of the vocalisations was expressed as the percentage of the overall range obtained. Once a number of VRPs had been rescaled in this way, group data could be handled
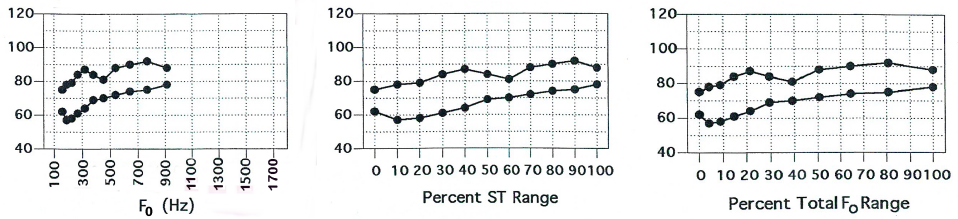
Figure 3.2: Taken from Baken, these figures depict well the visual transformation that occurs due to the normalisation of the VRP frequency scale. The data illustrated here is average data for female non-singers.

more agreeably. Many studies have used this technique in one way or another [136, 2, 161, 79, 94].

There are however two great disadvantages to this approach. 1) All frequency-dependent intensity information is lost. Orginal data interpretation is consequently impossible. 2) The graphical changes entailed by this rescaling method have important consequences. The original shape of the data is distorted considerably and the VRP of a group becomes uninteresting from a morphological standpoint. The original VRP becomes so expanded that certain characteristics, which might otherwise be quickly identified, lose their interesting singularity. Sulter as well as Coleman effectively illustrated the graphical effects of this type of normalisation. For the lack of a better solution and to enable data comparisons, both used the technique (Sulter proposed an alternative but needed to revert to a form of rescaling to compare his data with others). Figure 3.2, taken from Sulter *et al.* [160] depicts the deformations incurred by VRP rescaling.

Ellipses

A proposal to quantitatively assess VRP information, taking into account the two dimensions of the VRP, was initially introduced in 1983 and revisited in the 1990s [84, 83, 1]. The method mathematically prescribed ellipses to different sections of a VRP. Ellipses were based on five parameters: main and secondary axes, rotation and X,Y coordinates of a central point. The slope intersections of the ellipses were markers for register transitions. The authors departed from their observations of register manifestations in the VRP. They claimed that two vocal mechanisms were present: a phonation with high adductory activity and a phonation with high tensor activity. They also determined a mixed region they referred to as a transitory area.

This was a complicated method, given that the intention was to introduce a practical VRP evaluation for the clinic. A particular weakness of the ellipse analysis is found in the degree of arbitrariness in ellipse allocation to the VRP. Klingholz himself reports a lack of reliability in the computerised method's detections of
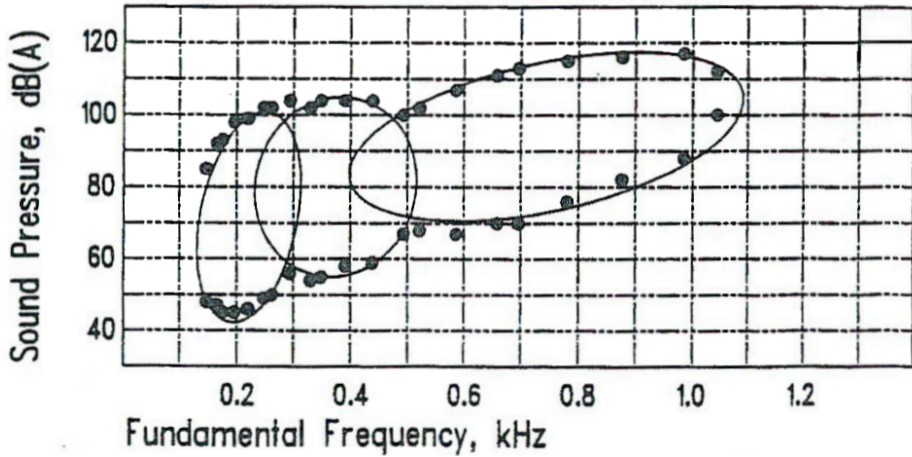
Figure 3.3: The fitting of ellipses in an attempt to approximate VRP contour points. In this example, three subareas are defined: the lower ellipse designates chest voice, the middle ellipse, a transitory space or voix mixte and the upper ellipse, head voice. Reprinted with permission [1].

register transitions, and stated that the investigator would need to manually modify these according to subjective judgment. The number of ellipses in such analysis is often limited to three, yet it remains unclear if this choice is automatic regardless of the voice recorded. Furthermore, ellipses can be made to fit the VRP data in numerous ways; and this approach would have difficulty in dealing with deviant data acquisition points. Figure 3.3 gives a general illustration of the method as first presented in the Klingholz & Martin article [84].

Indexing

Some attempts have been made to derive clinically relevant VRP indices.

Eichel [40], was perhaps one of the first to attempt to quantify an index that could be used for the evaluation of the VRP's graphical display. He introduced the "indifferent point", which he defined as the combination of the SFF and the SPL mean obtained for the four time repetition of relaxed counting from 5 to 10. From this indifferent point, a horizontal and a vertical line could be traced to the boundaries of the VRP contours. A third line was also drawn at a 36°angle as to trace a 1.33 dB increase per semitone. The extreme points of each of those lines were labelled by coordinates. The summation of the line segments corresponding to the lower coordinates defined the *Indifferent index*. This index hence related the indifferent point's position to the VRP's lower contour. The summation of the lines of upper coordinates defined the *Increase index*. This index served to characterise

the increase potential of F_0 /SPL in relation to the indifferent position. Finally the sum of both indices was named the *VRP index*.

It was found that these indices, together with maximum phonation time, would be practical in characterising the output power resources of a voice. However, this kind of approach was designed specifically for speech and was incapable of seizing subtle voice status changes. To the author's knowledge, no other VRP work has been based on this last evaluation scheme.

Heylen constructed the VRPi, a diagnostic index which he developed in his study of pediatric voice pathologies. This index was a combination of several VRP metrics with the subject's age. The aim for such an index was to quantify the functional vocal performance of children and thus facilitate diagnostic screening. This index relied on principles of discriminatory statistics (the Fischer discriminant analysis) and was shown to be efficient in distinguishing between healthy children and children with voice disorders. It was also shown that the index could help quantify the tracking of voice therapy progress. With this approach, a single value was attributed to the VRP, accounting for both metrics as well as the VRP's 2-dimensional representation. A rescaled version of the index resulted in referential values of -10 (a cut-off value indicative of pathology) and +10 (a cut-off value indicative of vocal health). The discriminative abilities of the index were reportedly weaker when the upper curve VRP slope metric was excluded.

This method seems indeed promising. A downside of the indexing is that it relies heavily on the subjective localisation of laryngeal mechanisms and consequently, slope assignment. Laryngeal mechanisms, especially in children, might be difficult to assess perceptually. The detection of "breaks" can be related to resonance issues and/or necessary changes of the voice source, and this differentiation is not always perceptually evident. Furthermore, VRP contours might not carry sufficient information to correctly identify transitory instances. For example, the crossing of formants and partials can also create abrupt contour changes.

Shape Descriptors

In his doctoral dissertation, Sulter introduced a structured approach to VRP evaluation. This approach centered on VRP shape description and on the importance of speech voice dynamics. The approach was also directed to individual VRP analysis. According to a shape quantification method elaborated by Zahn and Roskies (1972), the contours of the VRP were converted into a set of slope values as a function of length. From the line length and angle information provided by the set of slopes, Fourier Descriptors (FDs) could be derived. FDs are often used to measure, recognize and classify shapes. For example, FDs are classically employed in the analysis of handwriting. Discrete VRP shape changes were thus tracked, without distorting the original overall shape of the VRP. Figure 3.4 illustrates the process in three steps.

A weakness of the FD approach closely resembles one found in the former elliptical method: there is a dependence of the method on data acquisition. In this
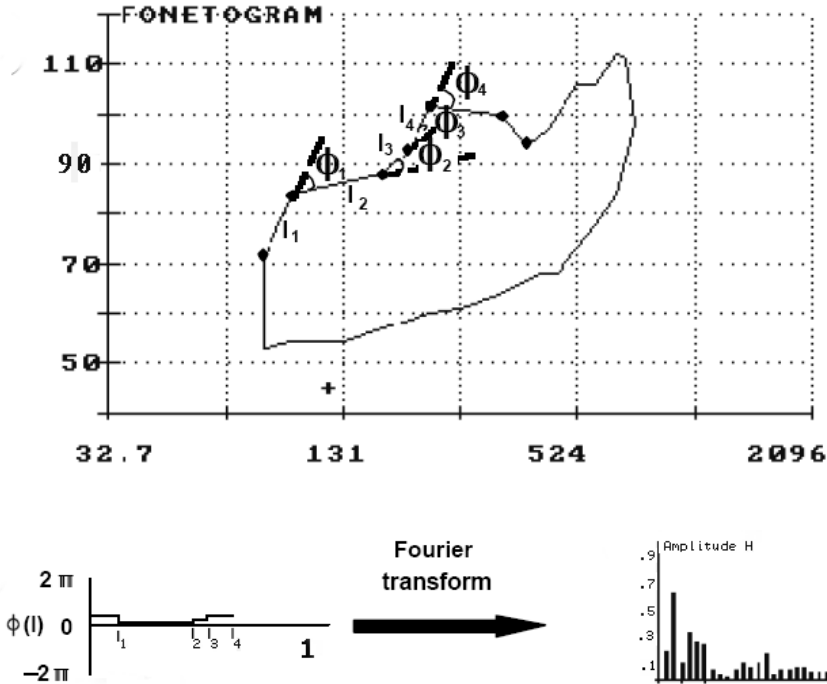
Figure 3.4: The Fourier descriptor approach in three steps, a) line lengths are calculated from lowest loud phonation to highest and angles between these line segments are calculated; b) line lengths and angles are plotted according to new axes; and c) The Fourier descriptors (general shape contributors are closest to the point of origin and higher descriptors are related to specific and detailed shape changes. The amplitude of the descriptors illustrate their particular contribution). The figure is reprinted, with permission, from the dissertation work of Arend Sulter [159].

instance, shape parameters are dependent on the total number of points in the VRP contour. This means that such an approach is restricted to a consistent selection of contour points.

50 % (median) overlap method

In 1990, Hacki presented a VRP averaging method [59]. The goal was to be able to create an average VRP without losing the intensity dependence on frequency. In this process of averaging, since there was no rescaling involved, subjects had to

be grouped meticulously. Hacki grouped his subjects according to voice category. Furthermore, this kind of necessary grouping could lead to a better characterisation of voice category differences. Overlapping VRP cells are accumulated and three curves are created to illustrate the 10, 50 and 90 % amount of overlay of VRP data.

The idea was further developed by Pabon *et al.*[115]. In an automated setting, the number of times that a cell (1 ST by 1 dB wide) falls inside a phonation contour is tracked and this count is visualised by a colour map (the darkest showing the highest accumulation and the palest, the lowest). At the same time as the programme scans the counts of the frequency scale, a vertical scan detects the changes at the extreme of the intensity range for each semitone that corresponds to the 50 % occurrences. In this way, a 50 % occurrence VRP contour can be traced, showing the average upper and lower intensity contours. In a later study[114], this technique is compared to another averaging method, and it emerges that the 50 % contour trace acts more like a median than a mean. The advantage of such a method is that it allows to track not only the contours of a group but also the inner VRP areas that can be used to reflect voice quality characteristics.

95 % Prediction intervals and mathematical transformations

This VRP norm building procedure can be summarised in four steps. This new approach to normative VRP data is borne out of the need to include some measure of variability to facilitate individual-to-group comparisons [63]. The method consists of first converting frequencies to semitones, and secondly placing the VRP's starting point at a same semitone (the semitone below the lowest phonation for the group). Unifying individual VRPs to one common start involves a translation process. Thirdly, a compression process ensures that all VRP-final semitones also coincide. Finally, the intensity points are interpolated over a detailed semitone scale (0.05 ST) and mean upper and lower intensities along with confidence intervals (95 %) are calculated per semitone value. The total VRP can then be rescaled by the same factor previously used in the compression phase and shifted in the opposite direction taken in the translation phase. Semitones are reconverted to semitones to yield the normative VRP. Figure, 3.5a 3.5b taken from the Heylen's methodological article, depicts some of the different phases involved in norm VRP building.

Contour Averaging by DFT

Recently Pabon *et al.* proposed a novel approach to VRP contour averaging, based on morphological modelling [114]. The underlying two-dimensional circularity of the Fourier Transform is exploited to characterise the shape of the VRP. Inspired by Sulter's earlier use of FDs, this technique also respects the individual shape characteristics of the VRPs. The method is distinct, however, in that it considers all absolute contour point positions of the VRP, and so the information of the contour is completely accounted for in the averaging process. With this method any VRP shape can be brought to a common uniform base. This base can then be

(a) The Translation\Compression Phase   (b) The Norm VRP
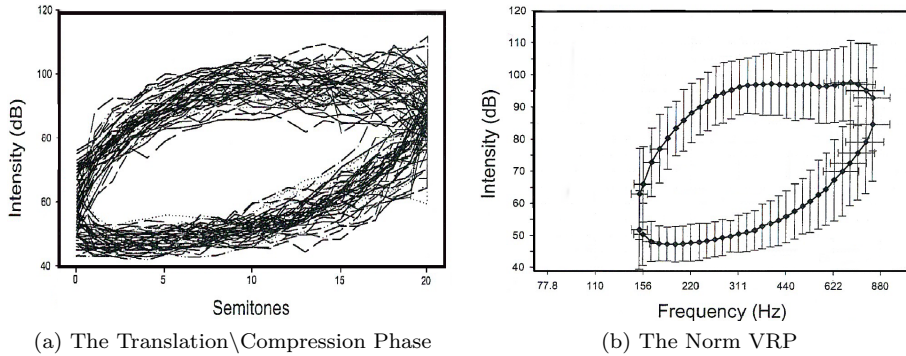
Figure 3.5: The 95% prediction intervals and mathematical transformations involved in building normative VRPs. Permission to reproduce these figures found in Heylen *et al.* was granted.

used as a platform for the comparisons of contours from a variety of sources. Local co-variation along the contour average can also be depicted. The method is however limited to VRP contours and does not offer greater improvements from the 50 % coverage averaging method which is able to account for interior VRP information as well. Figure 3.6 illustrates a central motivation for the DFT contour averaging approach.

3.7 VRP of the Singing Voice

As discussed in Chapter 1, diagnostic and evaluative methods used in voice care are mostly designed for the speaking voice, and are not necessarily directly applicable to the singing voice. Indeed, the performing singing voice requires specific attention in that it uses a range different from the speaking voice and possesses several other features not present in speech. The VRP is a useful resource that can assist in the improvement of the documentation and the understanding of the singing voice. After all, the study of the singing voice was the initial source of motivation for the elaboration of the VRP. Only later did the VRP serve to analyse non-singer's voices and disordered voice function. Coleman effectively comments on the asset of the VRP in the evaluation of voice: "the phonetogram allows to draft a balance of the vocal capacities in relation with the demands" [31]. Naturally this last statement applies to the general applicability of VRP recordings in the clinic, yet, it seems particularly suitable in the case of the singer patient.

The present literature review reveals that there is an extensive body of normative and reference phonetograms. In 1996, Heylen completed a helpful summary in his survey of the literature [66]. These resources however, seem limited mainly to the study of non-singers. Data published on the VRP of singers is relatively scarce
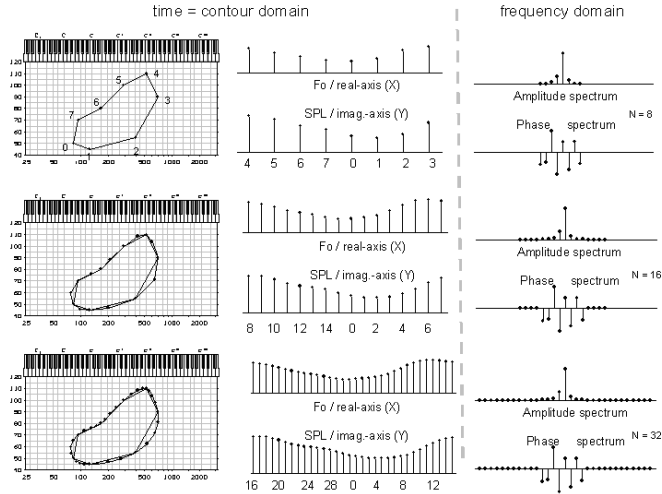
Figure 3.6: Interpolation by zero-padding. The figure displays an important part of the contour averaging by DFT. When contour points in the contour domain are numerous, the contour spectrum (mirroring the same number of points) will contain more high frequency components (+/-). Since these high-frequency components typically have very low amplitudes, they become instrumental for the zero-padding involved in the interpolation that is performed in the contour domain. This way, any VRP contour can be uniformly resampled to a predefined number of points. The figure is reproduced from Pabon *et al.* [114].

and when it exists, it presents great procedural and methodological incoherences. Furthermore, the definition of singing subject groups seems quite broad; sometimes including both sexes, diverse training experiences and genres of singing. Table 3.5 summarises the studies that were found concerning the singing voice and the VRP.

Some of the data collected in the studies included in Table 3.5 were digitised in order to enable a comparison of different singer VRPs. A focus on the female singing voice was given in this comparison since it was most pertinent to the work presented in this dissertation. Naturally, it would be inappropriate to digitise rescaled or normalised data, and therefore such results could be included only when the original frequency and intensity data points were recovered. This VRP comparison was performed regardless of the type of VRP recording and regardless of the approach (manual or computerised). A stricter comparison would benefit from more convergent data. Figure 3.7 demonstrates the VRP data collected for female singers from six different sources. The number of subjects per illustrated group varies from 8 to 42. Numerous differences exist between these studies. Most likely, they are due to factors that change from one experiment to another. For example, Hacki's data seemingly has more of a performance nature whereas Pabon and Sulter collected

physiological data. These last VRP norms coincide nicely with Lamesch's data. His recordings were restricted to a maximum frequency of 523.3 Hz and although they depict both mechanisms, M1 is certainly predominant in view of the fairly low range tested. The trend that is appreciable in such a comparison graph is that the overall slope of the singer's average VRP is quite similar despite experimental differences. Also clear is that the definition of the group and the VRP recording approach will greatly influence the final results. As Roubeau *et al.* importantly pointed out, there are differences in voice function and voice use among non singers, amateur singers and professional singers. Those differences are bound to impact the VRP information and should be maintained separately or at least, be well identified.
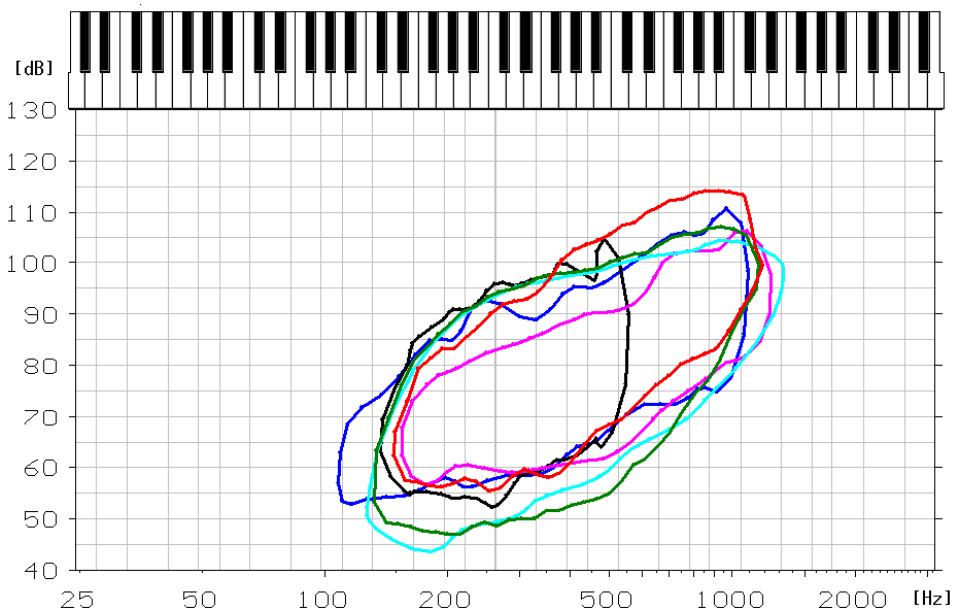
Figure 3.7: Singing voice VRP data compared (female voices only). **Sulter et al. in cyan** , N=42 trained females, Pabon in green, N=23 female singing students, **Åkerlund *et al.*in blue**, N=10 female professional singers,**LeBorgne & Weinrich in pink**, N=17 female graduate singing students, **Hacki in red**, N=10 sopranos, **Lamesch in black**, N=8 amateur and professional sopranos and mezzo-sopranos (constrained range).

3.8 Terminology

In the 1983 Schutte and Seidner standardisation paper [138], a brief discussion is given concerning terminology. Similar concerns as the ones that persist today, seem

to have been present at the time. From these recommendations, it is clear that a suitable term should account for the frequency and sound level plane limitations of the graphic display end result. However, the paper did not recommend a specific term.

Voice Range Profile is a fairly recent term that was adopted by the IALP in 1992. Multiple references to the VRP were already in use: "phonetogram" (a term given to create a voice test equivalent to the audiogram; the term has however a misleading phonetics connotation), "phonogram" (a term used by the Japanese teams but which conflicts with the concept of a *speech sound*; also a phonetic term), "courbes vocales" (named by the Frenchmen Calvet & Mahliac and limited to the contour), "Stimmfeld, Stemgebied or voice field," (frequently used in German, Dutch and by Sonninen and his lab respectively), "voice area", "voice profile", "phonational profile", "F_0-SPL profile" are other terms that are encountered in the literature.

For computerised phonetographs, it would be practical to adopt a term that refers not only to the contour or profile aspects of the voice, but also designates the inner VRP areas and the extra dimensions that can be added to the VRP. The terms "Voice Map" or "Voice Feature Map" are proposed here as possible adequate replacements for VRP.

In the course of the current dissertation work, some terminological issues were also identified in respect to the types of VRP recordings that are conducted in research or in clinical environments. Since instructions and investigation aims can completely change the information obtained in a VRP recording, it is suggested that VRPs should be labelled according to three different types of possible recordings. In fact, a protocol suggestion, including all three types of VRP recordings discussed, is appended to the dissertation.

VRP$_{phys}$

The "standard" or "classic" VRP refers to a physiological VRP (VRP$_{phys}$) measurement intended for the assessment of voice function (muscular strength, control and balance combined). This means that voice quality is not the aim and is usually disregarded completely. Here, the ideal recording would include phonation at the extremes of frequency and intensity that an individual is capable of producing. However, it is well acknowledged that such phonations, typically produced in drastic emotional or survival situations, are inhibited in studio and everyday situations. Hence, the physiological characteristic of the recording is somewhat relative. Sustained phonations are used to perform VRP$_{phys}$ recordings and that has most likely led people in naming these recordings "singing voice profiles" or "singing voice VRPs". This creates considerable ambiguities that should be avoided.

SRP

The speech range profile, or what is sometimes referred to as an habitual or speech VRP, distinguishes itself from the physiological VRP recording in that it specifically

aims at recording continuous speech. Thus, such recording captures voice behaviour patterns that are dependent on different contexts. This type of phonetographic recording was introduced quite early, with the appearance of the computerised phonetograph [152, 140, 56, 136]. This term is rather well coined and does not seem to pose any semantic challenges. In recent studies, it has attracted interest as an integral part of the complete VRP patient/subject evaluation [4, 42, 99, 10]. The tasks included in SRP recordings (reading, counting and/or spontaneous speech) remain disparate. However, it is agreed that the importance of such recordings lies in the testing of the dynamic range and habitual averages of the voice in speech.

Performance VRP (VRP$_{\text{perf}}$)

Wolf *et al.* maintained that "falsetto" measurements of the male singing voice were unnecessary since they did not reflect the performance realities of the male voice [182]. Whether this last statement was justified or not is not of interest here, rather, it is the line of thinking which most likely was a source of inspiration for future VRP research focusing on the singing voice. Coleman *et al.* strove to demonstrate differences between the physiological VRP measurement and what they termed a "musical VRP". The latter designated the *aesthetically* acceptable frequency range of the singer and was quickly endorsed by many other researchers working with the singing voice [126, 141, 48, 136, 3, 94]. Coleman, after discussing the impact of *vibrato* on overall VRP recordings, nonetheless recognised the importance of this singing voice characteristic and included it in his "musical VRP" recordings. However, it is often not alluded to (in the context of a "musical VRP") [141, 136] or is simply precluded [3, 4].

Unfortunately the concept of "musical VRP" or even the "musical range of phonation" is ambiguous. First of all, it is not clear if this "musical VRP" (a reduced frequency range) is designed as a replacement of the physiological VRP. Furthermore, this type of recording is defined by frequency range alone and does not address stamina, energy and other performance-relevant details such as softest/loudest phonations acceptable on stage and overall musicality. Moreover, the semantics of the term can lead to some interpretation inaccuracies, as "musical" is an adjective which typically designates that which contains the qualities of music, is harmonious, and melodious.

For this reason, the present work proposes the term *performance VRP* (VRP$_{\text{perf}}$ which will be henceforth adopted throughout the thesis. The VRP$_{\text{perf}}$ is here viewed in a similar way as the SRP. It is a context-based and behaviour-dependent recording. The VRP$_{\text{perf}}$ is contextual in the sense that it is a reflection of the voice use of a singer performing on the stage (in this case, the opera stage). To achieve this, visualisation of a stage and audience are encouraged during the VRP recording. The performance VRP is also behavioural in that it records the voice such used typically by individuals in their capacity of a singer. Just as the SRP, the performance VRP can be juxtaposed or superimposed onto the physiological VRP to allow comparisons between speech/singing and physiological voice capabilities.

Table 3.4: A summary of the most reported VRP characteristics

Characteristics	Metric	Definition	Study
Frequency	F_0 minimum	The lowest pitch	[182, 158, 23, 179, 37, 155, 136, 135, 170, 4, 49, 160, 7, 62, 106, 167, 145, 154, 42, 99, 180, 95, 10]
	F_0 maximum	The highest pitch	[182, 158, 37, 23, 179, 155, 136, 135, 170, 4, 49, 160, 7, 62, 145, 154, 99, 10, 180]
	F_0 range	F_0 maximum -F_0 minimum	[32, 33, 84, 48, 136, 51, 88, 3, 4, 160, 89, 79, 7, 109, 62, 100, 106, 167, 94, 145, 154, 21, 42, 99, 10, 180, 95]
	MFF	mean fundamental frequency	[84, 155, 40, 51, 4, 160, 79, 154, 145, 42, 10, 95]
Vocal Output	SPL minimum	The lowest SPL	[23, 37, 32, 33, 157, 141, 51, 136, 135, 3, 4, 49, 160, 11, 57, 7, 79, 62, 109, 102, 167, 94, 145, 154, 21, 70, 99, 10, 69, 180, 95]
	SPL maximum	The highest SPL	[182, 23, 37, 32, 33, 84, 155, 141, 51, 136, 135, 3, 4, 49, 160, 11, 57, 79, 7, 109, 62, 106, 102, 167, 94, 145, 154, 70, 99, 10, 69, 180, 95]
	SPL habitual	A comfortable sound level	[3, 4]
	SPL range	SPL maximum-SPL minimum	[32, 33, 87, 141, 48, 110, 136, 51, 40, 88, 13, 170, 4, 160, 89, 7, 100, 62, 142, 16, 154, 145, 42, 99, 180, 10, 95]
	SF	The intensity of the singer's formant	[141, 48, 88, 106]
(can also define voice quality)	Coefficient of sound	Quotient of SF and SPL maximum, in percent	[14, 141, 20, 145]
	mean (or Leq) SPL	Average SPL (or Equivalent continuous noise level)	[40, 51, 4, 79, 144, 145, 154, 10, 95]

Table 3.4:   (continued)

Characteristics	Metric	Definition	Study
Area	Area	The space contained within the VRP lower and upper contours	[158, 23, 110, 51, 88, 13, 1, 49, 160, 89, 109, 102, 16, 72, 167, 154, 145, 99, 69, 10, 180]
Shape	Shape	The morphological attributes of the VRP contour and area	[160, 62, 63]
Slope	Slope	The covariation relationship of F_0 and I	[14, 84, 155, 40, 51, 115, 3, 170, 160, 89, 79, 62, 112, 17, 113, 145, 154]
Smoothness	Smoothness	Regularity and evenness of the VRP contour (most often the lower contour)	[86, 85, 152, 51, 160, 7, 154, 69]
Quality	Jitter	Duration deviations from period-to-period	[117, 111, 113, 17, 154]
	Crest Factor	The ratio of RMS to peak amplitude, per period	[111, 113, 15, 17, 18]
	Hoarseness	Harsh, rough quality of voice associated with disphonia	[60]
	L_0	The level of the fundamental	[51]

Table 3.5: A survey of studies that address the singer's VRP. Computerised VRP studies are italicised. Mezzo-soprano=mezzo; soprano=sopr; baritone=barit; f=female; m=male; amat.=amateur; stud.=student and prof.=professional. The third column indicates the total subjects that participated in the study (in bold) and when reported, the breakdown of the subject distribution according to gender and voice classification.

Investigator(s) Year	Group Definition	N	dB weighting	Contribution /Aims	Results
Wolf et al. (1935) [182]	a) Various singing levels (m,f) b) Singers	a) **50** b) **5** barit-5	dB rel 1μW	Investigate the relative maximum intensities over ranges, and for different vowels	Vocal power increases smoothly with F_0 Vocal power levels are higher for [a],[e] than for [u],[i] Reference curves are drawn according to the highest intensities for each frequency found amongst the 50 singers
Stout et al. (1938) [158]	Male singers	**3**		Study the effect of intensity and frequency on vowel structure in singing	The increase of SPL with F_0 varies as a function of vowel articulation (the greatest increase noted for vowel [i] and the smallest for the vowel [a]) When frequency is constant, increase in intensity will enhance higher than 1800 Hz partials, a decrease of these partials relative intensity is found for the opposite condition
Calvet & Malhiac (1952)[23]	Boy choir singers (age 4-18) Recorded every 6 months	**± 100**	measured in phon	Study the effect of puberty in vocally trained boys	The measure of intensity is valuable in tracking voice maturity and training Despite error factors, the intensity curve of a voice is able to depict laryngeal possibilities of a voice.
Coleman et al. (1978)[33]	Girl singers (age 10-13)	**9**	not reported	Compare the musical VRP to the physiological VRP	Girls have smaller VRPs than adult females Musical VRPs are more restricted than the physiological VRPs (F_0 range, minimal/maximal vocal output for sustained phonation)

Table 3.5: (continued)

Investigator(s) Year	Group Definition	N	dB weighting	Contribution/Aims	Results
Bloothooft (1981)[14]	*Prof. singers*	**14**	Db(C))	Explore the possibilities of computerised phonetography	Voice mechanisms can be mapped in the VRP. The difference between total intensity and the intensity of the singer's spectral cluster peak is not only relevant to the singing voice but can be used to map spectral balance in the VRP.
Klingholz & Martin (1983)[84]	Singers	not reported	not reported	Report the elliptical analysis approach applied to the singing voice	Singers have higher SPLs Singers manage voice mechanism transitions smoothly Singers have greater frequency and dynamic ranges
Pedersen (1984)[122]	Girls form a singing school (age 8-19)	**47**	dB(A)	Investigate the possibilities of predicting voice puberty occurrence in girls	VRP discontinuities greater than 5 dB can be attributed to voice mechanism transitions A general change of VRP area in relation to menarche VRP is significantly related to age
Seidner *et al.* (1985)[141]	Prof. singers, singing stud.	**60** 30 prof., 30 stud. (sopr, mezzo, alto, tenor, barit, bass)	dB(A)	Investigate spectral qualities in simultaneous to VRP recording. Effects of vowels are explored ([a][i][u])	Male singers can be distinguished best by the difference between max. intensity and the singer's spectral peak resonance. Female singers are best differentiated with the intensity of the spectral peak resonance. There is no detectable effect of SPL dependence on vowel in female singing.

Table 3.5: (continued)

Investigator(s) Year	Group Definition	N	dB weighting	Contribution/Aims	Results
Gramming et al. (1988) [54]	Male prof. singers	**9**	dB(C)	Compare a mean VRP to the mean VRP of nine males, non-singers.	One difference found: male singers have a higher upper contour in the high frequencies range.
Konzelmann et al. (1989)[88]	Lay choir singers	**66,** (tenor-8, bass-16 sopr-27, alto-15)	not reported	Investigate the effects of vocal loading on the singer's VRP	Voice range, dynamics and area are greater for singers. Loading effects lead to greater metric values (especially for males) with exceptions for altos and soloists.
Pedersen (1990)[123]	*Choir girls (age 8-19)*	-		*Compare voice category, hormones, puberty stages to the VRP*	-
Hacki (1990)[59]	*Prof. singers*	**20,** *(sopr-10, alto-10)*	dB(A)	*Create averages for professional singers without rescaling the data.*	*Sopranos have a more restricted dynamic range in the mid-frequency range than altos.* *Altos are able to maintain a flatter lower VRP curve than sopranos*
Büttner et al. (1991)[20]	a) Prof. sopr and barit b)Beginner singers (at 0 lessons, 60 lessons and 90 lessons)	**a) 2** sopr-1, barit-1 **b) 374**	dB(A)	Propose the coefficient of sound as a potential measure for voice training and voice quality	Increase of the coefficient of sound with voice training for all voice categories. The coefficient of sound varies negligibly with vowel and frequency in professional singing voices.

Table 3.5: (continued)

Investigator(s) Year	Gender and Group Definition	N	dB weighting	Contribution/Aims	Results
Awan (1991)[3]	University choristers, mix gender, no voice categories	**20**	dB(C)	Investigate differences in VRP contours between trained and untrained adults	Singers have increased F_0 ranges, max/min and comfortable SPLs at all frequency levels except the extreme lowest part of the F_0 range.
Åkerlund et al. (1994)[80]	Female Western lyrical prof. singers (sopr, mezzo and alto)	**10**	not reported	Investigate if vocal behaviour is differently manifest in the VRP, mean sound levels and F_0 in speech	Singers have a greater vocal range and upper VRP contours. There was a significant difference in vocal output between the triad task and the sustained discrete pitch task.
Pedersen (1993)[119]	*Choir boys (age 13-15)*	**3**	*dB(A)*	*Longitudinal tracking of the effects of puberty on the VRP (VRPs recorded every 2 months for a year)*	*The VRP changes are noted in respect to total area, lowering of the minimum F_0 and more pronounced register dips. A general restricted dynamic flexibility is noted. The min. F_0 was significantly related to binding globulin (a sex hormone).*
Åkerlund & Gramming (1994) [77]	Female Western lyrical prof. singers (sopr, mezzo and alto)	**10** (4,4,2)	not reported	Investigate to what extent high P_s values contribute to higher upper VRP contours (or higher SPL in loud phonation)	The higher VRP upper contours found in singers are tied to the use of higher P_s Upper contour differences cannot be explained by P_s levels alone but also include acoustic strategies.

Table 3.5: (continued)

Investigator(s) Year	Group Definition	N	dB weighting	Contribution/Aims	Results
Sulter et al. (1994)[161]	*Female and male chorister (no voice categories)*	**85,** *(42-f 43-m)*	*dB(A)*	*Study the differences between trained and untrained groups*	*Characteristic shape difference between males and females. Males have greater min SPLs yet, females phonate louder at VRP extremes.* *Singers have larger dynamic ranges, especially in soft voice. No local minimum found at expected voice mechanism transitions.*
Mürbe et al. (1999)[106]	Conservatory students	**25,** (sopr-5, alto-5, tenor-5, barit-5 basses-5)	dB(A)	Longitudinal tracking of singing training with the VRP of students recorded over the span of 4-5 years	Increases of mean overall SPL for vowels [a],[i],[u]. Increases in SPL related to the singer's spectral peak resonance, a decrease of variation of overall SPL with frequency, especially related to the spectral peak resonance band.
LeBorgne & Weinrich (2002)[94]	*Conservatory graduate singing students*	*21 (sopr-17, mezzo-1, tenor-2, barit-1)*	*not reported*	*Tracking with the VRP intensive voice training over a 9 month period.*	*Expanded frequency ranges and lower minimum SPLs*
Roubeau et al. (2004)[131]	Amat. and prof. singers (both sexes)	**33** (21 amat., 11-f,10-m), (12 prof., 7-f, 5-m)	dB(A)	VRP recordings for separate voice mechanisms	Mechanism ranges are identical for males and females. Prof. singers have greater ranges than amateurs. Amateurs have greater ranges than non-singers. Prof. singers have a greater overlap between both mechanisms in max. intensity. Min. intensities are comparable for all subjects.

Table 3.5: (continued)

Investigator(s) Year	Group Definition	N	dB weighting	Contribution/Aims	Results
Hunter et al. (2006)[70]	*Prof. singers*	*4* (*sopr-1, mezzo-1, tenor-1, basse-barit-1*)	*dB(C)* (*later scaling to dB(A)*)	Investigate the relevance of a perceived VRP (based on equal-loudness)	*VRPs are similar to previous reports, some vowel variation effect is observed for all singers* *The perceived dynamic range (in the PVRP), is much greater than the the VRP's.* *A-weighting underestimates the most sensitive region of the ear (the location of the singer's spectral peak resonance) by nearly 10 dB* *The overall perceptual level construct (OPLC) allocates a single value associated with the auditory system and grades the perceptual difference between trained and untrained*
Lamesch (2007)[92]	Prof. singers	**2** (counter-tenor-**1**, sopr-1)	dB(C)	Investigate "voix mixte"	There is a large overlapping area between the M1 and M2 VRP contours M2 intensities are contained within M1
Lamesch (2008)[91]	Prof. and amat. singers	**20** (sopr-4, mezzo-4, counter-tenor-2, tenors-4, barit-5, basses-2)	dB(C)	Influence of the vowel on the phonetogram	Phonetogram contours differ across vowels in M1 but not in M2 Similarly, open quotient values vary across vowels in M1 but not in M2 Laryngeal mechanisms are important to include in the VRP recording of the singing voice.

Chapter 4

Methodology

This section gives an overview of different measurements and recording set-ups that were used throughout the course of the present work. Further details specific to a particular paper can be found in the methodological section of the respective paper.

4.1 Voice Measurements

Voice Range Profile

For all studies, with the exception of Paper V which did not include VRP recordings, VRPs were automatically recorded using *Phog*, (Version 2.00.10, Saven Hitech AB, Sweden). In parallel to the VRP recording, *Phog* records a corresponding audio file. This audio file enabled the Matlab processing of VRP recordings for Papers I, III and IV. Phonetograph settings were the same for all experiments, including the maximum standard deviation threshold taken over 7 periods and a 0.025 second minimum for the voicing threshold. The latter threshold was chosen so that even a single vibrato cycle excursion would be registered. This is an important detail to take into consideration when measuring the singing voice. Since the object of study was the performance voice of Western lyrical singers, vibrato was included *de facto* in VRP recordings. In Figure 4.1a, 4.1b and 4.1c the VRP differences for no vibrato, little vibrato or typical vibrato are illustrated.

For a major part of Paper I, complete Paper II and Paper III, recordings took place in a recording studio. The room's characteristics are listed in Table 4.1. Figures 4.2a and 4.2b illustrate the sound response characteristics of the room in question.

This room was sound treated and isolated yet not anechoic. Subjects performed alone in this room while the investigator attended to the recording in an adjacent room. Visual communication was possible through a window; however, subjects had no access to VRP feedback. A fixed omnidirectional microphone was used and adjusted to the height of each subject. Microphone-to-mouth distance was rigorously controlled between each task so as to keep a constant 30 cm distance

(a) no vibrato   (b) ± 45 cents vibrato   (c) ± 90 cents vibrato
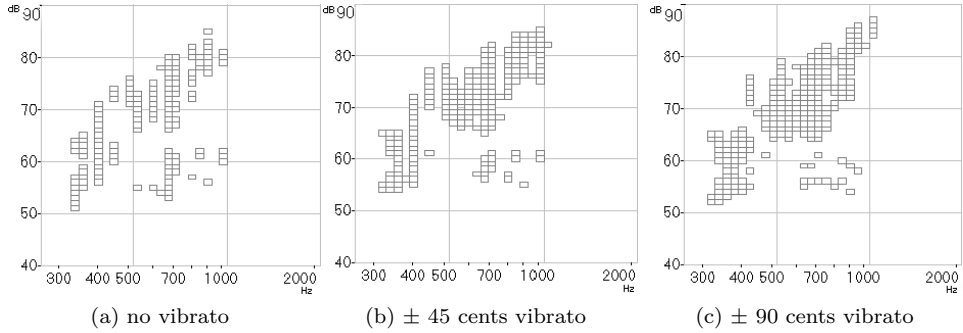
Figure 4.1: The vibrato impact on the overall VRP was explored by a repeated synthesis of a musical phrase. With a 0.025 s accumulated occurrence threshold, a) depicts results for no vibrato, b) for a ± 45 cents vibrato and c) for a ± 90 cents vibrato.

Table 4.1: Recording Studio Characteristics

Studio Characteristics	Measurements
Volume	45 m^3
Ceiling Height	2.74 m
Reverberation Time, T30	0.1 s
Reverberation Radius, across the spectrum	> 1.2 m
Absorbent depth	0.5 m

(a) Reverberation Radius   (b) Reverberation Time
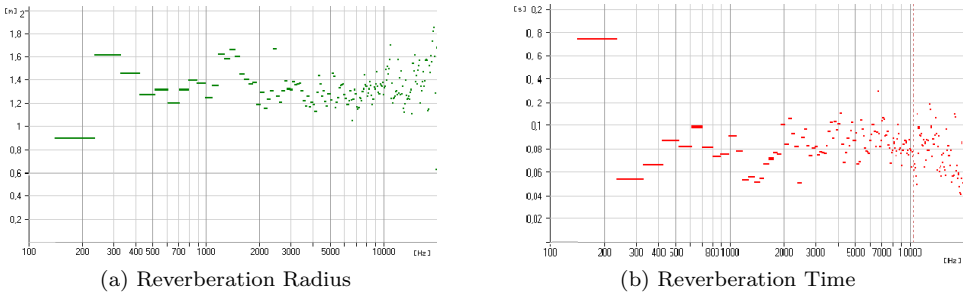
Figure 4.2: Recording Studio specifications

to the microphone. To help subjects maintain their position, pieces of tape were placed on the floor to monitor feet alignment and a waist-high divider wall (fiber glass wool covered by cotton material) was used to provide some back support and to delimit the stance position.

Part of the work for Paper I took place in Montreal. The recordings were performed with a portable platform (laptop and portable DSP card) and the same equipment as mentioned above. In this case, a typical audiology sound booth was used for the recordings. This meant that both equipment and investigator were present in the room as the subject performed the tasks. Visual feedback remained unavailable to the subject.

In the case of Paper IV, recordings were conducted in the University St-Luc Clinic. A slightly different portable platform was used for these recordings. Included were a laptop, a portable DSP card, a smaller two-channel preamplifier and a cardioid head-mounted boom microphone instead of a fixed omnidirectional microphone. The equipment as well as the investigator were in the same room with the subject. At times, a video camera as well as an observer were also present.

The equipment used for Studies I to IV is tabulated in Table 4.2.

Calibration

- Fixed microphone procedure: A Brüel & Kjær calibrator generating a 1000 Hz tone at 94.9 dB SPL re 20 μPa was used to calibrate the condenser microphone. *Phog*'s calibration settings were adjusted to match this reference tone. Calibration was performed for each subject. Due to some limitations of the software, the microphone-to-mouth distance needed to be increased to one meter for certain singers in order to avoid *Phog*'s saturation. In this case, a correction of 10.5 dB was applied at the data processing stage.

- Headset procedure: The calibrations performed for the clinical portion of the recordings were performed by help of white noise generation through a speaker. Positioned at the microphone, a quality sound level meter (LA-210, Ono Sokki, Japan), set to linear weighting was used to measure the speaker's output. *Phog*'s settings were adjusted according to the reading of sound level at that position and microphone-to-mouth distance was compensated for in order to obtain a 30 cm distance. The microphone placement was carefully monitored and the distance was systematically measured from between the front teeth to the boom for each subject.

- Pressure procedure: A pneumotach calibration unit was used to calibrate the pressure transducer. Readings at various increments of cmH_2O (20-10-5) were taken and recorded in file. These files were later used to calibrate the DC-coupled P_s channel of the recordings. This complete calibration procedure was performed for each subject.

- Accelerometer procedure: Since the accelerometers used in Paper II were provided by the NCVS, they were calibrated according to the NCVS accelerometer calibration protocol [124]. However, in Paper II, the SAL signal was later normalised and so calibrations were no longer necessary.

The Augmented Phonetograph

In order to record and map subject self-perceptions into the VRP, *Phog* was augmented with a hand-held device. [1] The button signal was recorded synchronously into the audio file where the button status information was stored in a vacant channel. Each depression of the button resulted in a fixed 73 ms pulse regardless of the duration or force of pressing. The maximum point of this pulse was retained for analysis. Button presses occurring in unvoiced portions were discarded by the system. In order to function also with an AC-coupled input, the detection scheme used interruptions in a sentinel tone, on a separate, inaudible channel, to signal the switched state of the button.

4.2   Measurement of Intraoral Pressure

The subglottal air pressure was an important dependent parameter in Paper II. To measure P_s, a non-invasive estimation method was adopted. This estimation method is based on the observation that intraoral pressure peaks, obtained during the elocution of a string of [p] occlusions in a series of [pae, pae, pae], can be deemed equal to P_s [129, 96]. Since the glottis is open during a [p]-occlusion, while the flow is interrupted, the intraoral pressure can be considered equal to the pressure under the glottis. This estimation is quite practical due to its non-invasive nature. However, it is also very sensitive and difficult to measure correctly. For example, the production of [p]'s needs to be fluent without concomitant modulation of the lung pressure. Meticulous attention must be directed towards this important detail. In Paper II, P_s was monitored using a storage oscilloscope while subjects performed the tasks. To obtain correct measurements, subjects needed frequent reminders to refrain from singing musically the [p]-occlusions .

When addressing the singing voice, some attention must be given to a few limitations of this type of measurement. In a study where the P_s estimation method was compared to direct P_s measurements, Kitajima and Fujita [81] found that the accuracy of the estimation method was quite high, as long as P_s was lower than 25 cm H_2O. This is perhaps not a matter of concern when addressing speech production where average normative values range from 6 to 10 cm H_2O, however, when addressing the singing voice, which generally requires higher pressures, estimations

---

[1]Internally, Phog uses a block-based signal-processing tool (Aladdin Interactive DSP 3.0 from Saven Hitech AB, Täby, Sweden) which interprets the signal model at run time. With some prior knowledge of the system, certain modifications can be made to the signal model without requiring recompilation. It is thus possible to prototype limited changes to the system with a short turn-around time. This is how the button mechanism was added.

yielded lower pressure values than for direct measurements. Also, the repeated [pae] string might not be true to typical P_s use in the context of singing where many various variables are constantly changing.

Table 4.2: Equipment Included in Studies I to IV

Equipment, set-up and data collection	I	II	III	IV
Phog (Version 2.00.10, Saven Hitech AB)	x	x	x	x
DSP sound card (BlueWaves LSI-PC/C32 board with DC coupled input)	x	x	x	
Mobile DSP sound card (CAC Bullet II DSP with AC coupled input)	x			x
Fixed personal computer (Microsoft Windows XP)	x	x	x	
Laptop (Microsoft Windows XP)	x			x
Omnidirectional microphone (Brüel & Kjaer, model 4003 or 2238)	x	x	x	
AKG microphone model 420, cardioid (headboom)				x
2 accelerometers (Thin Case BU-7135 Knowles Acoustics)		x		
Surgical glue (Mastisol®) and suture strips (TS 3101 Derma Sciences)		x		
Microphone preamplifier (model 2MP Line Audio Design)				x
Line amplifier (Nyvalla-DSP Audio Interface Box)	x	x	x	
Sound level meter with linear weighting (LA-210, Ono Sokki, Japan)	x		x	x
Sound level meter with slow A-weighting (Brüel & Kjaer 2238 Mediator)		x		
Electrical -12 dB pad	x		x	
Earphones used for prompting the subject	x		x	x
0.025 accumulated time threshold	x	x	x	x
75 cents maximum for F_0 standard deviation over 7 phonation periods	x	x	x	x
Pressure transducer (Glottal Enterprises PT series)		x		
Pneumotach (Glottal Enterprises Model MCU-4)		x		
Storage oscilloscope		x		
Regular clinical room				x
Recording studio (45 m^3, 3 m high ceiling)	x	x	x	
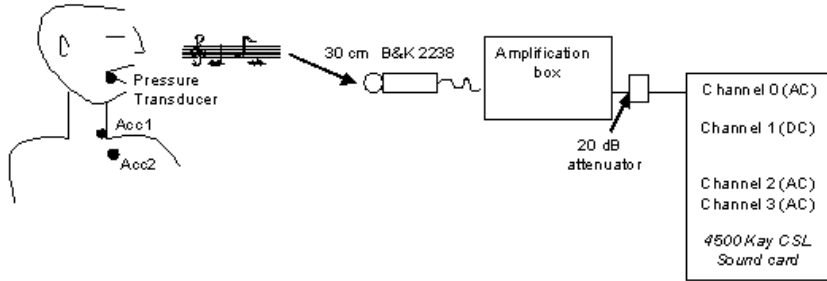
Figure 4.3: Experiment Set-up for the audio, aerodynamic and accelerometric measurements reported in Paper II. Acc1 indicates the position of the accelerometers at the jugular notch and Acc2 the sternum bone.

4.3   Measurement of Skin Acceleration Levels

In Paper II, skin acceleration levels (SAL) were recorded synchronously with SPL. SAL is a measure of tissue vibration and it can be recorded near the vocal folds by help of accelerometers that are attached to the skin. Previous research has demonstrated that the colliding forces of the vocal folds have little effect on the overall vibrations that are registered in the vicinity of the trachea and the sternum[162]. SAL was measured near the vocal folds as an noninvasive measure of phonatory activity. The vibrations of the thyroid and the sternum lamina are mainly related to the voice source; one obtains a highly voice source dominant signal with little vocal tract influence. This measure has gained particular interest in the voice science field since it could have the potential to be an estimate of the intensity of the glottal source rather than the intensity of the radiated sound. Very small accelerometers are used as phonation sensors. In Paper II, two accelerometers were attached to the skin of the neck. One accelerometer was fixed at the jugular notch (the anterior part of the neck and between the cricoid cartilage and the sternum) and the control accelerometer was fixed to the sternum bone. The accelerometers were glued to the skin with surgical adhesive, and suture tape secured the accelerometer body to the skin. The set-up for the experiment performed in study II is displayed in Figure 4.3.

4.4 Qualitative Instruments

Papers I, III, IV and V all included self-administered questionnaires. These questionnaires are reproduced verbatim in the article appendices.

Paper I: The questionnaire resembled the format of a typical medical health history form. It comprised 4 sections: (1) voice classification (three characteristic closed-ended questions and one opinion open-ended question); (2) body typology and hormonal cycle information; (3) vocal habits (a characteristic closed-end question on the singing voice, frequency questions relating to speech and singing voice use, and a verification of warm-up time prior to recording; (4) vocal health history (in open-ended question format).

Paper III and IV: The questionnaire was structured with two types of questions. Half of the questionnaire included opinion and self-classifying questions that were answered by the use of VAS scales. Each scale had two extreme semantic anchors organised systematically in the same rank: negative connotation to the left and positive connotation to the right. The other half of the questionnaire collected enumerative and descriptive responses.

Paper V: The structure of the questionnaire tested in this paper has been standardised [74] and works on the principle of Likert scaling. For 30 items, subjects crossed circles ranging from 0-4 where 0 was equivalent to "never" and 4 was equivalent to "always". A maximum of 120 points could be obtained. High scores were attributed to perceptions of severe vocal disorder and low scores were considered typical of a healthy vocal state. This form was made available in two different formats. The format used with healthy subjects was the same as described above and was available on the Internet as well as in hard copy. A VAS scale was included in patient questionnaires. A portion of the questionnaire was also designed to collect personal data and details relevant to vocal genre, level of training and context of vocal use. All these questions were formulated in a closed-ended fashion and one contingency option was included.

Tasks

In what follows, the different tasks used in the four first papers are illustrated. Detailed descriptions can be found in respective papers. A VRP protocol suggestion including all three types of phonetographic recordings, SRP, VRP_{phys} and VRP_{perf} is also included in Appendix A.

Paper I, III, IV: These studies included two speech tasks. The subjects first were asked to speak on the theme of their vocal warm-up routine for 1 minute. Other suggested themes were favourite performances or history of voice training. Secondly, subjects counted from 21 to 80. For each increment of 20, the subject paused and contextual instructions were given ("count to sleep a baby in arms without whispering", "have a typical dialogue with a friend", and "give a seminar to a minimum of 50 people").

The definition and the motivation of the VRP_{phys} were explained to each subject. Various examples of maximal and minimal phonation with no concern for voice quality were given, and the non-singing aspect of the exercise was stressed. Subjects first started at a mid-range pitch, in an *as soft as possible* dynamic and performed a long and slow descending pitch glide. Several glides were repeated and subjects were asked to perform similar vocal gestures in a shorter format. The instructor gave initial pitches and instructed on the length of the pitch glide. The same procedure was repeated in an ascending direction. Here, it was often useful to shorten the length of the initial glide in order to obtain stable soft phonation. When results were felt to be representative of the subject's capacity and willingness to phonate at a lowest effort possible and at lowest and highest pitches, the procedure was repeated in an *as loud as possible* dynamic. Subjects were often redirected to a non-singing or non-aesthetic phonation.

For the remainder of the recording, subjects were urged to use their singing voices *only.* Prior to each task, subjects were asked to visualise themselves on stage with an orchestral accompaniment and a full audience and to perform the task according to what would be musically and dynamically acceptable to them in such a context. Prompted musical tones (C-E-G-A musical notes) were sung in a *messa di voce.* The first pitch, G_4(for soprano), E_4(mezzo-sopranos) or C_4 (contraltos) were followed by lower increments of the pitch range. Next, singers sang again the initial pitch, after which higher increments of the pitch range were sampled. When prompted pitches exceeded the range deemed musically acceptable in performance, smaller pitch increments were prompted until the singer signaled the range completed.

Next, a musical triad exercise was performed. The same instructions as for the prior task were given to the subjects. Subjects were free to phrase and accentuate the task as they wished. The singer chose an appropriate mid-range point to begin a descending, (stage) soft rendition of the triad. Again, singers were left the judges of a stage-appropriate lowest pitch. When the lowest pitch had been sung, the initial chosen pitch was replayed and the triad was sung ascending in the upper part of the range. With the exception of the speech tasks, all tasks were performed on the vowel [a] (many variations of the vowel were accepted).

Finally, subjects performed an excerpt of their best audition piece with text. In cases where the triad task was repeated, the aria excerpt was performed in

between replications. An illustration for these various tasks is given in Figure 4.4a.

(a) Tasks Devised for Paper I, III and IV

(b) Tasks Devised for Paper II

Figure 4.4: The tasks utilised in four of the studies presented in this dissertation work

Paper II: The subjects performed a series of tasks in which different parameters were maintained fixed while one parameter was varied. For example, subjects sustained a pre-determined tone in one musical dynamic while alternating vowels in a slow tempo. Three music dynamics (*p, mf, ff*), at a comfortable and high soprano pitch were tested. An ascending scale ranging an octave was also performed. Singers simply chose a comfortable starting pitch and repeated each pitch three times. The scale exercise was performed for both, the vowel [a] or [i] and all three musical dynamics mentioned above. A third task consisted of an arpeggiated octave starting at 349 Hz in which each tone was repeated three times. This exercise was performed for five vowels (/i e a o u /) and for all three musical dynamics. The tasks are illustrated in Figure 4.4b.

Ethics

An ethical vetting from the *Regional Ethical Review Board in Stockholm ("Regionala etikprövningsnämnden i Stockholm"*, certificate 1358-31) was obtained for the studies included in this thesis work. Before the recordings were performed, subjects received written information on the project especially regarding the project's purpose. The protocol for the experiment in question was also made available in writing. Finally, subjects signed a consent certificate where subject rights and interests were identified clearly. For studies I and III, subjects were given a small remuneration for their participation. Subjects in Paper II were vocology students who all voluntarily agreed to participate in the experiment's recordings. Subjects involved in Paper IV were not remunerated but were given the opportunity to take part in an extended evaluation session. Participants in the control group for Paper V were recruited on a voluntary basis in rehearsals, music schools, master classes or on Internet.

Chapter 5

Aims and Results

5.1 Overall Goals

The following overall goals provided the main impetus for this dissertation work:

1) the creation of singer-related resources for the clinician

2) the adaptation of common clinical tools to the concerns and the reality of the singer

3) the creation and documentation of normative references for the singing population

It was imperative for this work to depart from clinical realities and to direct research in relation to relevant pre-existing tools. Experiments conducted for this thesis employed two clinical instruments which have rapidly become part of the standard battery of measurement tools for clinicians. The VRP, plotting the region of the fundamental-frequency and the intensity space over which a speaker or singer can phonate, was selected for its capability to be particularly sensitive to the singing voice. Attention was also given to the Voice Handicap Index (VHI), which has demonstrated sensitivity to the patient's experience of vocal disorder, giving more room to patient perception and thereby weighing considerably in the evaluation process.

5.2   Importance of the Present Work

The empirical studies included in this thesis have mainly investigated the quantitative use of the VRP in relation to the singing voice. With the clinical evaluation of the singing voice and basic research perspectives, the work generally focused on three informally reported/observed problems:

1. Singing voice complaints are often not accompanied by speaking voice difficulties

2. Singing voice disorders may be difficult for the clinician to detect perceptually

3. Published VRP data for any specific singing voice group is rather scarce

The process of the clinical evaluation of the singing voice remains a fairly subjective process, for which the clinician's own singing experience and knowledge are considered essential. This work was particularly concerned with cases for which no such singing voice expertise is available.

   Although it is well known that singers often have vocal complaints specifically related to singing, and not necessarily to speech, very little has been done to adjust the clinical process accordingly. This mismatch to patient needs can be assessed in reports such as that of Rosen and Murry [128]: singer VHI scores showed no differences between healthy and pathological voice groups. In the example of the VHI, singer adaptations of this psychometric instrument were not addressed until 2005 [105]. Sataloff and Benninger [134, 9] assert the importance of integrating performance with the overall clinical evaluation of the singer; yet, other than stroboscopy or high-speed vocal-fold imaging, there are no formal, objective procedures followed for such evaluation. The current work looks at singing-voice-specific tasks in relation to VRP recording, the relationship of SAL to SPL in singing, and tests a Swedish adaptation of the VHI for the singer.

   With respect to item 2 above, part of this work also attempts to fill the gap between patient and clinician perception. The problems, as experienced or as reported by the professional singer, are typically very subtle. These subtleties are usually not detected by mainstream voice function measures, not necessarily evident in the acoustic signal of the voice, and are even less obvious to the untrained ear. To the author's knowledge, no previous work has explicitly taken this issue to task. This work attempts to offer a novel solution to what can often represent a real challenge to the clinician with no singing voice training. In the process, the results have proved to be of interest also in generalized clinical and singing voice pedagogical frameworks.

   As described earlier in Section 3.7, few studies can be found on the VRP in conjunction to the singing. When such studies exist, they often fail to clearly identify different singing styles, singing proficiency or even voice classifications, perhaps because these precisions were deemed unessential for study purposes. The present work needed to closely consider these issues in creating representative and quality normative data of the singing voice.

5.3 Original Contributions

Paper I presents the VRP recordings of 30 well documented professional opera female vocalists, recorded in a controlled environment. A stage and singing relevant approach to VRP recording is compared to the usual physiological VRP. The differences found in the present studies indicate the importance of considering the performance aspect of the singing voice. Task design as well as voice classification are examined to throw light on their possible effects on VRP outcome. New singing voice VRP metrics are suggested that quantify the VRP area above 90 dB and dynamic extent in an F_0 independent way.

Paper II follows the work of Švec *et al.* in which SAL is usefully employed to estimate long-term SPL in speech [166]. The SAL-P_s relationship was investigated for the singing voice. The results demonstrate that for the singing voice, such a relationship is weaker than the one found for the SPL-P_s counterpart. SAL does show the possibility to facilitate VRP interpretation in that, compared to SPL, it is much less dependent on frequency.

Paper III tests a novel approach for merging the singer's self-perception into the VRP. Singers were provided with a button device that they used to indicate vocal difficulties as they sang. The feasibility of such an approach was confirmed by the results of a consistent button press behaviour in spite of sparse button press rates. The button information is not spurious and reflects an underlying cause.

Paper IV continues the work of Paper III with singer patients. Patterns of button pressing for this group were distinctly different from patterns observed in Paper III. Button presses do not seem to necessarily coincide with audible symptoms of voice difficulty. Consequently, the mapping of the singer's perception gives a new non-acoustic and singer-relevant information.

Paper V translates and adapts the VHI for singers. This work, based on the initial work of Morsomme *et al.*, clearly indicates the need to address the singer patient according to his/her needs and language.

5.4   Summary of Studies

In the following, an overview of each paper is presented with respect to aims and major findings.

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

Aims

Most reported studies that compare singing to speech recruit a mixture of different levels and styles of singers when conducting singing voice VRP recordings. Very few VRP studies have specifically focused on the singing-voice alone [94, 70, 131, 91]. The VRP is known to be particularly sensitive to gender, to age, as well as to vowel and to individual characteristics [160, 184, 58, 53, 52, 119]. It would follow that the VRP could also be dependent on training, profession and even style [160]. It was thus considered important to document and collect VRP phonetographic data for singers. Because of the lack of available experimental information on the elite singing voice and with the VRP sensitivity in mind, a very specific singing subject group was defined. This investigation's main aim was the investigation of a singing-voice-relevant approach to VRP recording: the effect of tasking, meaningful VRP features and voice classification effects. Thirty professional female opera soloists participated in the recordings and filled an extensive vocal health questionnaire.

The questionnaire responses outlined a vocally experienced and healthy group. The age distribution of the group was rather well balanced across all three voice categories, with ages ranging from 20 to 55 years. The mean age was 33.7, $\pm$ 8.8 years. Subjects had extensive training and professional experience ranging from 12 to 27 years. Overall, singers reported daily or more frequent training (excluding rehearsals and singing lessons or coachings) in sessions of little over than an hour. Subjects generally rated their daily use of both speech and singing voice as *moderate*. The questionnaire also included a health section where subjects reported body length and mass, physical activity, general medical history, voice health history as well as medical or homeopathic intake. 73.3 % of the group had a healthy BMI. In average, subjects were engaged in physical activity (of minimally 15 minutes) three times a week. Finally, 43 % of subjects reported no medicinal intake whatsoever. 23 % reported regular intake of hormonal contraceptives.

Results

This paper's major finding was that a VRP_{perf} differs considerably from a physiological VRP_{phys}. This difference is interesting as it uncovers the importance of examining vocal behaviour within context as much as possible. The performance aspect of voice seems detrimental to the complete assessment of the singing voice. Without a group of truly high-performance vocalists, these results most likely would not have been obtained.
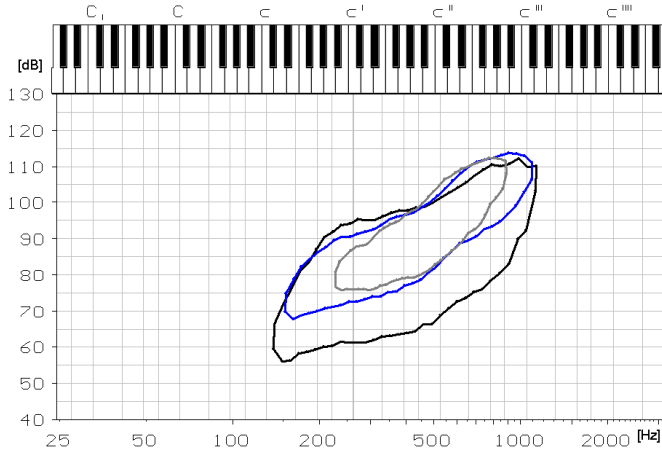
Figure 5.1: A physiological contour (black) average compared to two types of VRP_{perf} (the vocalise and the aria) averaged contours (shades of blue). This figure appears in Paper I as Figure 10.

Figure 5.1 depicts the differences that were found in physiological and performance averaged contours.

Most of the difference found was attributable to the amount of voice used in high-intensity regions of the VRP. Maximum intensity values however, did not significantly differ from VRP_{perf}. When task effects were examined, no significant tasking effect could be detected in the VRP_{perf} recordings, yet VRP_{phys} results, especially in concern to the lower curve, demonstrated potential tasking effects.

Two noteworthy VRP metrics for the singing voice were introduced: $Percent_{\geq 90dB}$ (the percentage of the voice area equal and/or above 90 dB) and the SPL_{ext} (the level difference between the upper and lower bounds of the contour, averaged from lowest to highest F_0). The only effect of voice category for the female singers studied here was observed for minimum and maximum frequency VRP features.

Paper II: An Exploration of Skin Acceleration Level as a Measure of Phonatory Function in Singing

Aims

In working towards the adaptation of existing clinical equipment to the reality of the singer, this paper examined the possibilities of further integrating voice function (voice source related information) to VRP recordings. SAL was used as a means to address voice function non-invasively and more directly. The relationship between SPL and F_0 is especially complicated in singing. For example, it could be useful

to reduce the variations within and across tones in VRP recordings. Because P_s drives the vocal folds and is a main determinant of voice intensity, its correlation to SPL and SAL was compared. It was hypothesised that SAL would correlate better to P_s and thereby be a suitable substitute for SPL. For VRP recording, such a substitution would imply a facilitated interpretation, since information displayed would be more directly related to voice function. Furthermore, the effects of vowel variation could then be reduced and thereby warrant the inclusion of different vowels in the clinical evaluation without incurring important signal variations. Because SAL is measured with contact microphones, more physical and vocal freedom could be given to subjects during recordings (a critical detail in recording singers) and the substantial influence of environmental noise during clinical recordings could be reduced.

Results

Three valuable outcomes will be mentioned here. Firstly and most importantly, the relationship between SAL and P_s could not warrant the replacement of SPL by SAL. Indeed, the correlation of SAL to P_s was rather weak, while the SPL data clearly followed the trends described in the literature, $\pm$ 12 dB per doubling of $\mathtt{P_s}$ (see Section 2.2). Figure 5.2 demonstrates those results.

Spectrally, SAL is dominated by the level of the first partial. This result is understandable given the low-passed nature of the SAL signal. Although increases of P_s mostly tend to boost higher spectral components, when they are compared to the dominant first partial, they probably remain too weak to affect the overall signal level. Further specific testing should be done to test this hypothesis.

Secondly, in the singing voice, SAL is capable of displaying more immediate information whereas SPL includes a variety of factors (dominated by F_0) that impact voice amplification. While F_0 is the factor that explains most of the SPL variation, the same is not true for SAL. Although changes were small, musical dynamics were better in explaining the variation observed in SAL. By substituting SAL for SPL on the VRP *y*-axis, a nearly rectangular VRP was obtained and the 11-12 dB per octave slope observed was reduced to almost no slope at all. Figure 5.3b depicts the effect of substituting SAL on the VRP's *y*-axis.

Thirdly, vowel variation was, in practice, negligible in SAL. This met initial expectations. It was found however, that vowel changes also led to little or no SPL variation. Consequently, SAL could not be proposed as a better candidate than SPL on the only premise that its signal was minimally impacted by vowels. Although this outcome was not expected, the fact that vowel variation is negligible in SPL is of interest in the context of VRP recording. It must be noted that for this experiment, only female singers were studied. Therefore, results can only have practical implications for female singing evaluation protocols.

(a) Design 1   (b) Design 2
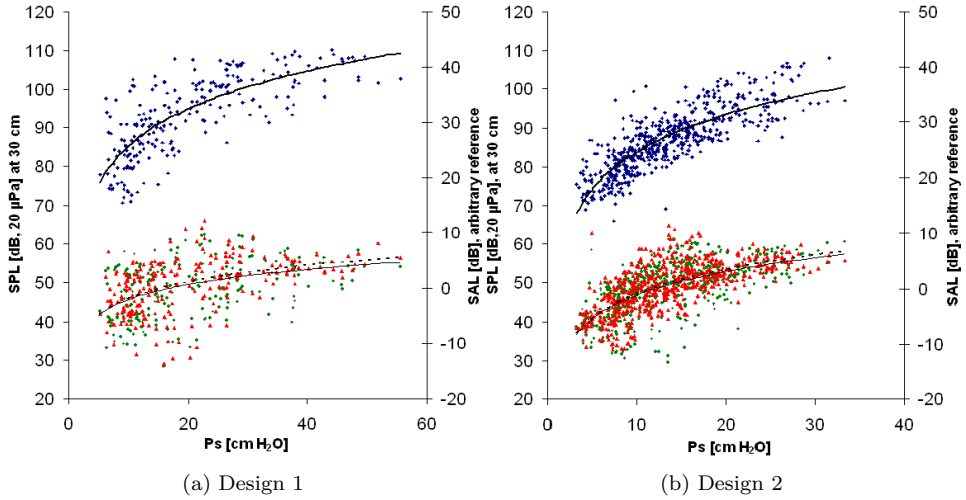
Figure 5.2: Correlations found for SPL and SAL_N (measured at the jogular notch) and SAL_S according to the division of the dataset of Study II into two statistical designs. The regression outcome for a) $Y_{SPL}=14ln(x)+53$, $r^2=0.5968$ and $Y_{SALn}=4ln(x)-11$, $r^2=0.1833$ and for b) $Y_{SPL}=13ln(x)+52$, $r^2=0.6732$ and $Y_{SALn}=6ln(x)-15$, $r^2=0.4171$. SPL is depicted with blue lozengess, SAL_N with green triangles and SAL_S with red lozenges. Both SAL measurements clearly demonstrate a weak correlation to P_s. These two figures appear as Figure 5 and 6 in Paper II.

(a) The Classical VRP with SPL   (b) The Modified VRP with SAL
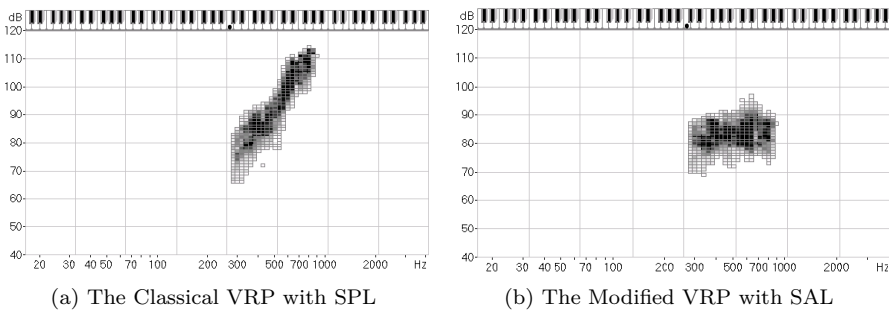
Figure 5.3: The typical VRP slope (a) changes considerably when SAL (measured here at the jugular notch) is substituted for SPL(b) on the VRPs y-axis. These two figures appear as Figure 7 and 8 respectively in Paper II. The outcome illustrated here is for a same subject and task.

Paper III: Not Just Sound: Supplementing the Voice Range Profile with the Singer's Own Perceptions of Vocal Challenges

Aims

The basis for this paper was to use the singer's self-perception of vocal discomfort and/or difficulty to attain further relevant information in the understanding of voice complaints directed specifically to the singing voice. If singing voice problems are often difficult to detect perceptually and even acoustically, perhaps part of the explanation lies in the singing experience. For this purpose, *Phog* was supplemented with a button device which, when pressed, mapped specific frequency and intensity combinations. In this way, non-acoustic but singer-relevant information could be included in the objective vocal measurements of the VRP and possibly fill the gap between external perception and internal experience. Furthermore, such an augmented VRP could succeed in isolating and visually identifying the subtleties of vocal artist problems.

Results

This paper validated a new tool, the button-augmented phonetograph. In order to do this, the consistency of the singer's button pressing was quantified by the amount of overlap found for button presses in different tasks. The reliability of the augmented phonetograph was supported by the consistent button pressing of subjects in task replications as well as across tasks. Similarity scores were on average higher for task replications and lowered somewhat across tasks, yet in both instances statistical proof of non-random behaviour could be demonstrated. Figure 5.4 summarises replication task and across task similarity scores for all subjects. Understandably, in healthy singers, vocal difficulties are of transitory nature. Nevertheless, the button device seemed to be used as a communicative tool during performance and results supported the use of the button-mediated responses as a new metric. In a questionnaire, singers positively graded the efficiency and the information displayed in the button-VRP. As could be expected of healthy singers, the button device was mostly used when vocal limits were visited at the extreme contour portions of the VRP. Figure 5.5 demonstrates the button pressing trend observed in healthy singers. No systematic pressing was found at important voice transitions or register areas of the VRP, a possible consequence of recruiting professional and experienced singers.

Paper IV: Not Just Sound II: an Investigation of Singer Patient Self-Perceptions Mapped into the Voice Range Profile

Aims

In continuation of the previous paper, the augmented phonetograph was used with a singer patient population. The objective was to assess how the button device
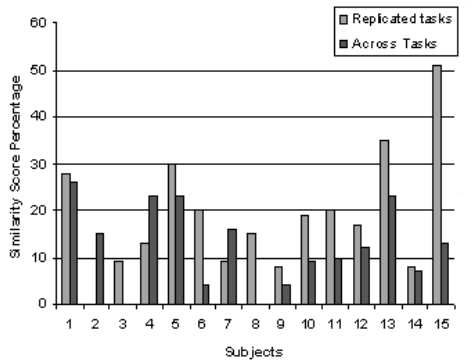
Figure 5.4: Similarity scores obtained for a replicated task and across different tasks. For the majority of subjects, similarity scores lowered when the button pressing behaviour was observed across different tasks.
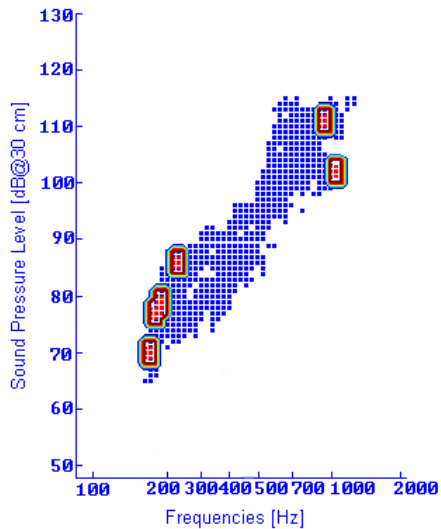
Figure 5.5: A Matlab reconstructed VRP displaying button presses and button regions for a healthy professional soprano. The tendency to press at the periphery of the VRP was a common observation for all 16 singers.

would be used in practice by patients presenting specific singing-voice complaints. While Paper III focused on answering the question, does the button augmented VRP work mechanically and practically? Paper IV sought to answer the questions: how can this type of VRP further assist the clinician in his/her work; and what do button presses tell us about the singer patient?

Results

The semi-structured type of questionnaire collected subjective ratings of overall voice control, impressions in using the button as well as the reasons for doing so. On average, the button press display was rated to be consistent with the recollection of the singing experience. Singer patients also confirmed that the button press map illustrated clearly typical areas of difficulty that they attributed to their pathological vocal state. These areas of button pressing were very divergent from those observed in the case of healthy singers. Button presses in the high frequency and intermediate level portion of the VRP were a recurrent pattern for this group of singer patients. As observed in Figure 5.6, not only were button presses concentrated in one distinct VRP region, but they also occured in inner VRP regions rather than on the periphery.

Answers to open ended questions confirmed that instructions concerning the use of the button device had been understood correctly and yielded interesting support material in understanding the vocal difficulties of singers. The main underlying reasons for pressing the button device were motivated by answers touching on concepts of lack of control, limited dynamic flexibility in the higher range, forcing, larynx height and tension. Surprisingly and opposite to what had been hypothesised, singer-patients had a lower rate of button pressing than healthy singers. Singer-patients were consistent in their use of the button device, although similarity score results were generally weaker than those found for healthy singers.

Paper V: RHI-s

Aims

In view of the frequency with which singers articulate voice complaints related to their singing voice, the Voice Handicap Index (VHI), an instrument which measures the voice handicap of a speaker, seems ill adapted to the reality of the singer patient. The aim of this paper was to create a Swedish version of the VHI to better evaluate the singer's need, language and reality. The VHI's Swedish equivalent is named the Röst Handicap Index (RHI) and so the Swedish version of the instrument adapted to singers was labelled as RHI-s. The work concentrated on verifying the validity, reliability, stability and the overall relevance of the RHI-s. The leading hypothesis was that the RHI-s can successfully evaluate voice handicap in the Swedish singer.
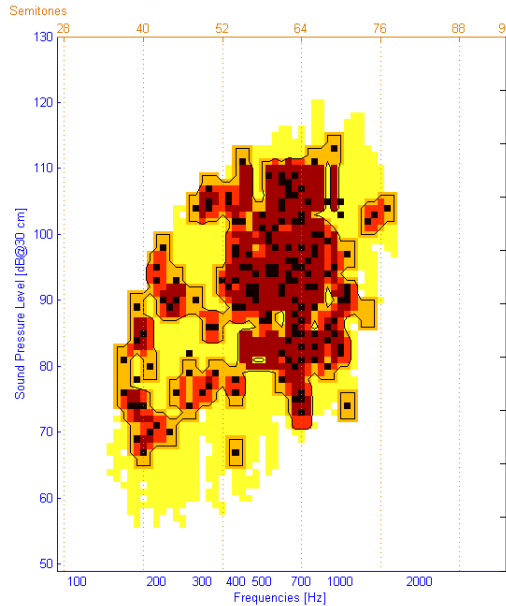
Figure 5.6: Total accumulated button presses for singer patients. The yellow-orange hue identifies areas of single button presses whereas the darker hue gradations underscore the amount of overlap obtained from comparing the initial task to its replication. This figure appears as part of Figure 8 (lower section) in Paper IV.

Results

The Swedish translation and adaptation of the Voice Handicap Index for singers was successful. A total of 96 healthy singers along with a group of 30 singer patients participated in the testing of the new instrument. Robust validity and reliability results were obtained. Singer-patient scores were significantly different from healthy-singer scores both in the test and the retest of the questionnaire (t-values were -10.8 with degrees of freedom 124, $p < 0.001$ and a power of 2.28). Figure 5.7 illustrates the test and retest differences between both groups. Indeed, patient scores were higher than healthy singers (patients had average scores of 54 $\pm$ 18 while healthy singers scored on average 22 $\pm$ 13).

A cut-off score of 31 identified the patient population with 100% sensitivity while the correct identification of healthy singers, the specificity, had an accuracy of 76%. Thus the risk for Type I error in diagnostics was not negligible. Because the RHI-s was not intended as a diagnostic tool, this trade-off between sensitivity and specificity was not a serious one. Unlike many other reports of the VHI, a very high correlation was found between the general self-rated severity of the singing problem

(a) RHI-s Test Scores   (b) RHI-s Retest Scores
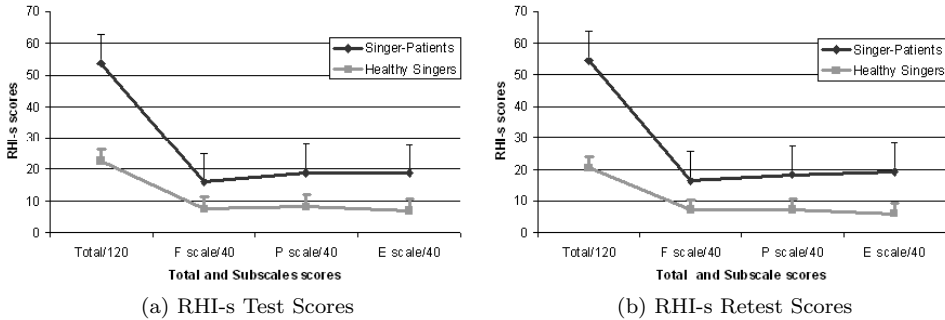
Figure 5.7: Singer-patient and healthy singer RHI-s scores for the test and the retest instances. Total scores were on a scale of 120 points while each subscale had a total of 40 possible points. The error bars depict the positive standard deviations.These figures can be found in Paper V, labelled as Figure 1a-b.

(VAS scaling) and the RHI-s score. The correlation found between the VAS and the RHI-s for the test was 0.74 and for the retest 0.84 ($p < 0.001$ respectively). These results helped establish the strength of the questionnaire's internal coherence. The reliability of the questionnaire was confirmed by high correlations between test and retest scores. When both groups were pooled together, Pearson's r was .91. This value lowered somewhat when groups were analysed separately and the correlation of .85 found for the singer patients was the highest of both groups. Indeed, singer-patient scores differed the least between the times of test and retest. When internal consistency was evaluated, high Cronbach's alpha were obtained for all items as well as for subscales. Despite this last result, a PCA analysis was conducted to verify the adequacy of the subscales. A four-component result indicated that perhaps the RHI-s items would be best explained by four categories or scales. When the four factor scores were analysed by ANOVA, factors 1 and 4 alone could best descriminate between healthy singers and singer patients. Finally, no other variables than sex could be identified as having an effect on RHI-s scoring. Interestingly, the difference between healthy singers and singer patients was greater for females.

Chapter 6

Discussion

6.1 General Discussion

This dissertation work dealt mainly with the adaptation of clinical methods and tools in relation to singing voice demands. The VRP and the VHI both have an extensive history of clinical usefulness and their sensitivity to specific population groups make them ideally suited for inclusion in the evaluation of the singing voice. Many more aspects of the clinical evaluation would no doubt need to be revised and adapted and this work only skims the surface of what needs to be a much bigger endeavour. By first working with descriptors of total vocal output and vocal health, the path to improved clinical measures of singing-voice laryngeal and acoustic function, will hopefully have been better established, as well as the basis for future solid evidence-based work.

As presented in Chapter 3, this work pursued three overall goals. The following discussion serves to relate the findings of each included paper to these goals.

First goal

The first goal was to <u>create singer-related resources for the clinician</u>. All five studies contribute to this aim, in one way or another. Paper I and V deliver well-defined tools with which the clinician can work. Coleman [32, 31] first put forth the idea that a separate VRP recording (a "musical range of phonation") should be made for singers. This was revisited by Awan [3]. Very little, however, has been done to further investigate this VRP recording approach. In Paper I, VRP_{perf} was suggested instead of "musical range of phonation" as it was considered important to combine both, a quality of phonation range and the dynamics used for stage performances. Furthermore, the energy and engagement necessary for performance were also important in the definition of the VRP_{perf}. The latter is not advocated as a replacement for the VRP_{phys} but rather, as its necessary complement. It is important to consider both: the total vocal capabilities of the singer regardless of performance, as well as how he/she uses the voice on stage. For example, a physiological VRP
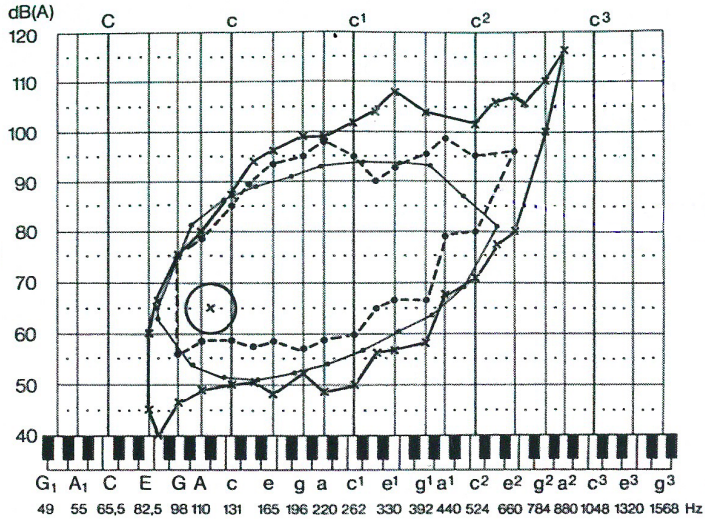
Figure 6.1: A Baritone singer with a singing voice problem: particular dynamic
restrictions between mechanisms 1 and 2. This is an example taken from Schultz-
Coulon which clearly demonstrates the need to do both a VRP_{phys} (the crossed
symbols) and a "musical range of phonation" VRP (the dotted line) recording
in order to properly assess the voice. The encircled cross depicts the habitual
speech frequency in relation to the VRPs while the thin solid line with filled circles
represents a normative reference.

should clearly illustrate the voice transitions between vocal mechanisms ("regis-
ter") whereas such information is skillfully concealed in a "healthy" and proficient
singer's performance and therefore, not readily available in the VRP_{perf}. A similar
example would be the analysis of the lower VRP curve. If singers are instructed
to "perform" a task, and this task is later analysed on a physiological basis, that
analysis is erroneous. Singers simply do not perform at phonation threshold levels.
When the correct task is used (as appreciated in Paper I) a performance VRP, as
a measure of behavioural voice production, becomes clinically relevant. Paper I
demonstrates that in the assessment of singer patients, the measurement of physio-
logical capabilities alone might be insufficient and even misleading in understanding
important and relevant aspects of the disordered singing voice. Schultz-Coulon had
put forth a similar claim in a clinical example of a baritone patient for which both
a VRP_{phys} and a "musical range of phonation" VRP were recorded. Figure 6.1
depicts the possible clinical limitations in conducting the VRP_{phys} alone. Without
the VRP representative of the singer's "musical range of phonation" the dynamical
restrictions in the transition from one laryngeal mechanism to the other would not
have been detected.

Luchsinger & Arnold [98] as well as Large [93] had concluded that, typically, a physiological range should exceed a "musical" range. Coleman [32] concluded rather that if the singer was highly skilled, the physiological and musical ranges would be equivalent. Later, Awan [3] and Sulter [160] both favoured the first mentioned conclusion. Paper I demonstrates that perhaps both conclusions are reasonable. On the one hand, both types of VRPs studied yielded similar minimum and maximum SPL and frequency VRP points, thus indicating that extreme vocal possibilities did not significantly change from one type of VRP to the other. On the other hand, it was found that in a VRP_{perf}, the voice use between these minimum and maximum reference points differed considerably from the voice use in a VRP_{phys}, in terms of both the upper and the lower contours. This difference underscores the importance of giving the correct context to the VRP tasking. The importance of the context is further corroborated by similar conclusions put forth by Emerich *et al.* in the analysis of actor SRPs and VRPs [42]. In Paper I, the context of the performance task incited the singer to sing in a more representative way, something terribly difficult and by definition not advocated, within the limits of a physiological task. In short, if VRP recordings do not include the performance aspect of the voice, then the voice status evaluation of a singer is incomplete. This would also apply to other types of voice measurements.

The capacity of a voice to produce loud sounds was shown to be of particular interest in the case of the Western opera singer. Indeed, the upper curve of the VRP_{perf} can be expected to exceed the one obtained in a VRP_{phys} if female singers are reverting to vocal tract amplification strategies and glottal-source efficiency typical to opera singing. According to the literature, such strategies should produce an acoustic gain of up to 30 dB. Thus, the VRP upper curve becomes especially interesting from both a voice function and an acoustic perspective. In a comparison of female singers and non-singers, Åkerlund obtained a higher VRP upper curve for singers [79]. He concluded that female singers seemed to tolerate and use higher P_s (possibly due to stiffer vocal folds in the higher range) and that more strategic acoustic amplification behaviours were possibly at play. These observations generally closely relate to the ones put forth in Paper I. Åkerlund's instructions to the singers were unfortunately not described in detail. One assumes that the differences he observed would have been even more pronounced if the singer had been singing in a performance context and was not limited to /pae-pae-pae/ phonations while holding a pressure catheter in her mouth.

In a performance context, failure to produce loud sounds could be a considerable handicap and an indication of voice function or technical failure. Other than by Hacki and Åkerlund [59, 79], the necessity for producing loud sounds has not been demonstrated (most likely due to the tested population groups). In Paper I, strict criteria ensured that VRP results would be representative of the professional Western opera female singer. When VRP metrics were tested, differences between the VRP_{phys} and the VRP_{perf} were best explained by four metrics. The SPL range, the Area (proven in the past to be especially discriminating between non-singer and singers), and the two metrics newly introduced in Paper I, the SPL extent and

Percentage of voice area over 90 dB were all significantly different. It follows that the dynamic aspects of the voice are instrumental in assessing a performer. These few metrics can thus help the clinician in assessing the singer's VRP more efficiently and understand where weaknesses affecting performance might lie.

In Paper V, a Swedish version of the VHI adapted for the singer (RHI-s) was created and validated. This instrument thereby becomes a tool which can be directly put to use in the clinic. With the RHI-s, the clinician will be able to finely tune his/her dialogue with the patient and achieve a better understanding of patient priorities. Patient motivation is often a challenge in the clinic, and yet it is a key element to successful rehabilitation. A clinical tool that can address the specific needs of a patient is likely to help increase that patient's motivation. Most importantly, with the RHI-s, the clinician who is not necessarily a singing-voice specialist (SVS) is given some means to work with the singer patient and to better decipher the impact of the complaint. The data collection for this paper also proved to be quite revealing as it exposed certain singer-patient trends that may have great clinical implications. The singer-patient group was mainly composed of soloists of contemporary commercial music (CCM) genres. Similar results are reported by Cohen *et al.,* [29]. Although RHI-s scores were not significantly different between these singers and other singer patients, the fact that voice care help is sought mainly by soloists of jazz, afro; blues; rock; pop and soul clearly identifies the direction for future preventive voice care and highlights the need for further research and improved comprehension of voice source and resonance aspects of the CCM genres. Furthermore, it was by far unexpected and interesting that singing genre, singing level and singing context did not have an effect whatsoever on participant scores. Paper V results do not support assumptions that professional singers experience greater voice handicaps than students or amateur singers. Rather, for singer patients, the general impact of a voice disorder seems to be more or less the same regardless of the singing genre, level and context.

Studies III and IV produced interesting results which can assist and support the clinician's work with the singer. By mapping the singer patient's self-perceptions into the VRP, the clinician is given a tool with which he/she can grasp more directly and distinctly the problem at hand. Since the voice problems are often left unperceived and/or occur very specifically, the visual markings imprinted in the VRP by the patient's button pressing help the clinician to locate and trace areas of concern. The clinical experiment with the button-augmented VRP was well received on the part of the clinician and was found to enhance communication not only between the clinician and the patient, but also between clinicians. The button markers can objectify something that until now remained intangible and subjective.

Second Goal

The second aim of the thesis was the adaptation of common clinical tools to the concern and the perception of the singers. The studies most relevant to this goal were Paper III and IV. In these studies, the VRP was augmented with a button

device which, with instructions, could be used while singing to map points of interest into the VRP. In these studies, singers were asked to press the button to signal instances of particular vocal difficulty and/or discomfort. The VRP was thus adapted to reflect not only maximum voice performance but also an element of subject self-perception. Others have added extra dimensions to the VRP: spectral information such as the energy related to the singer's formant cluster, voice quality aspects such as jitter, shimmer, and hoarseness, as well as voice source information like the open quotient [141, 88, 117, 91, 60]. Yet, the idea of mapping subjective information into an objective map like the VRP goes beyond the voice signal as such, in mirroring both the vocal status and experiences of the singer. The paper demonstrated that the button device instructions were well understood by the singers and that the motor task of button pressing during singing could be performed. Extensive research exists that addresses the combination of motor tasking to speech. The majority of such experiments follow one of three schools of thought: capacity theory, time-sharing models and functional distance theory [39]. They investigate the amount of load incurred by performing motor tasks during speech, on one level or another (e.g., lip movements, brain activity). Generally, such experiments have demonstrated reduced articulatory and semantic abilities during loading. In contrast to these experiments, Paper III and IV showed that singers were fairly consistent in button pressing within task replications and (in the case of healthy singers) across tasks. This result underlines that singing in itself is an act that requires attention to many simultaneous motor details. The act of singing not only combines semantic and musical dimensions, but it also includes rhythm, implicit and procedural memories and physical displacements. Hence, it is not surprising that singers did not demonstrate difficulties in performing the additional button task while singing. Conversely, generalising the button task to non-singers could potentially lead to task performance obstacles unseen in this work.

The lower occurrence of button pressing found in singer patients was unexpected. In view of the additional load that a voice disorder may incur, it may be that the button task becomes more difficult to manage. Indeed, there are psychological aspects related to the button task that need to be considered. In Paper III and IV, an average reaction time of 150 ms was accommodated into the task by asking the singer to phonate a minimum of 2 to 3 seconds per token. Furthermore, each button press was extended to a region. The button region was designed to account for proximity of button presses without actual overlap, and also for vibrato-induced variations. More sophisticated models, addressing the source of error in the use of a button device coupled to the VRP, could further improve the precision of the button-augmented VRP. Yet the button device as it has been tested here is precise enough in marking and mapping the perceptions of singers as they sing. It is rather the precision of the task that is most instrumental, as it will allow to better identify specific sensations and perceptions.

The button-augmented VRP has a practical aspect which might appeal to the singer's reality. Often, singers are practice-oriented people, skilled in demonstration and expression. To some, the analytical act involved in the verbalisation of

vocal problems might be unnatural. By pushing a button, singers could simply demonstrate their vocal difficulties and discomforts. For the singer, this might be of particular interest. This reasoning was corroborated by questionnaire responses. Both groups of singers rated highly the correspondence of VRP button markings to their singing experiences.

Distinctive group patterns of button pressing demonstrated the specificity of vocal difficulties. The singer patients pressed in the interior of the VRP at intermediate SPL and in the higher frequency range. Healthy patients only did so at VRP contour extremes. The pattern observed for singer patients was interesting as it occurred regardless of diagnoses collected in Paper IV. The results obtained for the singer patient group vindicates the importance of considering inner VRP areas rather than VRP contours alone. singer patients pressed predominantly within the 523 to 880 Hz frequency range, yet the reason for this remains unclear. Button presses could be expected to occur in regions related to voice mechanism transitions. Then again, the proficiency and voice classifications within the group were so diversified that group trends related to *passaggi* areas were practically impossible to assess. Trends were however much clearer in relation to SPL; most button presses were found to correspond to the *mf* dynamic segments of the *messa di voce*. In a way, the singer patient's button pattern might be visually depicting what, in the singing world, is commonly referred to as the "hole" in the voice. Singers might compensate successfully to sing at extreme intensities, yet at high frequencies, such compensatory behaviour might interfere with their aptitude to achieve the fine balance between vocal fold mass and subglottal pressure required in a gradual intensity progression towards *mf*. In turn, this difficulty in finding a proper balance might lead to increasing vocal effort. Further investigation of such a phenomenon promises to be of great interest for the singing population.

Paper V also involved some adaptation work. The VHI was remodeled to fit the singer's language and concerns. This psychometric instrument assesses the degree of voice disorder impact on the patient. High scores are related to a severe degree of impact while low scores signify hardly any impact at all (typical of a healthy state). Earlier studies had shown that singer patients scored lower than non-singer patients [128]. This was felt to be an indication of a lack of sensitivity and ability of the VHI to address the singer patient's reality. Indeed, when singers were provided with questions directly related to singing voice use, VHI scores were generally higher [104, 30, 107]. The Swedish version of a VHI adapted for singers (RHI-s) corroborates earlier results and is valuable for the proper assessment of the singer. Since many Swedes are actively engaged in choir singing, this work's ability to reliably appeal to all kinds of singers, especially choristers, was of great importance. Very little effect of the singing context ("sångsammanhang") and the singing levels could be identified in the results of Paper V and thus, the Swedish VHI for singers has succeeded in fulfilling its purpose.

In Paper II, it was shown that the substitution of SAL for SPL on the y-axis of the VRP can facilitate VRP interpretation. The influence of F_0 on the level information displayed is greatly reduced. However, this substitution fails in

transforming the VRP into a voice-source analysis tool. This work was important in that it demonstrated that adaptations of clinical tools to the singing voice cannot simply follow speech models. The use of SAL in speech does not directly extend to singing. The overall results obtained in this paper might at first seem counter-intuitive, in that SAL is influenced mostly by musical dynamics (as opposed to frequency in the case of SPL), but at the same time shows a weaker correlation to P_s. As outlined in Chapter 1, voice intensity (corresponding to musical dynamics) is steered by P_s. Indeed, the correlation between SAL and P_s does exist, but in comparison to the SPL (which is sensitive to both frequency and musical dynamics: P_s driven parameters) -P_s relationship, the correlation is much weaker. Further studies would be needed to investigate the subglottal pressure behaviour in relation to SAL. It is most likely that, due to the dominant first partial of the SAL signal, an increase of subglottal pressure will only result in negligible increases of higher spectrum energy and therefore will not impact the overall SAL signal.

Finally, the work of Paper I tested the necessity to include performance-like exercises in the acquisition of a VRP_{perf}. According to the results, the triad carrier (designed to resemble a typical vocalise) can be recommended in VRP_{perf} recordings of singers. The majority of singers showed some preference for the vocalise approach while the design itself did not yield accountable VRP_{perf} differences when compared to the discrete pitch task. This kind of result was unexpected, especially in view that previous research had demonstrated possible task differences [79]. In comparing a discrete pitch task and a triad task, Åkerlund *et al.* found that female singers could sing at higher levels in the discrete pitch task. In the context of Paper I, the contextual instructions seemed much more influential than the exercise itself. This said, there is a definite distinction between the VRP information obtained for a vocalise or a discrete pitch task and a sung aria excerpt. Paper I demonstrated that the performance type of tasks approximated the sung aria, yet there remained significant differences between the aria and the performance task, showing that such tasks are not fully representative of the voice used on the stage. Perhaps, future developments including stage recordings could elucidate further details concerning the performance aspect of the voice. In the meantime, the choice of a performance task depends on the objective of the investigation or the measurement. For example, in the context of Paper III and IV, a discrete pitch task in which a *messa di voce* could be executed was far more relevant in that the transition between soft and loud voice was believed to be key for the detection of vocal difficulties in singer patients. On the other hand, and in agreement with Åkerlund's results, the nature of the task in VRP_{phys} is very important in determining the outcome.

Third Goal

The third and final goal of this project was the creation and documentation of normative references for the singing population. In a textbook focused on the understanding voice problems, Colton, Casper & Leonard [127], claim that the lack of definition of a healthy speaking voice limits the setting of therapeutic goals

Table 6.1: Menstrual Cycle Information

| | Menstrual Cycle | | | |
	Soprano	Mezzo-soprano	Contralto	Group
Menses	4			4
Follicular	3	4	2	9
Ovulation	4	1	1	6
Luteal	5	3	2	10
Pregnancy			1	1
Menopause				

and the understanding of vocal deviations from the healthy state as well as their degree of severity. They criticise the lack of quantifiable and objective data. One understands that if this is the case for speech, it must be even more so the case for the singing voice. With appropriate norms against which to compare performance, a researcher or clinician might use total vocal function output results, such as the VRP, in diagnosing, assessing and the monitoring the voice. Data provided in Paper I are a first step towards such normative data of the singing voice. In order for this kind of data to be useful, subject group criteria need to be strict. Only professional female Western opera singers were included in this paper. It is suggested that, due to the VRP's sensitivity to individual characteristics like age, gender and training, the VRP is also capable of discriminating between levels of training/profession as well as the genre of singing (according to Frank [48], Seidner claimed that the VRP alone was not capable of doing so. This is most likely the case if only the VRP_{phys} is considered). Differences between genres of singing have been demonstrated not only on the acoustic level but also on the voice-source level [12, 168, 164, 27, 28]. The VRP is greatly influenced by these two vocal aspects and therefore, it can be expected that a VRP of another type of singer would not and should not be comparable to an opera singer's VRP. Since the VRP can be quite sensitive to age, the large age span of the subject group in Paper I (20 to 55) could have an impact on VRP results. For this reason, subjects were also asked to indicate their current menstrual cycle or menopause information. This data was not originally included in the article publication of this paper but was important in deciphering which of the effects, age or classification of voice, was more pertinent for the group's VRP analysis. Table 6.1 gives the group's menstrual cycle profile.

Although eight subjects were 40 years and older, there were no reported menopausal cases. Voice category changes in late career could also indicate a possible aging effect of the voice. To this effect, a questionnaire item addressed voice category changes. Subjects in Paper I only reported changes in relation to early training paths. The possible impact of age effect was thus discarded in the analysis of the VRP data and the possible impact of voice category was considered more pertinent for this group of professional singers.

The subject selection was very rigorous and subjects had to meet several criteria:

non-smoking, a minimum of 5 years of vocal training, free of voice complaint and an unproblematic vocal health history. Although these stringent group criteria might seem unimportant, they find support in the literature. For example, Roubeau *et al.,* in an explicit study of groups of non-singers, amateur and professional singers, concluded that clear subject group definitions were necessary. For example, amateur singers demonstrated an intermediate vocal behaviour to non-singers and professional singers [131] and so, mixing group definitions to include amateur singers and accomplished singers might jeopardize the conclusions of a study.

The recording procedure in Paper I was also very important. Simple details, such as the stance of the subject, were made to be as stage-like as possible. Each subject was asked to stand and to visualise themselves as if on stage. Frank and Donner emphasised this subtle but important difference in recording VRPs of the singing voice [48]. Paper I recordings included vibrato. As earlier shown in Figure 4.1a, vibrato can considerably impact the end results of a VRP recording. Coleman also attested to this but at the same time, stated that a "musical range profile" should factor in vibrato [31]. Awan did not include vibrato in his "musical" VRPs [3] and it is unclear if vibrato has been included in earlier studies of the singing voice and the VRP reported in the German literature.

Paper V also contributes in creating and documenting a normative reference for the singing voice. 96 healthy singers as well as 30 singer patients participated in RHI-s tests. The collected scores can help give some degree of expectations as what a typically healthy score should be. By means of ROC analysis, a cut-off value of 31 was deemed to successfully differentiate between singers with and without voice complaints. In the same line, a gauge of clinically significant change is reported by determining the critical limits of the test-retest mean differences. A change of total RHI-s score of more than 16 could be attributed to a voice status change. Similarly, a variation of more than 6 or 7 points on subscales scores could help track more precisely the nature of the voice-status change.

6.2 Limitations

General

For all of the experiments, the subjects were scarce, especially due to the fact that they should be representative of an elite or a specific population. Naturally, normative VRPs of singers should include many more voices and results obtained here need further confirmation. The same is true for the singer-patient tests, both in the case of Paper IV and Paper V. Paper V gave fair enough results given that the singing population of Sweden, although important, remains rather small when compared to more populated countries. Yet, a more effective comparison would require a patient group comparable in number to the control group. Moreover, in an ideal comparison, each control would be matched to a patient (taking into account at least variables such as age, sex, genre and level). In Paper IV, the overall subject criteria had to be relaxed and even abandoned. A more even distribution

in gender, singing genre, level and diagnosis, for example, would have allowed for a deeper assessment of the button pressing in relation to diagnosis.

Paper I

In this paper, the goal was to look at the differences between a VRP_{phys} and a VRP_{perf} , the idea being that a singer's evaluation should include a stage-relevant vocal performance. However, the recording context imposed certain limits as to how representative the performance could really be. Singers are not likely to find themselves in a very dampened acoustic environment where body movements and gestures are restricted as much as possible Still, singers are often at the mercy of vocally unfit scenography and need to comply with various singing positions (lying down, still, moving) as well as restrictions such as costumes and pre-established interpretations. In this light, the recording conditions of the experiment were not deemed inhabilitating nor less conducive to performance. Yet, the environmental (acoustic as well as physical) and the behavioural context to the voice use deserve some attention. Differences that were registered in the framework of this experiment are telling and could be more pronounced if a more realistic setting had been used. Much interest lies in studying the impact of different acoustic environments on the singer's voice use. Furthermore, experiments including virtual acoustic environments and even music accompaniment (that could be subtracted in a later processing stage) would be of great value in assessing the true difference between the reality of the stage and the studio. The study of voice use in its typical context has increased in value in the last decade (voice dosimetry being a main example) and certainly this holds for both the speaking and the singing voice.

Another issue concerns the quality of the VRP_{phys} data included in this paper. First, the inclusion of such a recording procedure was added somewhat later in the experiment, hence reducing the number of recordings per voice classification group. Only two mezzo-soprano recordings were obtained and therefore comparisons to the other voice categories were limited. Furthermore, a pitch glide task was used to record the VRP_{phys}. This task was chosen for its more or less rapid elicitation. The choice was also purposefully made following pilot recordings of two singers. In a sustained tone context, it was observed that singers had more difficulties to disregard voice quality and refrain from "singing" the tone. In fact, most singers had to be heavily coached for the physiological task. If the singer began a pitch glide without vibrato and voice quality, she was more likely to maintain that type of phonation for the entire glide than in the sustained tone task. These choices were certainly motivated but they also led to a much higher lower VRP curve than what is normally found in the literature. On that basis, the interpretation of phonation threshold pressure information derived from the lower VRP curve was not possible. However, this result was interesting as it indicated the critical importance of the task design in a physiological setting.

The task effect mentioned above could possibly have been avoided if subjects had had access to visual feedback. The VRP has been praised for its capacity to

provide immediate visual feedback [108][69]. In this setting, visual feedback was not made available for the main reason that button pressing information was collected in parallel for Paper III. It was important that visual feedback would not interfere with the subject's button pressing. In hindsight, providing visual feedback to the subject could probably be sufficient to compensate for the task effect of the pitch glide and help produce similar results as that obtained in other studies. Then again, most comparable studies do not report the use of visual feedback. The use of visual feedback is mostly reported in conjunction to investigations of therapy and voice status differences [154, 38].

Paper II

In this paper, the correlation between SAL and P_s was investigated. P_s measurements, as mentioned in the Methodology Section 4, are often difficult to collect, especially when the subject's attempt to sing the [pae] strings. Although the data collection was carefully monitored, the tasks involved in this experiment could have been structured so as to collect additional P_s tokens to yield a wider data range for the analysis.

Paper III and IV

The button-augmented VRP is a proof-of-concept idea which would require sophisticated and detailed models of motor, evaluation and judgement reaction time as well as vibrato to enable the precise analysis of both the proprioceptive and acoustic information behind the button pressing. However, the button augmented VRP in its present form succeeds in locating areas that deserve further consideration in the analysis of the singer patient's voice. The qualitative appeal of this kind of information can lead to quite interesting clinical applications, some of which can be found in the Discussion section of the respective paper. For these experiments, the question of task training posed some particular difficulties for the singer-patient population. The VRP recording was already considered to be extensive and thus training would have necessitated too much voice use. Both button pressing occurrences and group consistent behaviour were lowered in the case of the singer-patient group and this might be partially due to the lack of training for this group. Perhaps multiple recording sessions would be best suited for further experiments. It is suggested that the acquisition protocol entailing a training session followed by a rest period and a later experimental session would be better suited for a singer patient. The challenge then falls in the recruitment domain: voluntary participation might decrease in view of the considerable recording time required.

Paper V

The work performed with the Swedish adaption of the VHI for singers (RHI-s) was successful and the newly validated instrument will give a practical resource

to Swedish clinicians who work with singers. However, one might wonder if the VHI as a whole is based on the right kind of structure. Likert scaling is a closed-ended approach and psychological research has demonstrated that such structures, in the case of threatening questions, result in lower scoring and poorer overall results (for a singer, voice handicap related questions might indeed feel threatening). Furthermore, research shows that social desirability factors are higher in answering closed-ended questions. It is important to note that the RHI-s is well suited to assess the overall total health profile of an individual but is by far insufficient as a stand-alone assessment. This said, the VHI and the adaptation of the VHI to singers seem to adequately capture the essential impact that a voice disorder can have on an individual.

6.3   Future Work and Possible Applications

Ideas of future work and possible applications have been already touched upon, either in the main discussion of the thesis or in the respective discussion of the included papers. Here follow some questions and observations that are borne out of this dissertation work.

- By establishing clear group criteria and precisely exposing methodological procedures, a singing-voice database (including VRP and other relevant measures) could be developed. Subjects, especially elite level singers, are a continuous challenge to recruit, and controlled environment recordings are precious to research investigations. Such a database, as is suggested here, could form a wealth of research as well as clinical and education resources.

- Paper I, gathering data on the singer's VRP (physiological and performance) leads to many possible future steps. First and foremost, it would be essential and most interesting to compare the normative data obtained here to matching singer-patient data. Due to the level of proficiency of the singer group in Paper I, it would be important to compare VRPs of a similarly proficient group. Here, the taxonomy and level of usage schemes that have been previously elaborated can truly assist this kind of endeavour. Furthermore, it is suggested that group criteria should be strict in terms of singing genre. It would be most interesting to collect normative VRPs by singing genre and compare them to each other. The results obtained here are interpreted according to the voice technique employed for opera singing and it would be informative to assess the impact of other genres on the VRP_{perf}.
  Another interest lies in the systematic and experimental investigation of the minimum SPL and the SPLext produced in the high range of the voice. To do this, one would have to acount in the VRP for the singer's choice of register/voice mechanism. A recording of separate voice mechanisms such as performed initially by Wolf et al. [182] and by Stout [158] and again revisited by Roubeau et al. [131] and Lamesch et al.[91] could be most informative.

Such mechanism-defined VRPs could yield more immediate information on the dynamic flexibility of the singer, the effect of training against the natural F_0/SPL interaction, and the necessary pressure applied to the vocal folds in the high. Also, it could possibly be an important feature to compare across voice categories where range is somewhat the same, but dynamic flexibility in the higher portion of the voice is not. It could be useful to highlight the most recurring overall VRP shapes and form normative series accordingly.

- Paper III and IV, investigating the possibility of supplementing the VRP with the singer's perception, were explorative studies that also lead to many possible new avenues of investigation. As mentioned earlier, it would be of great research interest to finely tune the button-augmented VRP in order to render possible the acoustic analysis of the area highlighted by the button marker. The principle underlying the button VRP was that vocal difficulties of singers, which are tied to vocal effort, are not necessarily perceived nor detected in the acoustic signal. The button markings in the VRP might be able to guide further the analysis of the singer patient's voice and uncover unnoticed, yet perhaps key details, in the audio signal. The button-VRP seems ideally suited for pre-post voice-therapy monitoring and could be a promising asset to future evidence-based studies. It would be interesting to include such a tool in the long-term rehabilitation process of singer patients. The loading issues above named could be more specifically identified and the subject could become a more active participant in the overall rehabilitation process. The issue of diverging perceptions is a very interesting one in that this divergence might be interfering in the rehabilitative process. The button-augmented VRP may be an ideal tool in working towards understanding and bridging perceptual differences, and thus, improve the definition of a common goal for the clinician and the patient. Tests could be developed in which both the clinician and the patient are requested to use the button device to independently mark the VRP according to identical instructions. With different colour mapping the divergence in perception could be mapped and this information could become pertinent to vocal progress and therapy efficiency. Aside these few suggestions, the button-augmented VRP can be seen as a promising assistance to diagnostic procedures where problematic frequencies and intensity combinations are mapped out to facilitate the laryngeal examination. Some pegagogical aspects were also discussed in Paper IV. When using the button-augmented VRP in voice lesson contexts, an enhanced learning might be promoted, due to the terminal retrospective biofeedback involved in the button pressing. The performance is uninterrupted, yet the markers are in place to allow educative discussion and analysis. Not only would the student learn from this exercise, but the teacher would also gain some knowledge as to the student's perceptions of difficulties and challenges. Similarly as what has been mentioned above, a large part of successful voice training is the training and shaping of the student's or the patient's perception. The

button-augmented VRP could be of some assistance in reaching this aim.

- Paper II, exploring the SAL as an estimate of SPL, yielded interesting results concerning the subglottal pressure behaviour in connection to skin acceleration levels. It would be of research interest to investigate in further detail what occurs in the SAL spectrum as subglottal pressure increases.

- Paper V, testing a Swedish translation of the Voice Handicap Singers for singers, implied test-retest of the adapted VHI for singers. This adaptation is validated and useful but it could be interesting to consider the score differences for the same individual between the original standard and the adapted test. To the author's knowledge, some clinicians do this informally in their communications with other voice professionals and the difference between the scores can help further identify and define the patient's voice complaint.

6.4   Main Conclusions

- It is of importance to consider the stage/performance facet of the voice in the voice status assessment of the singer.

- The vocal proficiency of a studied group impacts considerably the VRP results obtained. The level of training as well as the genre of singing are key group criteria in producing normative VRP singer data .

- In creating VRP_{perf} norms for singers, there are no task design effects on overall results. In the work performed for this thesis, the pertinence of grouping subjects/patients according to voice classification was not shown. However, this variable should be considered in studies of larger singing populations.

- There are aspects of the singing voice that are not necessarily clearly identified in an acoustic signal but that become possible to study when the singer's self-perception is mapped to the VRP (physiological or performance).

- When tests are adapted to the reality and the needs of the singer (like the RHI-s), scores and responses become more representative of the voice complaint.

- Singing is different from speech and therefore voice evaluation equipment and evaluation task instructions should be adapted consequently.

- In singing, P_s is more strongly related to SPL than it is to SAL. More research is needed to understand the SAL-P_s relationship.

Bibliography

[1]   R Airainer and F Klingholz. Quantitative evaluation of phonetograms in the case of functional dysphonia. *J Voice*, 7(2)(2):136–141, 1993. (Cited on pages 25, 36, 37, and 48).

[2]   SN Awan. *Phonetographic assessment of voice characteristics in youn adults iwth and without vocal training*. PhD thesis, Kent State University, 1989. (Cited on pages 23 and 36).

[3]   SN Awan. Phonetographic profiles and f0-spl characteristics of untrained versus trained vocal groups. *J Voice*, 5:41–50, 1991. (Cited on pages 14, 45, 45, 47, 47, 47, 47, 48, 52, 77, 79, and 85).

[4]   SN Awan. Superimposition of speaking voice characteristics and phonetograms in untrained and trained vocal groups. *J Voice*, 7(1):30–37, 1993. (Cited on pages 25, 45, 45, 47, 47, 47, 47, 47, 47, 47, 47, and 47).

[5]   SN Awan. The aging female voice: acoustic and respiratory data. *Clin Linguist Phon*, 20(2-3):171–180, 2006. (Cited on page 25).

[6]   RJ Baken and RF Orlikoff. *Clinical Measurement of Speech and Voice, 2nd Edition*. Thomson Delmar Learning, 2000. (Cited on pages 5, 5, 8, 17, and 28).

[7]   A Behrman, CJ Agresti, E Blumstein, and G Sharma. Meaningful features of voice range profiles from patients with organic vocal fold pathology: A preliminary study. *Journal of Voice*, 10(3):269–283, 1996. (Cited on pages 25, 47, 47, 47, 47, 47, 47, and 48).

[8]   M Belhau and J Oates. Best practice guidelines for the management of clients with voice problems. In *Proceedings of the 27th World Congress of the International Association Logopedics and Phoniatrics*, 2007. (Cited on pages 5, 6, and 10).

[9]   M Benninger, B Jacobson, and A Johnson. *Vocal Arts Medicine: The Care and Prevention of Professional Voice Disorders, 3rd edition*. Thieme Medical Publishers, Inc, 1994. (Cited on pages 4 and 66).

[10] K Berglund and AE Hallin. Fonetogram för röstfriska män; normaldata, jämförelse med granulompatienter samt metoddiskussion. Master's thesis, Karolinska Institutet, Institute for clinical science and technology, May 2007. (Cited on pages 23, 34, 45, 47, 47, 47, 47, 47, 47, 47, 47, and 48).

[11] G Böhme and G Stuchlik. Voice profiles and standard voice profile of untrained children. *J Voice*, 9(3):304–307, 1995. (Cited on pages 25, 47, and 47).

[12] E Björkner. Musical theater and opera singing–why so different? a study of subglottal pressure, voice source, and formant frequency characteristics. *J Voice*, 22(5):533–540, Sep 2008. (Cited on page 84).

[13] DM Bless and RJ Baken. International association of logopedics and phoniatrics (ialp) voice committee discussion of assessment topics. *J Voice*, 6(2):194–210, 1992. (Cited on pages 26, 47, and 48).

[14] G Bloothooft. A computer-controlled device for voice-profile registration. In *Proceedings IX Conference of the Union of European Phoniatricians*, 1981. (Cited on pages 29, 32, 47, 48, and 50).

[15] G Bloothooft and D Binnenpoorte. Towards a searcheable resource of phonetograms. In *Proceedings Voicedata 98*, 1998. (Cited on page 48).

[16] G Bloothooft, M Dorssen, K van Koppen, P Pabon, and M van Wijck. On the comparison of computerised phonetograms. In *Proceedings Int. Conf. Phonetic Sciences*, 1999. (Cited on pages 33, 47, and 48).

[17] G Bloothooft and J P H Pabon. Vocal registers revisited. 1999. (Cited on page 48, 48, and 48).

[18] G Bloothooft and J P H Pabon. *Qualities of Voice Emeritus in On Speech and Language: Studies for Sieb G. Nooteboom*, pages 17–26. 2. LOT Occasional, 2004. (Cited on page 48).

[19] WS Brown, RJ Morris, DM Hicks, and E Howell. Phonational profiles of female professional singers and nonsingers. *J Voice*, 7(3):219–226, 1993. (Cited on page 34).

[20] M Büttner, W Seidner, and P Eichorst. Der "Klangkoeffizient" - ein beachtenswerter Parameter bei der Messung von Singstimmprofilen. *Sprache-Stimme-Gehör*, 15:135–138, 1991. (Cited on pages 19, 47, and 51).

[21] EH Buder and T Wolf. Instrumental and perceptual evaluations of two related singers. *J Voice*, 17(2):228–244, 2003. (Cited on page 47 and 47).

[22] M Bunch and J Chapman. Taxonomy of singers used as subjects in scientific research. *J Voice*, 14(3):363–369, 2000. (Cited on page 10).

[23] J Calvet and G Malhiac. Courbes vocales et mue de la voix. *J Fr Oto-Rhino-Laryngol*, 1:115–124, 1952. (Cited on pages 24, 47, 47, 47, 47, 48, and 49).

[24] LM Carroll, RT Sataloff, RJ Heuer, JR Spiegel, SL Radinoff, and JR Cohn. Respiratory and glottal efficiency measures in normal classically trained singers. *J Voice*, 10(2):139–45, 1996. (Cited on page 4 and 4).

[25] SH Chen. Voice range profile of taiwanese normal young adults: a preliminary study. *Zhonghua Yi Xue Za Zhi (Taipei)*, 58(6):414–420, 1996. (Cited on page 25).

[26] TF Cleveland. A clearer view of singing voice production: 25 years of progress. *J Voice*, 8(1):18–23, 1994. (Cited on page 4).

[27] TF Cleveland, R E Stone, J Sundberg, and J. Iwarsson. Estimated subglottal pressure in six professional country singers. *J Voice*, 11(4):403–409, Dec 1997. (Cited on page 84).

[28] TF Cleveland, J Sundberg, and R E Stone. Long-term-average spectrum characteristics of country singers during speaking and singing. *J Voice*, 15(1):54–60, 2001. (Cited on page 84).

[29] SM Cohen, BH Jacobson, CG Garrett, JP Noordzij, MG Stewart, A Attia, RH Ossoff, and TF Cleveland. Creation and validation of the singing voice handicap index. *Ann Otol Rhinol Laryngol*, 116 (6):420–6, 2007. (Cited on page 80).

[30] SM Cohen, JP Noordzij, CG Garrett, and RH Ossoff. Factors associated with perception of singing voice handicap. *Otolaryngol Head Neck Surg*, 138(4):430–4, 2008. (Cited on page 82).

[31] RF Coleman. Sources of variation in phonetograms. *J Voice*, 7(1):1–14, 1993. (Cited on pages 25, 27, 33, 34, 41, 77, and 85).

[32] RF Coleman, J H Mabis, and J K Hinson. Fundamental frequency-sound pressure level profiles of adult male and female voices. *J Speech Hear Res*, 20(2):197–204, 1977. (Cited on pages 20, 24, 25, 35, 47, 47, 47, 47, 77, and 79).

[33] RF Coleman and JB Mott. Fundamental frequency and sound pressure level profiles of young female singers. *Folia Phoniatr*, 30:94–102, 1978. (Cited on pages 16, 47, 47, 47, 47, and 49).

[34] RH Colton, JK Casper, and M Hirano. *Understanding Voice Problems: A Physiological Perspective for Diagnosis and Treatment*. Lippincott Williams & Wilkins, 1996. (Cited on page 5).

[35] RH Colton and H Hollien. Phonational range in the modal and falsetto registers. *J Speech Hear Res*, 15(4):708–713, 1972. (Cited on page 14).

[36] H Daffern. *Distinguishing Characteristics of Vocal Techniques in the Specialist Performance of Early Music*. PhD thesis, Department of Music and Electronics, University of York, 2008. (Cited on page 11).

[37] PH Damsté. The phonetogram. *Pract Otorhinolaryngol (Basel)*, 32(3):185–187, 1970. (Cited on page 47, 47, 47, and 47).

[38] PH DeJonckere, I van Wijck, and R Speyer. Efficacy of voice therapy assessed with the voice range profile (phonetogram). *Rev Laryngol Otol Rhinol (Bord)*, 124(5):285–289, 2003. (Cited on pages 25 and 87).

[39] C Dromey and E Shim. The effects of divided attention on speech motor, verbal fluency and manual task performance. *J Speech Lang Hear Res*, 51:1171–1182, 2008. (Cited on page 81).

[40] HW Eichel. The voice field index, a practicable method for the quantitative evaluation of voice field data. *Sprache-Stimme-Gehör*, 12:63–64, 1988. (Cited on pages 37, 47, 47, 47, and 48).

[41] ME Elias, RT Sataloff, DC Rosen, RJ Heuer, and JR Spiegel. Normal strobovideolaryngoscopy: variability in healthy singers. *J Voice*, 11(1):104–7, 1997. (Cited on page 5).

[42] KA Emerich, IR Titze, JG Svec, PS Popolo, and G Logan. Vocal range and intensity in actors: a studio versus stage comparison. *J Voice*, 19(1):78–83, 2005. (Cited on pages 23, 45, 47, 47, 47, 47, and 79).

[43] C Ericsdotter. *Articulatory-Acoustic Relationships in Swedish Vowel Sounds*. PhD thesis, Stockholm University, 2005. (Cited on page 19).

[44] G Fant. *Acoustic Theory of Speech Production*. The Hague: Mouton & Co, 1960. (Cited on page 18).

[45] G Fant. Preliminaries to analysis of the human voice source. *STL-QPSR*, Report 4:1–27, 1982. (Cited on page 17).

[46] G Fant and Q Lin. Glottal source - vocal tract acoustic interaction. *STL-QPSR*, 28(1):13–27, 1987. (Cited on page 18).

[47] NH Fletcher and TD Rossing. *The Physics of Musical Instruments*. Springer-Verlag Berlin, Heidelberg, New York, 2005. (Cited on page 19).

[48] F Frank and F Donner. Die Bedeutung der Stimmfeldmessungen für den Gesangsunterricht aus phoniatrischer und gesangspädagogischer Sicht. *Sprache-Stimme-Gehör*, 10:93–97, 1986. (Cited on pages 45, 47, 47, 47, 84, and 85).

[49] FJ Gamboa, A del Palacio, A Mico, A Nieto, and I Cobeta. Fonetograma en lesiones laríngeas por abuso vocal. *Acta Otorrinolaringol Esp*, 45(1):31–6, 1994. (Cited on pages 47, 47, 47, 47, and 48).

[50] M Garnier, N Henrich, J Smith, and J Wolfe. Soprano's secondo passaggio: a resonance adjustment and or a laryngeal transition to the whistle voice. *J Acoust Soc Am*, 123 (5):3243, 2008. (Cited on page 20).

[51] P Gramming. *The Phonetogram - an Experimental and Cllinical Study*. PhD thesis, Malmö University, Sweden, 1988. (Cited on pages 23, 26, 35, 47, 47, 47, 47, 47, 47, 48, 48, 48, and 48).

[52] P Gramming and L Akerlund. Non-organic dysphonia. II. Phonetograms for normal and pathological voices. *Acta Otolaryngol*, 106(5-6):468–476, 1988. (Cited on pages 25, 26, and 68).

[53] P Gramming and J Sundberg. Spectrum factors relevant to phonetogram measurement. *J Acoust Soc Am*, 83(6):2352–2360, 1988. (Cited on pages 18, 26, and 68).

[54] P Gramming, J Sundberg, S Ternström, R Leanderson, and W H Perkins. Relationship between changes in voice pitch and loudness. *JVoice*, 2(2):118–126, 1988. (Cited on page 51).

[55] M Gross. Stimmfeldmessung, eine objektive Methode der Stimmdiagnostik. *Sprache Stimme Gehör*, 4:100–101, 1980. (Cited on pages 28 and 32).

[56] T Hacki. Die Beurteilung der quantitativen Sprechstimmleistungen: Das Sprechstimmfeld im Singstimmfeld. *Folia Phoniatr*, 40:190–196, 1988. (Cited on pages 32 and 45).

[57] T Hacki. Comparative speaking, shouting and singing voice range profile measurement: physiological and pathological aspects. *Logoped Phoniatr and Vocol*, 21:123–129, 1996. (Cited on page 47 and 47).

[58] T Hacki. Tonhöhen- und Intensitätsbefunde bei Stimmgeübten: Vergleichende Sprechstimmfeld-, Rufstimmfeld- und Singstimmfeldmessung. *HNO*, 47(9):809–815, 1999. (Cited on page 68).

[59] T Hacki, B Frittrung, Ch Zywietz, and Ch Zupan. Verfahren zur statistischen Ermittlung van Stimmfeldgrenzen - Das Durchschnittstimfeld. *Sprache-Stimme-Gehör*, 14:110–112, 1990. (Cited on pages 39, 51, and 79).

[60] M Heinemann and H Gabriel. Möglichkeiten und Grenzen der Stimmfeldmessung: Vorstellung des Heiserkeitsfeldes als ergänzende Methode. *Sprache-Stimme-Gehör*, 6:37–43, 1982. (Cited on pages 48 and 81).

[61]   N Henrich, C d'Alessandro abd B Doval, and M Castellengo. Glottal open quotient in singing: measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency. *J Acoust Soc Am*, 117(3):1417–30, 2005. (Cited on page 18).

[62]   L Heylen. *De Klinische relevantie van het fonetogram*. PhD thesis, University of Antwerpen, 1997. (Cited on pages 23, 25, 34, 47, 47, 47, 47, 47, 47, 48, and 48).

[63]   L Heylen, F L Wuyts, F Mertens, M DeBodt, and P H Van De Heyning. Normative voice range profiles of male and female professional voice users. *J Voice*, 16(1):1–7, 2002. (Cited on pages 25, 40, and 48).

[64]   L Heylen, F L Wuyts, F Mertens, M DeBodt, J Pattyn, C Croux, and P H Van De Heyning. Evaluation of the vocal performance of children using a voice range profile index. *J Speech Lang Hear Res*, 41(2):232–238, 1998. (Cited on page 25).

[65]   L Heylen, F L Wuyts, F Mertens, M DeBodt, J Pattyn, and P H Van De Heyning. Comparison of the results of the frequency and the intensity data of the bsgvd with phonetogram characteristics. *Acta Otorhinolaryngol Belg*, 50(4):353–360, 1996. (Cited on pages 25, 28, and 34).

[66]   LG Heylen, FL Wuyts, FW Mertens, and JE Pattyn. Phonetography in voice diagnoses. *Acta Otorhinolaryngol Belg*, 50(4):299–308, 1996. (Cited on pages 25 and 41).

[67]   M Hirano, S Saito, M Sawashima, S Hiki, and H Hirose. Guidline for vocal function tests. *Jpn J Logop Phoniatr*, 23(2):164–7, 1982. (Cited on page 25).

[68]   EB Holmberg, RE Hillman, and JS Perkell. Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *J Acoust Soc Am.*, 84(2):511–529, 1988. (Cited on page 18 and 18).

[69]   EB Holmberg, E Ihre, and M Södersten. Phonetograms as a tool in the voice clinic: Changes across voice therapy for patients with vocal fatigue. *Logoped Phoniatr Vocol*, 32(3):113–27, 2007. (Cited on pages 23, 25, 47, 47, 48, 48, and 87).

[70]   EJ Hunter, JG Svec, and IR Titze. Comparison of the produced and perceived voice range profiles in untrained and trained classical singers. *J Voice*, 20(4):513–526, 2006. (Cited on pages 19, 23, 25, 47, 47, 54, and 68).

[71]   EJ Hunter, IR Titze, and F Alipour. A three-dimensional model of vocal fold abduction/adduction. *J Acoust Soc Am.*, 115(4):1747–1759, 2004. (Cited on page 13).

[72] Y Ikeda, T Masuda, H Manako, H Yamashita, T Yamamoto, and S Komiyama. Quantitative evaluation of the voice range profile in patients with voice disorder. *Eur Arch Otorhinolaryngol*, 256 Suppl 1:S51–S55, 1999. (Cited on pages 25 and 48).

[73] N Isshiki. Regulatory mechanism of voice intensity variation. *J Speech Hear Res*, 128:17–29, 1964. (Cited on page 20).

[74] BH Jacobson, A Johnson, C Grywalski, A Silbergleit, G Jacobson, MS Benninger, and CW Newman. The voice handicap index (vhi) development and validation. *Am J Speech Lang Pathol*, 6:66–70, 1997. (Cited on page 61).

[75] E Joliveau, J Smith, and J Wolfe. Vocal tract resonances in singing: The soprano voice. *J Acoust Soc Am*, 116 (4) Pt. 1:2434–2439, 2004. (Cited on pages 19, 20, and 20).

[76] RD Kent. *The MIT Encyclopedia of Communication Disorders*. MIT Press, 2004. (Cited on page 14).

[77] L Åkerlund. Average loudness level, mean fundamental frequency, and subglottal pressure: comparison between female singers and nonsingers. *J Voice*, 8(3):263–70, 1994. (Cited on page 52).

[78] L Åkerlund. Averages of sound pressure levels and mean fundamental frequencies of speech in relation to phonetograms: comparison of nonorganic dysphonia patients before and after therapy. *Acta Otolaryngol (Stockh)*, 113(1):102–8, 1994. (Cited on page 25).

[79] L Åkerlund. *Clinical and Experimental Aspects of Phonetogram Analysis*. PhD thesis, University Hopsital, Malmö Department of Oto-Rhino-Laryngology, 1996. (Cited on pages 23, 36, 47, 47, 47, 47, 47, 48, 79, 79, and 83).

[80] L Åkerlund, P Gramming, and J Sundberg. Phonetogram and averages of sound pressure levels and fundamental frequencies of speech: Comparison between female singers and nonsingers. *J Voice*, 6(1):55–63, 1992. (Cited on page 52).

[81] K Kitajima and F Fujita. Estimation of subglottal pressure with intraoral pressure. *Acta Otolaryngol (Stockh)*, 109:473:478, 1990. (Cited on page 58).

[82] F Klingholz. Die Stimme des Sangers im Phonetogramm. *Laryngorhinootologie*, 68(1):62–66, 1989. (Cited on page 26).

[83] F Klingholz. Computer aided evaluation of phonetograms. *Computer Methods Programs Biomed*, 37(2):127–35, 1992. (Cited on pages 5, 33, and 36).

[84] F Klingholz and F Martin. Quantitative evaluation of the voice field. *Sprache-Stimme-Gehör*, 7:106–110, 1983. (Cited on pages 25, 32, 36, 37, 47, 47, 47, 48, and 50).

[85] F Klingholz, F Martin, and A Jolk. Die Bestimmung der Registerbrüche aus
      dem Stimmfeld. *Sprache-Stimme-Gehör*, 9:109–111, 1985.   (Cited on page
      48).

[86] S Komiyama. Phonogram-a new method for evaluating voice characteristics.
      Otologia (Fukuoka), 18:428–440, 1972.   (Cited on pages 25 and 48).

[87] S Komiyama, H Watanabe, and S Ryu. Phonographic relationship between
      pitch and intensity of the human voice. *Folia Phoniatr (Basel)*, 36(1):1–7,
      1984.   (Cited on pages 25 and 47).

[88] U Konzelmann, M Moser, and G Kittel.   Stimmfeldmessungen bei
      Chorsängern vor und nach Stimmbelastung unter besonderer Berücksichti-
      gung des Sängerformanten. *Sprache-Stimme-Gehör*, 13:112–118, 1989. (Cited
      on pages 47, 47, 47, 48, 51, and 81).

[89] MN Kotby, MY Abul-Ela, and AA Orabi. Voice range profile as a quantitative
      measure of vocal function: Normative data. In MN Kotby, editor, *Congress
      Proceedings of XXIII IALP Congress*, pages 46–49, 1995.   (Cited on pages
      33, 47, 47, 48, and 48).

[90] J Koufman.   Levels of vocal usage.   http://www.voiceinstituteny.com/
      Voice/ab/voice-disorders2.htm, 2008-08-06.   (Cited on page 10).

[91] S Lamesch, B Doval, and M Castellengo. Phonetograms of laryngeal source
      parameters for different vowels and laryngeal mechanisms. *J Acoust Soc Am*,
      123(5):3243, 2008.   (Cited on pages 25, 54, 68, 81, and 88).

[92] S Lamesch, R Expert, M Castellengo, N Henrich, and B Chuberre. Investi-
      gating voix mixte: A scientific challenge towards a renewed vocal pedagogy.
      In M. Marin & J. Ross (Eds.) K. Maimets-Volt, R. Parncutt, editor, *Proceed-
      ings of the third Conference on Interdisciplinary Musicology (CIM07)*, 2007.
      (Cited on page 54).

[93] J Large. Acoustic study of register equalization in singing. *Folia Phoniatr
      (Basel)*, 24:39–61, 1973.   (Cited on page 79).

[94] WD LeBorgne and BD Weinrich. Phonetogram changes for trained singers
      over a nine-month period of vocal training. *J Voice*, 16(1):37–43, 2002. (Cited
      on pages 4, 16, 23, 25, 36, 45, 47, 47, 47, 53, and 68).

[95] T Leino, A-M Laukkanen, I Ilomäki, and E Mäki. Assessment of vocal capac-
      ity of finnish university students. *Folia Phoniatr*, 60:199–209, 2008.   (Cited
      on pages 25, 47, 47, 47, 47, 47, 47, and 47).

[96] A Löfqvist, B Carlborg B, and P Kitzing. Initial validation of an indirect
      measure of subglottal pressure during vowels. *J Acoust Soc Am*, 72:633–635,
      1982.   (Cited on page 58).

[97] TM Loucks, D Duff, JH Wong, and R Finley-Detweiler. The vocal athlete and endotracheal intubation: a management protocol. *J Voice*, 12(3):349–59, September 1998. (Cited on page 4).

[98] R Luchsinger. Physiologie der Stimme. *Folia Phoniatr (Basel)*, 5(2):58–127, 1953. (Cited on page 79).

[99] E Ma, J Robertson, C Radford, S Vagne, R El-Halabi, and E Yiu. Reliability of speaking and maximum voice range measures in screening for dysphonia. *J Voice*, 21(4):397–406, 2007. (Cited on pages 23, 45, 47, 47, 47, 47, 47, 47, and 48).

[100] A McAllister. Acoustic, perceptual and physiological studies of ten- year-old children's voices. *TMH-QPSR*, 38(1):75–103, 1997. (Cited on page 47 and 47).

[101] A McAllister, E Sederholm, J Sundberg, and P Gramming. Relations between voice range profiles and physiological and perceptual voice characteristics in ten-year-old children. *J Voice*, 8(3):230–239, 1994. (Cited on page 25).

[102] MR Van Mersbergen, K Verdolini, and I R Titze. Time-of-day effects on voice range profile performance in young, vocally untrained adult females. *J Voice*, 13(4):518–528, 1999. (Cited on pages 23, 33, 47, 47, and 48).

[103] J Montojo, G Garmendia, and I Cobeta. Comparación entre los resultados del fonetograma manual y el fonetograma automático. *Acta Otorrinolaringol Esp*, 57(7)(7):313–18, 2006. (Cited on page 33).

[104] D Morsomme, M Gaspar, J Jamart, M Remacle, and I Verduyckt. Adaptation du voice handicap index à la voix chantée. *Rev Laryngol Otol Rhinol*, 128(5):305–314, 2007. (Cited on page 82).

[105] D Morsomme, C H Simon, J Jamart, M Remacle, and I Verduyckt. Proposition d'adaptation du voice handicap index à la voix chantée. *Rev Laryngol Otol Rhinol (Bord)*, 126(5):305–313, 2005. (Cited on page 66).

[106] D Mürbe, J Sundberg, J Iwarsson, F Pabst, and G Hofman. Longitudinal study of solo singer education effects on maximum spl and level in the singers' formant range. *Logoped Phonaitr. Vocol*, 24(4):178–186, 1999. (Cited on pages 16, 23, 47, 47, 47, 47, and 53).

[107] T Murry, A Zschommler, and J Prokop. Voice handicap in singers. *J Voice*, [Epub ahead of print], 2008. (Cited on page 82).

[108] C Neuschaefer-Rube. Comparison of voice performance in auditory and visually assisted phonetography. *Logoped Phonaitr. Vocol*, 22(1):3–8, 1997. (Cited on page 87).

[109] CH Neuschaefer-Rube, F Sram, and S Klajman. Three-dimensional phoneto-graphic assessment of voice performance in professional and non-professional speakers. *Folia Phoniatr*, 49:96–104, 1997. (Cited on pages 25, 47, 47, 47, and 48).

[110] AC Ohlsson, B Järvholm, A Löfqvist, PE Näslund, and R Stenborg. Vocal behaviour in welders-a preliminary study. *Folia Phoniatr (Basel)*, 39(2):98–103, 1987. (Cited on pages 47 and 48).

[111] JPH Pabon. Objective acoustic voice-quality parameters in the computer phonetogram. *J Voice*, 5(3):203–216, 1991. (Cited on pages 33, 48, and 48).

[112] JPH Pabon. The phonetogram in singing voice analysis and synthesis. In *Proceedings of the Stockholm Music Acousic Conference (SMAC)*, volume 79, pages 228–33. Royal Swedish Academy of Music, 1993. (Cited on pages 33 and 48).

[113] JPH Pabon. Dynamics and spectral slope. Talk given during the course of Pevoc4, 2001. (Cited on page 48, 48, and 48).

[114] JPH Pabon, A Lamarche, and S Ternström. Fourier descriptor analysis and unification of voice range profile contours: method and applications. *Journal of Speech, Language and Hearing Research*, 2009. (in review). (Cited on pages 12, 32, 40, 40, and 42).

[115] JPH Pabon, A McAllister, E Sederholm, P Gramming, and J Sundberg. Dynamics and voice quality aspects of children's voices as manifested in com-puter phonetograms. 1988. (Cited on pages 40 and 48).

[116] JPH Pabon, A McAllister, E Sederholm, and J Sundberg. Dynamics and voice quality information in the computer phonetograms of children's voices. presented at the Child Voice Symposium, Stockholm Sept. 1998. (Cited on page 33).

[117] P Pabon and R Plomp. Automatic phonetogram recording supplemented with acoustical voice-quality parameters. *J Speech Hear Res*, 31(4):710–22, 1988. (Cited on pages 32, 34, 48, and 81).

[118] MF Pedersen. Computed phonetograms in adult patients with benign voice disorders before and after treatment with a nonsedating antihistamine (lo-ratadine). *Folia Phoniatr (Basel)*, 43(2):60–67, 1991. (Cited on page 33).

[119] MF Pedersen. A longitudinal pilot study on phonetograms/voice profiles in pre-pubertal choir boys. *Clin Otolaryngol Allied Sci*, 18(6):488–491, Dec 1993. (Cited on pages 25, 52, and 68).

[120] MF Pedersen, B Frokjaer, JF Pabst, HK Schutte, T Hacki, and HL Hansen. Standardization voice range profile measurements (phonetography / stimm-feldmessung). In *Proceedings of the XXIIIrd Congress International Association Logopedics Phoniatrics*, pages 41–5., 1995. (Cited on page 27).

[121] MF Pedersen and H Lindskov-Hansen. Computerized phonetograms for clinical use. In *Proc XXth congress International Association Logopedics and Phoniatrics (Folia phoniatrica)*, pages 170–1, 1986. (Cited on page 32).

[122] MF Pedersen, H Lindskov-Hansen, T Lindskov Hansen, and E Munk. A phonetograph for use in clinical praxis. *Acta otolaryngol*, S412:138, 1984. (Cited on pages 32 and 50).

[123] MF Pedersen, S Moller, and P Krabbe. Voice categories compared with phonetograms, androgen,estrogens and puberty stages in 8-19 year old choir girls. *J Research Singing*, 13:14, 1990. (Cited on page 51).

[124] PS Popolo, JG Svec, K Rogge-Miller, and IR Titze. Technical considerations in the design of a wearable voice dosimeter. Poster presented at the First Pan-American / Iberian Meeting on Acoustics / 144th Meeting of the Acoustical Society of America, Cancun, Mexico, December 2-6 2002. (Cited on page 58).

[125] SL Radionoff. Preparing the singing voice specialist revisited. *J Voice*, Volume 18 (4):513–521, 2004. (Cited on page 4 and 4).

[126] A Rauhut, E Stürzebecher, H Wagner, and W Seidner. Messung des Stimm-feldes. *Folia Phoniatr*, 31:119–124, 1979. (Cited on pages 28 and 45).

[127] R Leonard RH Colton, JK Casper. *Understanding Voice Problems: A Physiological Perspective for Diagnosis and Treatment*. Lippincott Williams & Wilkins, 3rd edition edition, 2006. (Cited on page 83).

[128] CA Rosen and T Murry. Voice handicap index in singers. *J Voice*, 14(3):370–7, 2000. (Cited on pages 66 and 82).

[129] M Rothenberg. A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. *J Acoust Soc Am*, 53(6):1632–1645, 1973. (Cited on page 58).

[130] M Rothenberg. Così' Fan Tutte and What it Means or Nonlinear Source-Tract Acoustic Interaction in the Soprano Voice and Some Implications for the Definition of Vocal Efficiency. In T Baer, C Sasaki, and KS Harris, editors, *Vocal Fold Physiology: Laryngeal Function in Phonation and Respiration*, pages 254–263. College Hill Press, San Diego, 1986. (Cited on page 18).

[131] B Roubeau, M Castellengo, P Bodin, and M Ragot. Phonétogramme par registre laryngé. *Folia Phoniatr*, 56(5):321–33, 2004. (Cited on pages 25, 53, 68, 85, and 88).

[132] B Roubeau, C Chevrie-Muller, and C Arabia. Control of laryngeal vibration in register change. In J Gauffin and B Hammarberg, editors, *Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms*, pages 279–286. Singular, San Diego, 1991. (Cited on page x).

[133] RT Sataloff. *Professional Voice: The Science and Art of Clinical Care, 2nd Edition*. Plural Publishing Inc, 1997. (Cited on page 3).

[134] RT Sataloff. Evaluation of professional singers. *Otolaryngol Clin North Am*, 33(5):923–56, 2000. (Cited on pages 4 and 66).

[135] HJ Schultz-Coulon. *Stimmfeldmessung*. Springer-Verlag Berlin, Heidelberg, New York, 1990. (Cited on page 47, 47, 47, and 47).

[136] HJ Schultz-Coulon and S Asche. Das "Normstimmfeld" ein Vorschlag. *Sprache-Stimme-Gehör*, 12:5–8, 1988. (Cited on pages 36, 45, 45, 45, 47, 47, 47, 47, 47, and 47).

[137] HK Schutte. *The Efficiency of Voice Production*. PhD thesis, Groningen University, 1980. (Cited on pages 17, 23, and 25).

[138] HK Schutte and W Seidner. Recommendation by the Union of European Phoniatricians (UEP): standardizing voice area measurement/phonetography. *Folia Phoniatr (Basel)*, 35(6):286–288, 1983. (Cited on pages 25 and 43).

[139] M Södersten, S Ternström, and M Bohman. Loud speech in realistic environmental noise: phonetogram data, perceptual voice quality, subjective ratings, and gender differences in healthy speakers. *J Voice*, 19(1):29–46, 2005. (Cited on page 25).

[140] W Seidner. Objektive qualitätsbeurteilung der stimme mittels dynamikmessungen. *Z Klin Medicina (Kaunas)*, 40:1521–1525, 1985. (Cited on page 45).

[141] W Seidner, H Krüger, and KD Wernecke. Numerische Auswertung spektraler Stimmfelder. *Sprache-Stimme-Gehör*, 9:10–13, 1985. (Cited on pages 16, 19, 45, 45, 47, 47, 47, 47, 47, 50, and 81).

[142] W Seidner and J Wendler. *Die Sängerstimme:Phoniatrische Grundlagen des Gesangs*. Wilhelmshafen, Heinrichshofen, 1997. (Cited on page 47).

[143] M Sihvo, P Laippala, and E Sala. A study of repeated measures of softest and loudest phonations. *J Voice*, 14(2):161–169, 2000. (Cited on page 25).

[144] M Sihvo and E Sala. Sound level variation findings for pianissimo and for-
     tissimo phonations in repeated measurements. *J Voice*, 10(3):262–268, 1996.
     (Cited on pages 25 and 47).

[145] N Siupsinskiene. Quantitative analysis of professionally trained versus un-
     trained voices. *Medicina (Kaunas)*, 39(1):36–46, 2003. (Cited on pages 25,
     47, 47, 47, 47, 47, 47, 47, 47, 47, 48, and 48).

[146] P Sjölander and J Sundberg. Spectrum effects of subglottal pressure variation
     in professional baritone singers. *J Acoust Soc Am*, 115:1270–3, 2004. (Cited
     on page 17).

[147] NP Solomon and M Dimattia. Effects of a vocally fatiguing task and sys-
     temic hydration on phonation threshold pressure. *J Voice*, 14(3):341–62,
     2000. (Cited on page 15).

[148] NP Solomon, LE Glaze, RR Arnold RR, and M van Mersbergen. Effects of
     a vocally fatiguing task and systemic hydration on men's voices. *J Voice*,
     17(1):31–46, 2003. (Cited on page 15).

[149] NP Solomon, P Ramanathan, and MJ Makashay. Phonation threshold pres-
     sure across the pitch range: preliminary test of a model. *J Voice*, 21(5):541–
     50, 2007. (Cited on page 15).

[150] A Sonninen. The role of the external laryngeal muscles in length-adjustment
     of the vocal cords in singing. *Acta Otolaryngol*, 130 (suppl.):1–102, 1956.
     (Cited on page 13).

[151] A Sonninen, P Hurme, and R Toivonen. Computer voice fields in the study
     of normal phonation. *Acta PHoniatrica Latina*, 10 (3):345–356, 1988. (Cited
     on pages 29, 33, and 34).

[152] A Sonninen, P Hurme, and E Vilkman. Computer voice fields of connected
     speech (6th edition). In P Hurme, editor, *Papers in Speech Research*, pages
     93–111, 1985. (Cited on pages 26, 29, 32, 45, and 48).

[153] A Sonninen, E Vilkman, P Hurme, and R Toivonen. Computer voice fields
     in basic phonation research: Rotation vs. gliding in cricothyroid articula-
     tion. *Scandinavian Journal of Logopedics and Phoniatrics*, 12(1):20–28, 1987.
     (Cited on page 32).

[154] R Speyer, G H Wieneke, I vanWijck Warnaar, and PH Dejonckere. Effects
     of voice therapy on the voice range profiles of dysphonic patients. *J Voice*,
     17(4):544–556, 2003. (Cited on pages 25, 34, 47, 47, 47, 47, 47, 47, 47, 47,
     48, 48, 48, 48, and 87).

[155] AJ Stecher. *Die quantitative Bewertung trainierter und untrainierter menschlicher Stimmen im Stimmfeld*. PhD thesis, Universität München, 1983. (Cited on pages 23, 47, 47, 47, 47, and 48).

[156] JC Stemple, LE Glaze, and BK Gerdeman. *Clinical Voice Pathology: Theory and Management*. Cengage Learning, 2000. (Cited on page 5).

[157] RE Stone and PA Ferch. Intra-subject variability in f0-splmin voice profiles. *J Speech Hear Disord*, 47(2):134–137, 1982. (Cited on page 47).

[158] B Stout. The harmonic structure of vowels in singing in relation to pitch and intensity. *J Acoust Soc Am*, 10:137–146, 1938. (Cited on pages 24, 47, 47, 48, 49, and 88).

[159] AM Sulter. *Variation of voice quality features and effects of voice training in males and females*. PhD thesis, Rijksuniversiteit te Groningen, 1996. (Cited on pages 23, 23, and 39).

[160] AM Sulter, HK Schutte, and DG Miller. Differences in phonetogram features between male and female subjects with and without vocal training. *J Voice*, 9(4):363–377, 1995. (Cited on pages 14, 25, 25, 34, 36, 47, 47, 47, 47, 47, 47, 47, 48, 48, 48, 48, 68, 68, and 79).

[161] AM Sulter, HP Wit, HK Schutte, and DG Miller. A structured approach to voice range profile (phonetogram) analysis. *J Speech Hear Res*, 37(5):1076–1085, 1994. (Cited on pages 33, 36, and 53).

[162] J Sundberg. Chest wall vibrations in singers. *J Speech Hear Res*, 26:329–40, 1983. (Cited on pages 15 and 60).

[163] J Sundberg. *The Science of Singing*. Dekalb: Northern Illinois University Press, 1987. (Cited on pages 16, 18, 19, and 20).

[164] J Sundberg, TF Cleveland, Jr RE Stone, and J Iwarsson. Voice source characteristics in six premier country singers. *J Voice*, 13(2):168–83, 1999. (Cited on page 84).

[165] J Sundberg and J Gauffin. Waveform and spectrum of the glottal voice source. In B Lindblom and S Öhman, editors, *Frontiers of speech communication research, Festschrift for Gunnar Fant*, chapter Waveform and spectrum of the glottal voice source, pages 301–322. Academic Press: London, 1979. (Cited on page 18).

[166] JG Švec, IR Titze, and PS Popolo. Estimation of sound pressure levels of voiced speech from skin vibration of the neck. *J Acoust Soc Am*, 117(3):1386–1394, 2005. (Cited on page 67).

[167] LC Teles-Magalhaes, MI Pegoraro-Krook, and R Pegoraro. Study of the elderly females' voice by phonetography. *J Voice*, 14(3):310–321, 2000. (Cited on pages 25, 47, 47, 47, 47, and 48).

[168] M Thalén and J Sundberg. Describing different styles of singing: A comparison of a female singer's voice source in "classical", "pop", "jazz" and "blues". *Logoped Phonaitr. Vocol*, 26(2):82–93, 2001. (Cited on page 84).

[169] IR Titze. The physics of small-amplitude oscillation of the vocal folds. *J Acoust Soc Am*, 83(4):1536–1552, 1988. (Cited on page xi).

[170] IR Titze. Acoustic interpretation of the voice range profile (phonetogram). *J Speech Hear Res*, 35(1):21–34, 1992. (Cited on pages 25, 26, 47, 47, 47, and 48).

[171] IR Titze. *Principles of Voice Production*. Prentice Hall Inc, 1994. (Cited on pages 14, 18, and 19).

[172] IR Titze. Toward standards in acoustic analysis of voice. *J Voice*, 8(1):1–7, 1994. (Cited on pages 16, 19, 19, 20, 27, and 34).

[173] IR Titze. A theoretical study of F0-F1 interaction with application to resonant speaking and singing voice. *J Voice*, 18(3):292–298, 2004. (Cited on page 18).

[174] IR Titze and J Sundberg. Vocal intensity in speakers and singers. *J Acoust Soc Am*, 91:2936–2946, 1992. (Cited on page 17).

[175] IR Titze, D Wong, MA Milder, SR Hensley, and LO Ramig. Comparison between clinician-assisted and fully automated procedures for obtaining a voice range profile. *J Speech Hear Res*, 38(3):526–535, Jun 1995. (Cited on page 33).

[176] K Verdolini, Y Min, IR Titze, J Lemke, K Brown, M van Mersbergen, J Jiang, and K Fisher. Biological mechanisms underlying voice changes due to dehydration. *J Speech Lang Hear Res*, 45(2):268–281, 2002. (Cited on page 15).

[177] K Verdolini, IR Titze, and A Fennell. Dependence of phonatory effort on hydration level. *J Speech Hear Res*, 37(5):1001–1007, 1994. (Cited on page 15).

[178] E Vilkman, A Sonninen, and P Hurme. Observations of voice productions by means of computer voice profiles. In *Proc. XXth congress International association of Logopedics and Phoniatrics*, pages 370–71, 1986. (Cited on page 32).

[179] GT Vogelsanger. Experimentelle Prüfung der Stimmleistung beim Singen. *Folia Phoniatr (Basel)*, 6(4):193–227, 1954. (Cited on pages 24, 47, and 47).

[180] JM Wingate, WS Brown, R Shrivastav, P Davenport, and CM Sapienza. Treatment outcomes for professional voice users. *J Voice*, 21(4):433–49, 2007. (Cited on pages 25, 47, 47, 47, 47, 47, 47, and 48).

[181] SK Wolf and W J Sette. Some applications of modern acoustic apparatus. *J Acoust Soc Am*, 6:160–168, 1935. (Cited on page 24).

[182] SK Wolf, D Stanley, and WJ Sette. Quantitative studies on the singing voice. *J Acoust Soc Am*, 6:255–266, 1935. (Cited on pages 24, 24, 45, 47, 47, 47, 49, and 88).

[183] FL Wuyts, MS De Bodt, G Molenberghs, M Remacle, L Heylen, B Millet, K Van Lierde, J Raes, and PH Van De Heyning. The dysphonia severity index an objective measure of vocal quality based on a multiparameter approach. *J Speech Lang Hear Res*, 43:796–809, 2000. (Cited on page ix).

[184] FL Wuyts, L Heylen, F Mertens, M Du Caju, R Rooman, PH Van De Heyning, and M DeBodt. Effects of age, sex, and disorder on voice range profile characteristics of 230 children. *Ann Otol Rhinol Laryngol*, 112(6):540–548, 2003. (Cited on page 68).

[185] EM Yiu and RM Chan. Effect of hydration and vocal rest on the vocal fatigue in amateur karaoke singers. *J Voice*, 17(2):216–227, 2003. (Cited on pages 23 and 25).

Part II

Included Papers

Paper I

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 aquisition? 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 the fundamental frequency (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 et al., 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 [2,3-10] 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 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 m^3, 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, Saven Hitech, Sweden). 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 [19]. 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 and1b).

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 [20]. 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, [21]

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, 22]. This expectation could even possibly extend to the SRP [23]. 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 [15,21,25]. 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 et al.,'s reports of a logistic regression, the enclosed area metric was best at differentiating female untrained voices from female singers [1, 22].

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 (Percent$_{\geq 90dB}$).

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, 22]; 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: Rge, f_{min}, f_{max}, SPL_{rge}, SPL_{min}, SPL_{max}, SPL_{ext}, Area, $Percent_{\geq 90dB}$. 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, Chicago Il.

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 3.1 to 2.8 octaves (38 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 31% 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 the speech fundamental frequency (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 illustrate the results for the contour averaging of the physiological and performance tasks.

Tables 6 a) and 6 b) are 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 Pilais'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 /hormones	2	2	3	7
C-over the counter /homeopathic	3	2		5
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 denotefrequency 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.
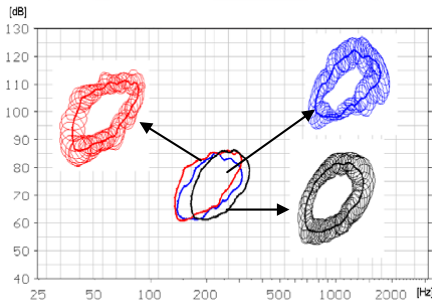

Figure 1 Average SRP contours for the counting speech task 1b (N= 30, soprano in black (16), mezzosopranos in red (8) and contralto in blue (6). The insets show the two-dimensional standard deviation as ellipses, whose orientation also suggests the local covariation of F_0 with SPL.
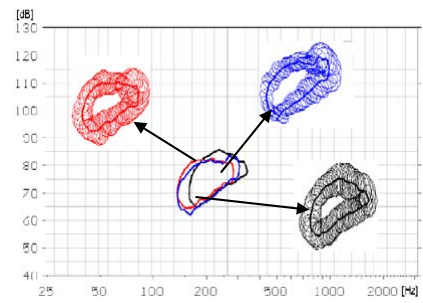

Figure 2 Average SRP contours for the spontaneous speech task 1a (N= 30, soprano in black (N=16), mezzosopranos in red (8) and contralto in blue (6). Insets show the standard deviations as for Figure 1.

	Task	Category							
		Soprano		Mezzosoprano		Contralto		Total	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
f_{max}	Fysio	1315.2	223.86	1176.66	96.01	1186.5	244,45	1249.65	220.26
	Pitch	1295.13	115.89	1147.25	181.50	982.23	187,41	1141.54	39.70
	Vocalise	1173.02	79.70	1061.70	195.46	988.00	154,12	1074.24	58.66
	Excerpt	970.61	128.42	942.07	150.36	800.97	183,40	904.55	27.67
f_{min}	Fysio	151.52	11.72	113.65	13.89	128.58	8,97	138.19	17.77
	Pitch	174.46	19.71	131.71	16.74	160.37	22,19	155.51	2.73
	Vocalise	169.72	24.23	144.93	35.39	158.20	21,24	157.61	7.46
	Excerpt	258.22	36.90	241.41	41.46	215.05	26,61	238.23	7.61
Rge	Fysio	3.14	0.27	3.41	0.06	3.22	0,38	3.2	0.31
	Pitch	2.92	0.20	3.14	0.25	2.64	0,49	2.90	0.15
	Vocalise	2.83	0.25	2.91	0.42	2.67	0,34	2.80	0.08
	Excerpt	1.93	0.18	1.99	0.26	1.90	0,27	1.94	0.05
SPL_{max}	Fysio	112.75	8.00	115.50	2.12	113.17	7,33	113.69	6.65
	Pitch	114.00	3.58	112.38	6.09	114.60	5,13	113.66	1.27
	Vocalise	113.50	2.92	114.00	5.61	108.83	2,32	112.11	1.75
	Excerpt	111.50	4.27	110.50	5.98	108.67	5,61	110.22	0.90
SPL_{min}	Fysio	56.38	3.46	55.50	0.71	58.00	2,16	51.13	2.31
	Pitch	63.75	3.75	64.13	5.03	60.00	4,47	62.63	0.64
	Vocalise	66.38	4.49	65.00	4.11	65.83	4,12	65.74	0.22
	Excerpt	71.44	6.25	74.00	3.63	67.50	6,09	70.98	1.47
SPL_{rge}	Fysio	56.38	6.82	60.00	1.41	55.17	8,48	62.56	6.67
	Pitch	50.25	3.47	48.25	8.99	54.60	7,09	51.03	2.80
	Vocalise	47.13	3.83	49.00	8.75	43.00	3,35	46.38	2.99
	Excerpt	40.06	5.28	36.50	5.15	41.17	4,96	39.24	0.17
SPL_{ext}	Fysio	26.85	4.79	29.07	4.41	27.47	3,36	27.38	4.62
	Pitch	17.34	2.63	15.94	3.83	19.64	3,31	17.64	0.60
	Vocalise	16.16	4.00	18.24	6.11	13.69	2,31	16.03	1.90
	Excerpt	14.24	2.75	12.79	3.58	13.86	4,14	13.63	0.70
Area	Fysio	762.75	275.58	725.00	4.24	734.50	63,36	744.44	155.04
	Pitch	528.06	104.94	534.63	137.60	522.00	131,30	528.23	17.33
	Vocalise	541.94	144.60	642.50	215.72	478.33	102,24	554.26	57.34
	Excerpt	338.44	78.20	312.50	76.90	324.17	79,68	325.03	1.39
Percent $\geq 90dB$	Fysio	30.66	9.57	28.13	1.20	30.87	4,50	30.68	8.57
	Pitch	46.59	7.32	41.50	8.99	39.66	9,28	42.58	1.07
	Vocalise	48.74	13.32	41.43	8.66	37.95	11,86	42.70	3.51
	Excerpt	50.38	15.25	53.99	12.95	46.50	14,51	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.
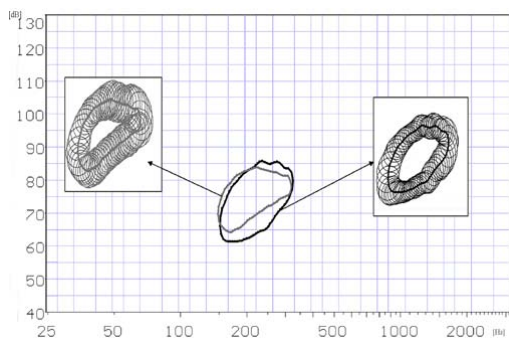
Figure 3 Average SRP contours for two types of speech, counting (1b) and spontaneous (1a), compared. Counting is in black and spontaneous speech is in grey. (N=30) Insets show standard deviations as for Figure 1.
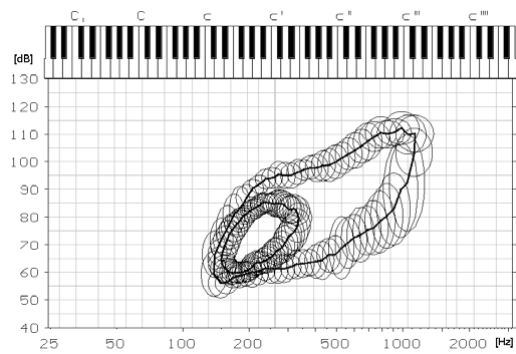

Figure 4a. Average SRP contours for speech (counting, inner contour) and for the physiological task (outer contour). Covariation ellipses are included. The lower minimal curves align closely, as do the lower portions of the upper contours. (counting, N=30, physiological, N=16).
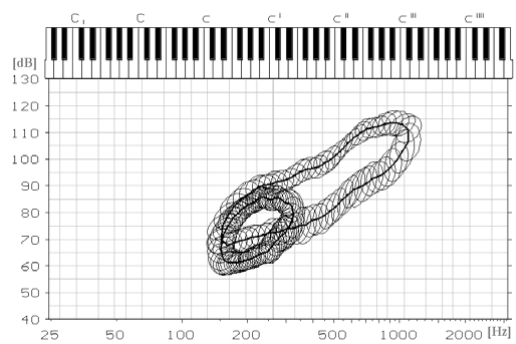

Figure 4b. Average SRP contours for speech (1b) and for the sung performance task , including covariation ellipses (N=30).
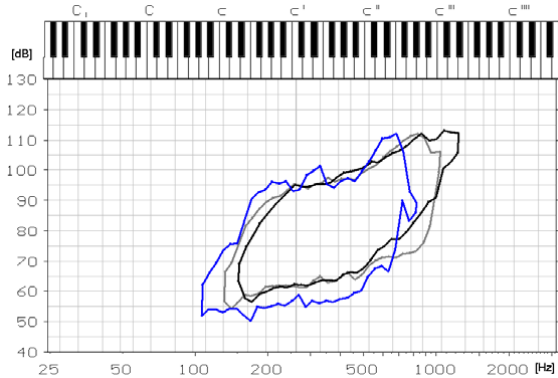
Figure 5 Average VRP contours for the physiological task (Task 2); for sopranos in black (8), for mezzosopranos in blue (2) and for contraltos in light grey (6), superimposed for comparison.
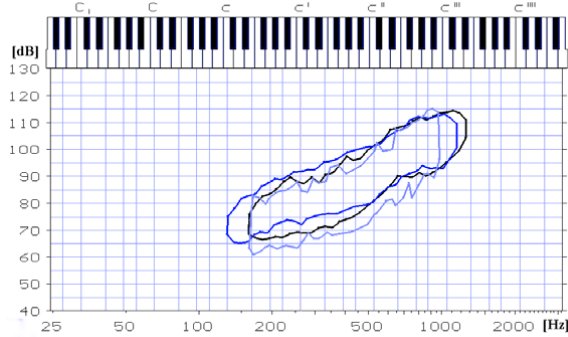

Figure 6 a) Average performance VRP contours for the discrete pitch task (Task 3). Sopranos are in black (16), mezzosopranos in blue (8) and contraltos in light grey (6).
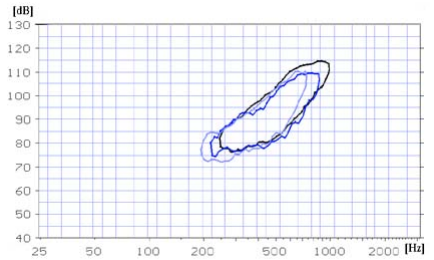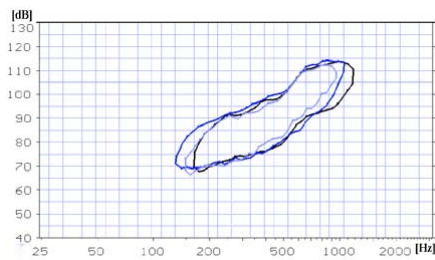

Figure 6 b) Average performance VRP contours for the vocalise task (Task 4) and **c)** Average performance VRP contours for the aria excerpt (Task 5). Sopranos are in black (16), mezzosopranos in blue (8) and contraltos are in grey (6).

	Z	Asymp. Sig. (2-tailed)
$f_{max2} - f_{max}$	-0.835	0.404
$f_{min2} - f_{min}$	-1.711	0.087
$rge_2 - rge$	-1.511	0.131
$SFF_2 - SFF$	-2.714	**0.007**
$SPL_{max2} - SPL_{max}$	-0.364	0.716
$SPL_{min2} - SPL_{min}$	-3.275	**0.001**
$SPL_{rge2} - SPL_{rge}$	-2.518	0.012
$SPL_{avg2} - SPL_{avg}$	-1.330	0.184
$SPL_{ext2} - SPL_{ext}$	-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 SPLmax (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	1
REGWR Test	Contralto	989.75		165.78		2.60
	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 $Percent_{\geq 90dB}$ 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 range. 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
MaxFreq	Aria	929,07		
	Vocalise		1200,39	
	Discrete Pitch		1106,33	
	Physiological		1249,65	
MinFreq	Aria		245,11	
	Vocalise	160,24		
	Discrete Pitch	160,80		
	Physiological	138,19		
OctRge	Aria	1,94		
	Vocalise		2,93	2,93
	Discrete Pitch		2,82	
	Physiological			3,20
MinSPL	Aria			71,33
	Vocalise		65,90	
	Discrete Pitch		63,21	
	Physiological	51,13		
RgeSPL	Aria	39,33		
	Vocalise		46,80	
	Discrete Pitch		50,45	
	Physiological			62,56
ExtSPL	Aria	13,77		
	Vocalise	16,22	16,22	
	Discrete Pitch		17,35	
	Physiological			27,38
Area	Aria	328,67		
	Vocalise		556,03	
	Discrete Pitch		528,83	
	Physiological			744,44
Perc90	Aria		50,56	
	Vocalise		44,63	
	Discrete Pitch		43,99	
	Physiological	30,68		

Alpha = 0.01

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 ($p<0.01$).
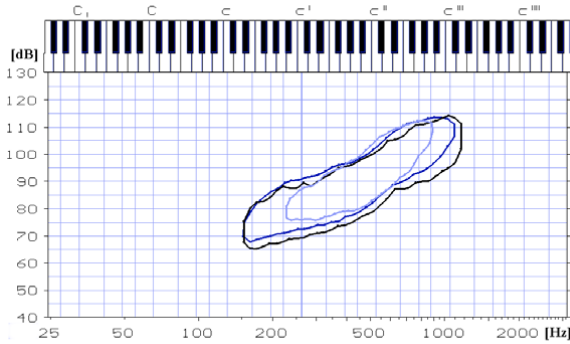
Figure 7. Average VRP contours for the discrete pitch task (task 3) in black, vocalise task (task 4) in blue; aria excerpt (task 5) in light grey (N=30).
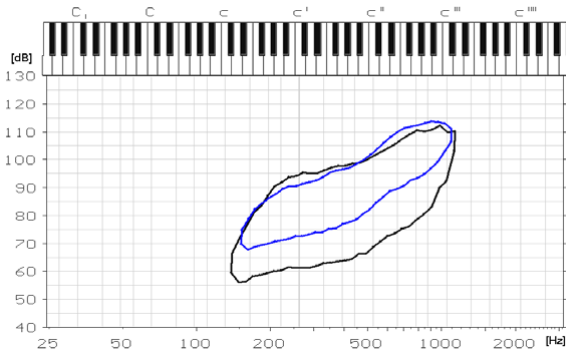

Figure 8. Comparison of averaged contours for the physiological (Task 2-black) and vocalise (Task 4-blue) tasks, for a group of 16 singers.

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 [27]. 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 [29]. 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 5 for detailed SRP results.

According to Hacki, a speech profile in normal cases should be approximately ⅓ of the VRP [30]. 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 3.1 octave range. In other words, the SRPs recorded in our study occupied the bottom third of the VRP_{phys}, covering 31 to 37% of its total area. When related to the performance profiles, SRPs covered 40-41% of the VRP_{perf}. area. Figures 4 a) and 4 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 [31]. 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"[30]. 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, 23]. 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 [33]. 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 [34].   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.
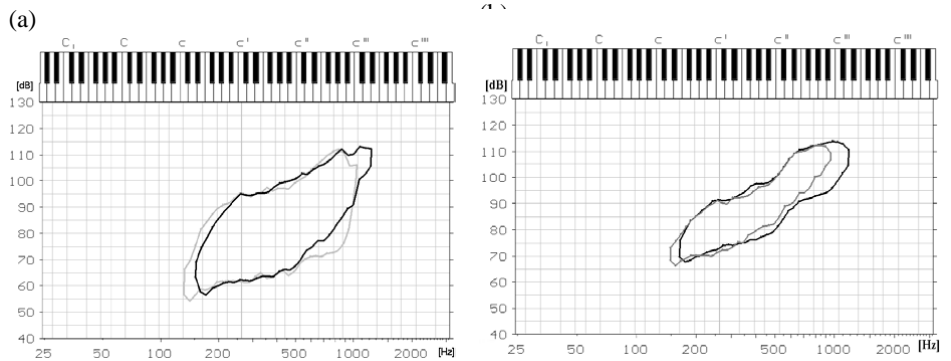

Figure 9 Average VRP contours for the physiological task (a) and the performance tasks (b). Sopranos in black (N=8) and contraltos in light grey (N=6).

Effects of Task

Reich *et al.* (1989, 1990) tested thoroughly the effect of different tasks in recording the frequency ranges of children and adults [35]. In those studies it was concluded that continuous tasks such as glissandi or small steps tasks 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* [36].

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 [36]. 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 et al.'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 et al. conclude 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 $Percent_{\geq 90dB}$ 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 $Percent_{\geq 90dB}$, 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 [34]. 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 [30]). 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 [37]. 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.
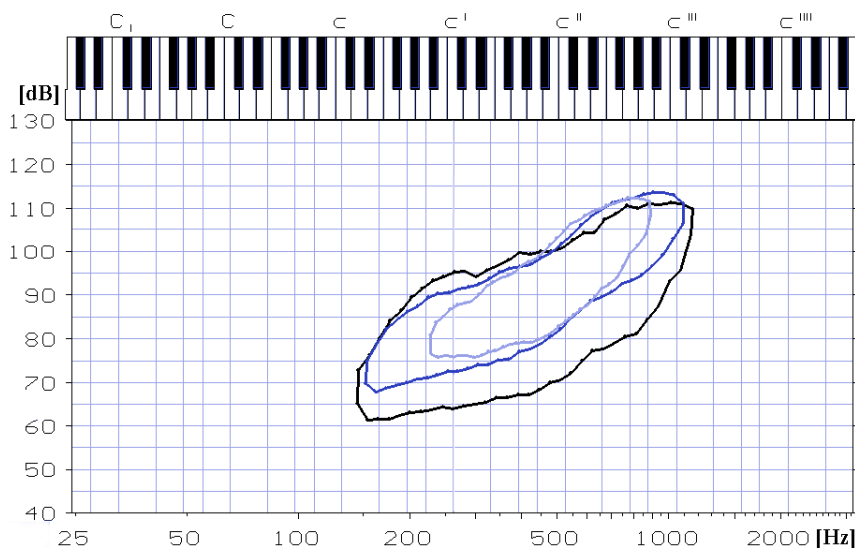
Figure 10. Physiological (in black) and performance contours (aria in light grey and vocalise in blue) for merged soprano, mezzo and contralto groups (N=30). At low levels, the performance contours are well contained by the physiological contour and even align at the low maximum curve rise. At high levels, however, the performance contours exceed the physiological contour, in the uppermost region of the maximum curve.

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 et al.'s result of actors studied in both on-stage and in-studio monologues seems to support this [36]. 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.
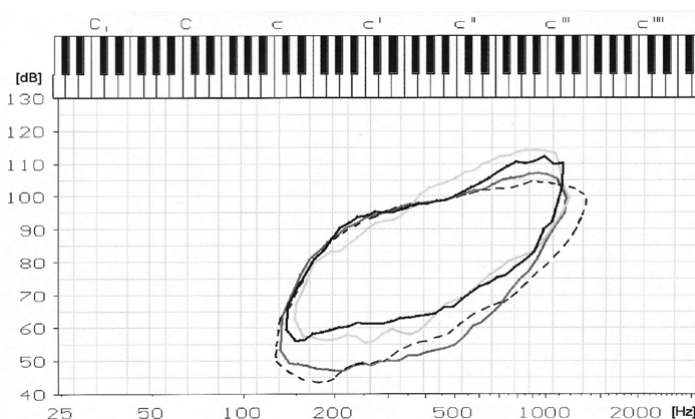

Figure 11. Contour averaging for singer groups. Data is representative of physiological VRPs. Lamarche data in dark blue, N=16, professional classical singers. Sulter data in broken line, N=42, choir singers with +-2 years experience [1]. Pabon data in blue, N=23, classical singing students (unpublished). 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 Western 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 (KayPENTAX, Lincoln Park/NJ) [11]. (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 phonetograph 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 mimimum 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 not only based 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

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, Komiyama 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. Bunch M, Chapman J. Taxonomy of Singers Used as Subjects in Scientific Research. *JVoice.* 2000;14(3):363-369.
20. Schutte H K, Seidner W. Recommendation by the Union of European Phoniatricians (UEP):Standardizing Voice Area Measurements/Phonetography. *Folia Phoniatr.* 1994;35:286-88.
21. Lamarche A, Ternström S. An Exploration of Skin Acceleration Level as a Measure of Phonatory Function in Singing. *JVoice.* 2008;22(1):10-22.
22. Awan S N. Phonetographic profiles and F0-SPL characteristics of untrained versus trained vocal groups *JVoice.* 1991;5:41-50.
23. Awan S N. Superimposition of speaking voice characteristics and phonetograms in untrained and trained vocal groups. *JVoice.* 1993;7:30–37.
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. Kitzing P. *Glottografisk frekvensindikering.En undersökningsmetod för mätning av röstläge och röstomfång samt framställning av röstfrekvensdistributionen.* PhD thesis, Lund University 1979.
30. Hacki T. Comparative speaking, shouting and singing voice range profile measurement: physiological and pathological aspects *Logoped Phoniatr Vocol.* 1996;21:123-129.
31. Coleman R F. Sources of variation in phonetograms. *JVoice.* 1993;7:1–14.
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. 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.
35. 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.
36. 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.
37. 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.
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.

Paper II

An Exploration of Skin Acceleration Level as a Measure of Phonatory Function in Singing

Anick Lamarche and Sten Ternström

Stockholm, Sweden

Summary: Two kinds of fluctuations are observed in phonetogram recordings of singing. Sound pressure level (SPL) can vary due to vibrato and also due to the effect of open and closed vowels. Since vowel variation is mostly a consequence of vocal tract modification and is not directly related to phonatory function, it could be helpful to suppress such variation when studying phonation. Skin acceleration level (SAL), measured at the jugular notch and on the sternum, might be less influenced by effects of the vocal tract. It is explored in this study as an alternative measure to SPL. Five female singers sang vowel series on selected pitches and in different tasks. Recorded data were used to investigate two null hypotheses: (1) SPL and SAL are equally influenced by vowel variation and (2) SPL and SAL are equally correlated to subglottal pressure (P_S). Interestingly, the vowel variation effect was small in both SPL and SAL. Furthermore, in comparison to SPL, SAL correlated weakly to P_S. SAL exhibited practically no dependence on fundamental frequency, rather, its major determinant was the musical dynamic. This results in a non-sloping, square-like phonetogram contour. These outcomes show that SAL potentially can facilitate phonetographic analysis of the singing voice.

Key Words: Singing voice—Skin acceleration level—Phonetogram—Vocal function—Vowel variation—Across tone fluctuations—Differences between singing and speech.

INTRODUCTION

Rationales

The vocal folds are a vibrating system and phonatory problems are likely to be most pronounced

at certain frequencies and/or amplitudes of vibration. The phonetogram offers a convenient mapping of vocal effort and fundamental frequency (F_0), and might therefore be useful in delimiting problem areas. In speech, the phonetogram or the voice range profile is used extensively in research and clinical settings.[1–13] This technique has also been applied to the classical singing voice.[14–18] However, classical singers train to maximize vocal output by means of vocal tract modifications. This implies that, in a phonetogram of a singing voice, the relationship of sound pressure level (SPL) to F_0 and vocal effort differs from that in a speech phonetogram. This difference is important and needs to be considered in the interpretation of

phonetograms. Indeed, in recording phonetograms of singing voices, SPL can vary within tone due to vibrato and across tones in regard to singer specific vowel modifications. Since these variations, to a large extent, are consequences of the vocal tract acoustics and are not directly related to phonatory function, it would be useful to minimize them when phonation is the primary object of study. As a measure of vocal function, the electroglottogram or EGG has advantages. It is minimally influenced, if at all, by vowel production. On the other hand, Askenfelt et al as well as Baken demonstrated that EGG has limitations in depicting vocal fold oscillations. The EGG does not have any microphonic capacities and furthermore, it cannot provide any information for the open phase of vocal fold oscillation.[19,20]

The skin acceleration level (SAL), if measured in the vicinity of the vocal folds, is another potential measure of phonatory activity. If the objective is to evaluate phonatory function, it is more relevant to estimate the intensity of the glottal source rather than the intensity of the radiated sound. In a 1983 study of chest wall vibrations, Sundberg noted that vibrations measured at the thyroid and the sternum are primarily determined by the voice source and to some extent modified by subglottal resonances. Hence it might be expected that vibrations measured at these locations would be less influenced by changes in the vocal tract. Moreover, it becomes a possible alternative for the vertical axis in the phonetogram and a replacement for SPL. SAL is mainly a measure of tissue vibrations rather than a measure of acoustic pressure and it is easily recorded near the vocal folds. One might also expect the vocal fold collisions to generate shock waves in the surrounding tissues. However, Sundberg investigated possible influences of colliding forces of the vocal folds and concluded that their contribution to vibrations recorded at the thyroid and sternum lamina is negligible.[21]

The subglottal pressure (P_S) drives the voice source. P_S is a main determinant of vocal loudness in speech and in singing and the literature demonstrates how P_S relates to SPL for both speech and singing. Therefore, it could be interesting to observe how SAL and SPL differ in their relationship to P_S. Generally, SAL seems to have the potential

to: (1) facilitate phonetographic analysis of the singing voice, (2) allow inclusion of all vowels in clinical evaluation, (3) address directly and unobtrusively the voice source, (4) allow singers more vocal and physical freedom during recordings, and (5) reduce influence of environmental noise on the recorded signal.

Earlier work

Accelerometers have occasionally been applied to speech and voice research, for example, in research on nasalization,[22–24] F_0 extraction,[19,25] frequency perturbation,[25] and alternative recording devices.[26] Recent studies have looked at SAL as an estimator for speech glottal characteristics[27] and also as an estimator for SPL for speech.[28] Švec et al showed that a near-to-linear relationship between SPL and SAL can be used to estimate long-term average SPL values in speech. Their data clearly show an established SAL correlation to long-term SPL: higher SPL corresponds to higher SAL in speech. The primary motivation for the current investigation was to find alternatives to the study of phonatory function in singers that would facilitate the interpretation of phonetograms. The questions to be answered were (1) does SAL vary less across vowels than does SPL? (2) is SAL more correlated to subglottal pressure than SPL? and finally (3) how does SAL measured in singers compare to findings for speech by Švec et al?

Hypotheses

The first question that was stated above leads to null hypothesis A: SPL and SAL are equally influenced by vowel variation. Our second question leads to null hypothesis B: SAL and SPL are equally correlated to P_S.

METHOD

To test these hypotheses, a number of singing tasks were designed to exercise variations in vowel, musical dynamic, and F_0 over a typical female singing range. Musical dynamic was included to obtain systematic variation in P_S. Through statistical analysis, the variance thereby incurred in SAL was compared to the variances incurred in SPL and P_S.

Each subject was instructed to warm up, before her arrival at the recording session, for a minimum of 5 minutes, and according to their personal warm-up routines. On arrival, the experimental procedure and tasks were explained. Subjects familiarized with the equipment and made a few trials. All recordings were performed at the NCVS Laboratories in Denver, Colorado. Recordings took place in a sound-isolated booth. Singers were asked to use a stage stance throughout the recording process. The experimenter was present to coach through different tasks as well as to monitor P_S signals on the oscilloscope.

Acoustic, aerodynamic, and accelerometric signals were recorded with the following equipment. Two accelerometers (Thin Case BU-7135; Knowles Acoustics, St. Louis, MO): one attached vertically at midline on the jugular notch and the other at midline on the sternum bone. Attachments and use of Mastisol surgical glue (Mastisol; Ferndale Laboratories, Ferndale, MI) and Suture-Strips (TS-3101; Derma Sciences, Elgin, IL) followed the protocol established in Popolo et al.[29] The airborne signal was recorded at 30 cm from the subject's mouth with the microphone of the sound level meter (Brüel & Kjaer 2238 Mediator, A weighted-slow; Brüel & Kjaer, Naerum, Denmark). Intraoral pressure during stop-plosives /p/ was measured with a pressure transducer (PT-series; Glottal Enterprises, Syracuse, NY).[30] Subjects were given the transducer to hold at the labial commissure during the performance of phonation tasks. The subjects familiarized themselves with the equipment and received brief oscilloscope feedback to facilitate the positioning of the

pressure transducer in their mouth and achieve a stable intraoral pressure during /p/ occlusions.

The microphone/sound level meter was connected through an amplifier to channel 0 of a model 4500 Kay CSL sound card (KayPentax, Lincoln Park, NJ). The pressure transducer was connected to channel 1 (DC) of the same card and the accelerometers to channels 2 and 3. A 20-dB attenuator pad (DGS pro-audio; Mouser Electronics, Mansfield, TX) was used when necessary to prevent clipping of the microphone signals (Figure 1 depicts the setup schematics). The sampling rate was 44 100 Hz. The four channels were recorded in synchrony and the resulting files were read and edited with *Cubase S.L.* (Version 1.07 build 97\2004 SE; Steinberg Media Technologies GmbH, Hamburg, Germany).

Calibration

Microphone and pressure transducer calibrations were performed at the beginning and the end of each subject's session. Accelerometer calibration followed NCVS-established calibration procedures for speech dosimetry (A. Starr, personal communication, August 2005). For the sound level calibrations, each subject phonated at three loudness levels and gains were adjusted to avoid clipping. The *Cubase S.L.* program was set to record position and a calibrator (Brüel & Kjaer 4231) was used to produce a 94-dB SPL re 20-µPa tone. Finally, pressure transducer calibrations were performed with a pneumotach calibration unit (Glottal Enterprises, Model MCU-4). Readings at 20, 10, and 5 cm water column were taken and recorded in the *Cubase S.L.* program.
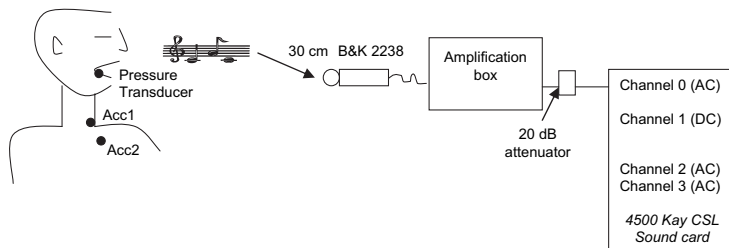

FIGURE 1. Schematic representation of the experimental setup. Accelerometers were attached at the jugular notch (Acc 1) and on the sternum bone (Acc 2) according to a protocol developed by Popolo et al.[29]

Subjects and vocal tasks

Five female singers, three sopranos and two mezzo-sopranos, aged 20 to 30 years, participated in the recordings. Each singer had obtained a university certification in voice performance or formal classical training. Levels ranged from bachelor to DMA. It must be specified that only one singer met the criteria established for a professional singer.[31] All singers reported good vocal health.

The subjects performed three tasks (Figure 2):

(1) Sustain a tone at D5 (587 Hz) while singing a /pi pe pa po pu/ series in a slow tempo. This task was performed at three intensity levels (*piano*, *mezzo forte*, and *forte*). The exact task was then repeated at G5 (784 Hz).

(2) Sing an ascending scale of an octave starting at a preferred F_0 and repeat each F_0 three times using the vowels /a/ and /i/. Again, this task was performed at all three intensity levels mentioned above, with /p/ occlusions preceding the vowel. Two subjects chose a C3 to C4 (131–262 Hz) scale, one a G4 to G5 (392–784 Hz), and two others D4 to D5 (294 Hz–587 Hz).

(3) Arpeggiate an octave from F4 (349 Hz), repeating each F_0 three times. The task was performed at all three intensity levels and included /p/ occlusions and all /i e a o u/ vowels.

Each performance was carefully monitored and the tasks were repeated if, for example, the

FIGURE 2. Illustrations of the three tasks that subjects completed during recording.

oscilloscope displayed unstable P_S signals or if singers believed they could perform higher dynamic contrasts. At the end of each recording session, subjects filled out questionnaires concerning their voice and vocal experiences.

Data processing

Recorded files were truncated from 24- to 16-bit samples and they were losslessly compressed in *Flac* (Frontend 1.7.1, FLAC, http://flac.sourceforge.net). Each channel was saved separately and reopened as a .wav file and converted to .smp format with a file conversion utility (Audiofil; Hitech Development AB, Täby, Sweden). Files were then reorganized back into synchronized four-channel files. The pressure value corresponding to the onset of phonation was taken as the pressure immediately before the release of the plosive /p/ (Figure 3). In measuring P_S, pressure tokens were discarded if the /p/ occlusion and phonation were not perfectly aligned. This was seen in the case where singers did not always succeed in keeping a sustained *legato* from one plosive occlusion to the next. Tokens

were also discarded if they displayed instability, too much sharpness, or when a breath was taken. L_{eq} values were computed over the initial 200 milliseconds of each vowel sound, following the /p/. All signal manipulations and measurements were done using the Soundswell Signal Workstation 4.0 (Hitech Development AB, Täby, Sweden).

To make phonetograms of the microphone and accelerometer signals, the signal files were resampled to 16 kHz per channel. This was a requirement of the computerized phonetograph (Phog 2.0, Hitech Development AB, Täby, Sweden). Conventional phonetograms as well as SAL phonetograms were made of the complete recordings of each subject.

Statistical analysis

A univariate general linear model–based analysis of covariance (ANCOVA) was designed. Dependent variables were defined as SPL, SAL_N (SAL for notch), and SAL_S (SAL for sternum) and independent variables as F_0, Dynamic, Vowel, and Subject. A univariate format was preferred to
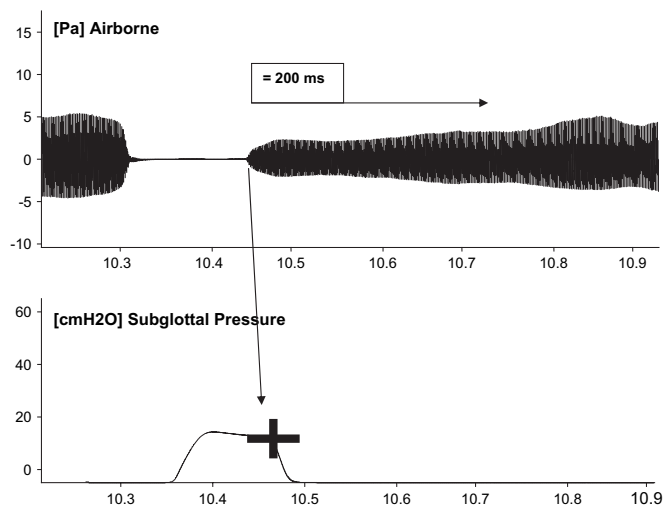

FIGURE 3. An example of the analysis points selected in audio and pressure signals. The intraoral pressure at *p*-release was used as an approximation of the subglottal pressure driving the first 200 milliseconds of subsequent phonation.

a multivariate to assess dependent variable behavior in isolation. Subject, Dynamic, and Vowel were treated as fixed factors, while F_0 was defined as a covariate.

The data were organized into two factorial designs, each with a balanced data set representing different levels of factors. For design 1, data recorded from the first task was combined with data from the third task. In the tasks for design 1, the subjects changed only the vowel or dynamic from token to token, while holding the F_0 constant (Table 1). For design 2, data from the second task were used (Table 2), in which subjects changed only the F_0 from token to token. Dividing the data into two groups by tasks should offer some insight as to the importance of tasks in the overall outcome. The division of the data also offers some indication of the reliability of the behavior observed across designs.

SAL was not calibrated against a reference level since SAL can be expected to vary from subject to subject, due to physiology and possible variations in transducer attachment. The intersubject variation in SAL is not relevant to this study. Rather, the SAL data were normalized by subject means, thereby excluding the expected variations in the gain of the SAL signals. This was done for each subject by computing the intrasubject average SAL within one design and one attachment (notch/sternum), and then subtracting the personal average from the raw SAL values. The SAL data were not normalized for standard deviation, since the variance in SAL is one of the outcomes of the experiment. Testing was performed using statistical software (*SPSS* Version 13.0; SPSS, Chicago, IL). The significance threshold was set to $P \leq 0.05$.

TABLE 1. *Statistical Design 1 (60 Tokens per Subject) Used to Obtain a Balance Set of Data for a General Linear Model ANCOVA Analysis*

Independent Variables	Statistical Label	Number	Definition
F_0	Covariate	4	349, 440, 587, 784 Hz
Subject	Fixed factor	5	2 mezzos, 3 sopranos
Vowel	Fixed factor	5	/pa/ /pe/ /pi/ /po/ /pu/
Dynamic	Fixed factor	3	*p—mf—f*

F_0 was defined as a covariate in the final model used.

RESULTS

Figure 4 depicts the collected data before overall statistical treatment and normalization of SAL. For each subject, the means and standard deviations obtained for three dependent factors are illustrated. SPL tended to be similar across all subjects, but both SAL signals showed some intersubject variation.

Design 1

The ANCOVA results for design 1 are given in Table 3. Interestingly, the chosen statistical model explained most of the variance in the data (see percentages in Table 3), provided that F_0 was defined as a covariate. It can be seen from the P values that the factors F_0 and Dynamic had a significant effect (at $P \leq 0.05$) on all the dependent variables: SPL, SAL_N, SAL_S, and P_S. This is, of course, as expected for SPL and P_S, since the voice output level rises with both F_0 and Dynamic. The expectations for SAL are not obvious. The factor Subject also had a significant effect on SPL and P_S (individuals differ in their choice of vocal power), but not on the two SAL measures, since they had been previously normalized. The Vowel factor was a significant source of variation in SPL and P_S; however, the percentages of explained variance for the Vowel factor are much smaller than for other significant factor percentages. In Table 3, further comparisons of F values and percentage explained variance show that for SPL, F_0 was clearly the dominant source of variation (stronger even than Dynamic); whereas for SAL, Dynamic was the dominant source of variation. For all four dependent variables, Vowel was the weakest source of variation.

For the dependent variable SPL, there were no significant interactions between the fixed factors. For the two SAL measures and for P_S, there were significant but small interaction effects between Subject and Dynamic (1–3%). This means that different subjects produced slightly different increments in SAL and P_S between *piano*, *mezzo forte*, and *forte*. For P_S, there was also a significant interaction between Subject and Vowel; in other words, different subjects would exhibit different changes in P_S when changing vowel. Since we are not

TABLE 2. *Statistical Design 2 (144 Tokens per Subject) Used to Obtain a Balance Set of Data for a General Linear Model ANCOVA Analysis*

Independent Variables	Statistical Label	Number	Definition
F_0	Covariate	8	C major scale <262–523 Hz>
			D major scale <294–587 Hz>
			G major scale <392–784 Hz>
Subject	Fixed factor	5	2 mezzos
			3 sopranos
Vowel	Fixed factor	2×3 repetitions	/papapa/ /pipipi/
Dynamic	Fixed factor	3	p—mf—f

F_0 was defined as a covariate in the final model used.

here concerned with individual behaviors, these interactions will not be discussed further.

Design 2

The ANCOVA results for design 2 are summarized in Table 4. Again, the statistical model seems to explain most of the variance (see percentages in Table 4). The overall pattern in the outcome was the same as was observed in design 1, with F_0 being the dominant source of variation for SPL, while Dynamic was the dominant source of variation for both the SAL measures. Generally, some significance levels were higher than those found for design 1 and Vowel variation presented a different pattern of significance; both SPL and SAL_N were significant for Vowel. On the other hand, SAL_S and P_S did not significantly change with Vowel. In this design, the F values were larger than those found for design 1, with the exception of Vowel for P_S but percentages of the explained variance are slightly lower. Nonetheless, design 2 confirms the most striking result of this study: in both statistical designs, F_0 was the dominant variation factor for SPL, while Dynamic was the dominant variation factor for SAL.

In the second design, practically all interactions between the fixed factors were significant, for all dependent variables, but their influence was small (1–3% of the variance explained). For the purpose of this study, these interactions do not seem to warrant a more detailed discussion.

Subglottal pressure

Finally, the P_S-SPL and P_S-SAL correlations were computed (Figures 5 and 6). SPL showed clearly *higher* correlation to P_S than did SAL.

This was true for both design 1 and design 2. The null hypothesis B, that SPL and SAL are equally correlated to P_S, is therefore rejected.

Examples of phonetograms are shown for one subject in Figures 7–9.

DISCUSSION

The first hypothesis of this study called for the investigation of the presence of vowel variation in SAL. Similar to the work done in Švec et al,[28] vowel-induced variation in SAL and SPL was compared (Figures 10 and 11). Generally, speech phonetograms are recorded with the /a/ vowel to avoid variation in SPL between open and closed vowels. In the soprano singing tasks of the present
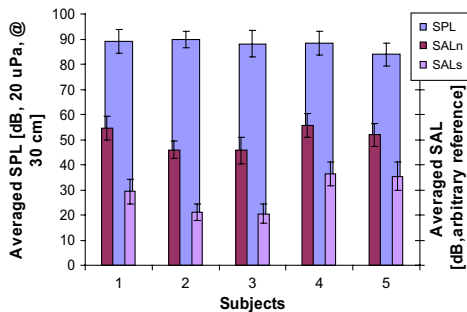

FIGURE 4. Subject means and standard deviations for three dependent variables: SPL, SAL_N, and SAL_S. For all subjects, SPL was quite similar. In this figure, SAL_N and SAL_S are not normalized and some intersubject variation can be observed. Intersubject differences in mean SAL were not important for the aims of this study and were canceled by normalizing to the subject mean.

TABLE 3. *Design 1, Test Between-Subject Effects*

Variables	SPL	SAL$_N$	SAL$_S$	P_S
P				
F_0	0,000	0,006	0,000	0,000
Subject	0,000	(1,000)	(0,999)	0,000
Vowel	0,050	0,175	0,110	0,000
Dynamic	0,000	0,000	0,000	0,000
F				
F_0	**486,676**	7,821	66,811	**291,398**
Subject	22,108	(0,010)	(0,025)	103,510
Vowel	**2,406**	**1,603**	**1,911**	**6,074**
Dynamic	62,287	**155,722**	**171,808**	90,457
% of Explained variance				
F_0	**49**	1	9	**24**
Subject	9	0	0	34
Vowel	**1**	**1**	**1**	**2**
Dynamic	13	**51**	48	15
R^2 (% of explained variance by model)	78	63	69	87
Observed power for $\alpha = 0.05$				
F_0	1,000	,795	1,000	1,000
Subject	1,000	(,052)	(0,055)	1,000
Vowel	0,686	,490	0,572	0,985
Dynamic	1,000	1,000	1,000	1,000

Highest *F* values, all *F* values for vowel, and corresponding percentages of explained variance are given in bold to clearly depict the magnitude poles in the data. Values for the factor Subject and dependent variables SAL$_N$ and SAL$_S$ are in parentheses, since SAL was normalized for each subject. Frequency has a dominating influence on SPL variation but it is Dynamic which dominates SAL$_N$ and SAL$_S$. Interestingly, the Vowel factor does not explain much of the variance for either SPL or SAL.

experiment, vowel variation was found to be an almost negligible source of SPL variation, when compared to the other experimental factors. F_0 was the dominant factor in terms of variation in SPL, and this result is supported by the literature. It is known that SPL in speech increases by approximately 9 dB per octave.[32] The corresponding slope values observed in this study ranged from 20 to 30 dB per octave. Hence, in soprano singing, F_0 has a considerably stronger influence on SPL than it does in speech. This could be due to the F_1-F_0 matching that is conventional in high-pitched female singing. This matching would presumably become more precise with rising F_0.

The near absence of vowel variation in SPL and the strong F_0 dependency observed here both confirm the need for differentiation between speech and singing behaviors. Singers operate their vocal instrument characteristically on many different levels, and a number of compensations can be at play in the production of an equal loudness and timbre across vowels. This raises the question whether vowel variation in P_S could be indicative of compensatory adjustments at the voice source. If so, we would expect the Vowel factor to be a stronger source of variation for P_S than for SPL. However, the ANCOVA results show that this was not the case. Rather, the present data suggest that if the singers systematically modify P_S with vowel, then such modifications are very small.

Overall, the results suggest that SAL cannot be proposed as a useful replacement for SPL merely on the grounds that it is a signal with reduced vowel variation. The first null hypothesis driving this work is, therefore, *not* rejected by our findings: in the singing voice tasks used, vowel changes caused little or no SPL variation, and in practice were negligible also in SAL.

TABLE 4. *Design 2, Test Between-Subject Effects*

Variables	SPL	SAL_N	SAL_S	P_S
P				
F_0	0,000	0,000	0,000	0,000
Subject	0,000	(0,000)	(0,000)	0,000
Vowel	0,000	0,000	0,055	0,206
Dynamic	0,000	0,000	0,000	0,000
F				
F_0	**1863,362**	145,144	503,107	**1040,738**
Subject	105,972	(11,081)	(38,642)	104,369
Vowel	**36,538**	**57,639**	**3,702**	**1,604**
Dynamic	539,606	**763,143**	**838,468**	534,905
% of Explained variance				
F_0	**44**	5	16	**28**
Subject	10	2	5	11
Vowel	**1**	**2**	**0**	**0**
Dynamic	25	57	53	29
R^2 (% of explained variance by model)	84	75	78	85
Observed power				
for $\alpha = 0.05$				
F_0	1,000	1,000	1,000	1,000
Subject	1,000	(1,000)	(1,000)	1,000
Vowel	1,000	1,000	0,485	0,244
Dynamic	1,000	1,000	1,000	1,000

Highest *F* values, all *F* values for vowel, and corresponding percentages of explained variance are given in bold to clearly depict the magnitude poles in the data. Values for the factor Subject and dependent variables SAL_N and SAL_S are in parentheses, since SAL was normalized for each subject. Design 2 statistics show the trend observed in design 1 where Frequency is a dominant factor for SPL and Dynamic is the important factor for SAL. Here as well, the Vowel factor has a very small effect.

Although vowel variation was the primary topic of this study, other outcomes revealed some potentially useful aspects of SAL. The pronounced dependency of SPL on F_0 (20–30 dB per octave) is practically eliminated in SAL. For this study, the clear reduction of the influence of F_0 in SAL when compared to SPL is very interesting. F_0 remains a statistically significant source of variation across all dependent variables. However, although F_0 is significant for SAL, *F* values and percentages of explained variance are much lower than those for Dynamic, and thus indicate a weaker source of variation. The literature gives explanation for the reduced F_0 variation in SAL. In his study of chest wall vibrations, Sundberg demonstrated how sternum displacement amplitude lines up along a 12-dB slope when plotted according to F_0 and a constant vocal effort.[21] Because acceleration is the second derivative of displacement, it is expected that the frequency related slope in SAL will have 12 dB less in inclination than for the displacement slope. This essentially agrees with the outcome of the present study.

For phonetography, these results point to the necessity for clear differentiation between the analysis of speech and high F_0 soprano singing. In speech studies, SPL estimation by SAL alone is successful whereas in the case of singing, this type of estimation would need to account for F_0.

In what concerns Dynamic, changes are somewhat smaller in SAL than in SPL (Figures 12 and 13). Nevertheless, the Dynamic variation in SAL is more explanatory than the one observed in SPL. SPL embeds a combination of different factors that work together in amplifying the voice. As seen in results obtained above, F_0 is the most important of these factors. Since there is a reduced F_0 effect, the source of variation in SAL is mostly attributed to the Dynamic factor. Indeed, results demonstrate clearly the dominant influence of
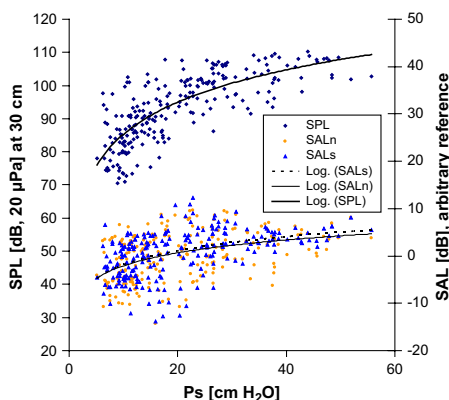
FIGURE 5. Correlations found for SPL, SAL_N, and SAL_S to P_S according to statistical design 1. The regression outcomes were $Y_{SPL} = 14 \ln(x) + 53$, $r^2 = 0{,}5968$, $Y_{SAL_N} = 4 \ln(x) - 11$, $r^2 = 0{,}1833$, and $Y_{SAL_S} = 4 \ln(x) - 12$, $r^2 = 0{,}2455$. SAL for both attachments was only very weakly correlated to subglottic pressure, whereas SPL followed trends documented in the literature. The null hypothesis that SPL and SAL are equally correlated to P_S is therefore rejected.
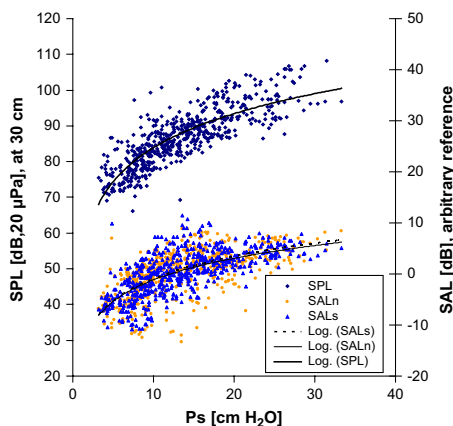

FIGURE 6. Depicted here are the correlations found for SPL, SAL_N, and SAL_S to P_S according to statistical design 2. Slopes and regressions are defined by the following equations: $Y_{SPL} = 13 \ln(x) + 52$, $r^2 = 0{,}6732$, $Y_{SAL_N} = 6 \ln(x) - 15$, $r^2 = 0{,}4171$, and $Y_{SAL_S} = 6 \ln(x) - 15$, $r^2 = 0{,}4771$. Clearly, SAL for both attachments is only very weakly correlated to subglottic pressure, whereas SPL reflects trends documented in the literature. The null hypothesis that SPL and SAL are equally correlated to P_S is therefore rejected.

Dynamic as a source of variation in SAL. This points to the potential of SAL variation to display more immediate information and could result in interesting implications for phonetograms.

The phonetograms in Figures 7–9 exemplify clearly the type of results obtained when SAL is substituted for SPL on the phonetogram *y*-axis. The distribution shape of phonetogram changes from steeply inclined to horizontal and almost rectangular. SAL might therefore simplify the interpretation of the phonetogram, by showing results without the usual bias due to F_0, which is even stronger in singing than in speech.

The second hypothesis concerned the subglottal pressure. Since P_S drives the vocal chords, the expectation was for P_S to have equal or more correlation to SAL at the notch and at the sternum than what is observed for SPL. This expectation was not borne out by the results. Nevertheless, the P_S-SPL relationship was similar to that which has been reported in previous literature. In speech, Fant originally established a 9.5-dB theoretical increase in SPL for every doubling in P_S.[33] In singers, Sjölander and Sundberg, in agreement with Schutte's studies,[34] observed that the decibel

increase was higher. They established an average of 12 dB.[35] It is interesting to note that those reports addressed only the male singing voice. According to our results, the relationship between
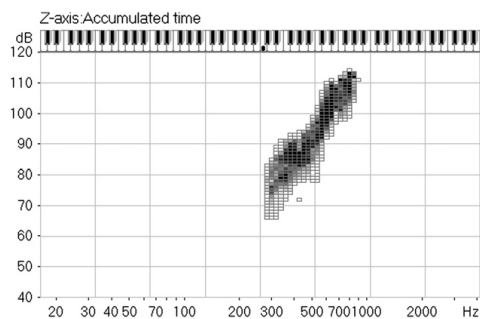

FIGURE 7. Aggregate phonetogram of all tasks performed by subject 2. The format below is the standard display used in clinics and in experiments, with SPL on the *y*-axis and log frequency on the *x*-axis. The phonetogram exhibits a pronounced slope with frequency, and shows a 20- to 30-dB increase per octave.
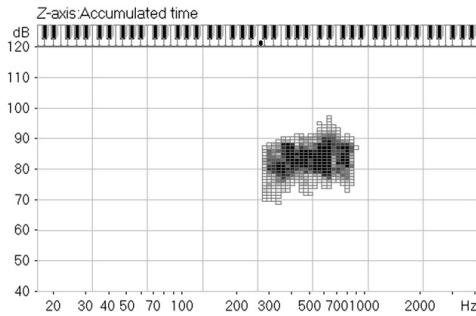
FIGURE 8. Alternative phonetogram of all tasks performed by subject 2, with SAL$_N$ (at the jugular notch) on the *y*-axis rather than SPL. Although a slight slope remains, the observed dominance of the Frequency factor in the traditional phonetogram is almost gone. In this format, the Dynamic factor is the major source of level variation. This has the potential to simplify phonetogram interpretation.

SPL and P_S reported for male singers also holds true for female voices. From the regression equations shown in Figures 4 and 5, we find that SPL on average increased +12 dB per doubling of P_S in design 1 and +11 dB in design 2.

Unexpectedly, SAL showed weaker correlation to P_S than what was found for SPL. These findings oppose the null hypothesis above mentioned. A possible explanation could be the spectral
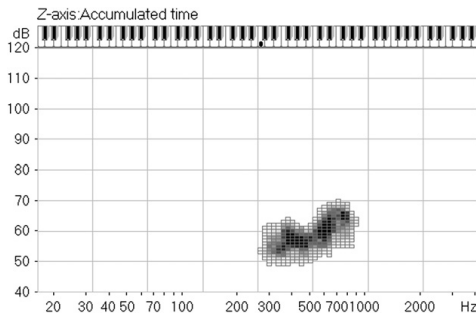

FIGURE 9. Alternative phonetogram of all tasks performed by subject 2, with SAL$_N$ (at the jugular notch) on the *y*-axis rather than SPL. Although a slight slope remains, the observed dominance of the Frequency factor in the traditional phonetogram is almost gone. In this format, the Dynamic factor is the major source of level variation. This has the potential to simplify phonetogram interpretation.
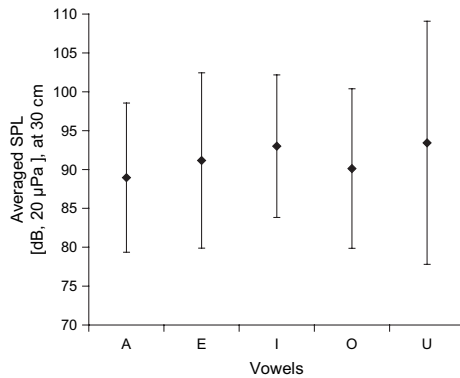
FIGURE 10. Overall means and standard deviations in SPL, by vowel. The means changed less than 4 dB between vowels, which would probably be negligible in practice.

characteristics of the skin acceleration signal. SAL is dominated by the level of the first partial. If the effect of increasing P_S is mostly to boost the rather weak higher partials, then there would be very little effect on the overall signal level. This issue would require further study. Until this is clarified, the SAL correlation to P_S does not in itself support the use of SAL as an alternative to SPL or as a method for vocal function quantification.
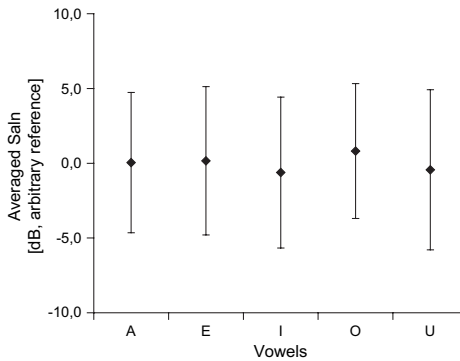

FIGURE 11. Overall means and standard deviations in SAL$_N$, by vowel. In comparison to the vowel variation observed in SPL, the variation in SAL$_N$ was even smaller, and did not have any important effect on the signal. The outcome for SAL$_S$ was practically identical.
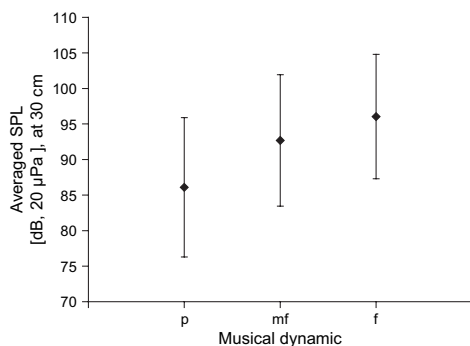
FIGURE 12. Overall means and standard deviations in SPL, by Dynamic. The Dynamic variation in SPL was considerable, with a 10-dB difference in the means; however, it was smaller than the Frequency variation, which dominated SPL.

CONCLUSION

From this study, it appears that SAL does have the potential to (1) facilitate phonetographic analysis of the singing voice, (2) allow singers more vocal and movement freedom during recordings, and (3) reduce influence of environmental noise on the recorded signal. Nonetheless, SAL remains an indirect assessment of vocal function. It would be necessary to investigate further the spectral properties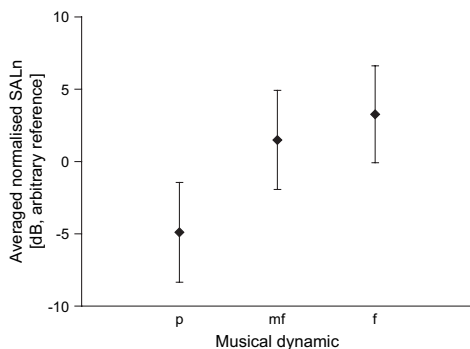 of the skin acceleration signal to assess in detail the behavior of the fundamental in SAL and how precisely dynamic is displayed.

SAL proves to be potentially useful to phonetography, but for different reasons than those initially expected. There are two main observations which have important consequences. First, SAL offers a signal which is minimally influenced by F_0 and therefore, is able to clearly illustrate effects of the musical dynamic. And second, vowel variation, when addressing the high singing voice, is practically negligible in both SPL and SAL. This finding is important in that it underpins differences between speech and singing. In singing, phonetograms might not be as influenced by the use of different vowels as they are in speech. This fact would allow for much more freedom in performing phonetograms of the singing voice given that singing tasks involving different vowels and song/aria excerpts could be used.

In the process of this investigation other pertinent questions were encountered. For example, are the SPL variations that are due to vibrato smaller in SAL phonetogram output? Since it is established that vowel variation for the singing voice has a minimal influence, it would be interesting to explore also the other type of fluctuations observed in real-time acquisition of phonetograms.

FIGURE 13. Overall means and standard deviations in SAL_N, by Dynamic. The Dynamic factor in SAL_N (7 dB) was important in that it dominated the variation in SAL_N: the standard deviations were smaller than the differences in the means. Again, the same was observed for SAL_S.

REFERENCES

1. Damste PH. The phonetogram. *Pract Otorhinolaryngol (Basel)*. 1970;32:185–187.
2. Dejonckere PH. Effects of phoniatric therapy on the phonetogram. *Acta Otorhinolaryngol Belg*. 1980;34:309–317.
3. Gramming P, Åkerlund L. Non-organic dysphonia II. Phonetograms for normal and pathological voices. *Acta Otolaryngol*. 1988;106:468–476.
4. Gramming P, Sundberg J. Spectrum factors relevant to phonetogram measurement. *J Acoust Soc Am*. 1988;83:2352–2360.
5. Pabon JP, Plomp R. Automatic phonetogram recording supplemented with acoustical voice-quality parameters. *J Speech Hear Res*. 1988;31:710–722.

6. Titze IR. Acoustic interpretation of the voice range profile (phonetogram). *J Speech Hear Res*. 1992;35:21–34.

7. Coleman RF. Sources of variation in phonetograms. *J Voice*. 1993;7:1–14.

8. Sulter AM, Wit HP, Schutte HK, Miller DG. A structured approach to voice range profile (phonetogram) analysis. *J Speech Hear Res*. 1994;37:1076–1085.

9. Behrman A, Agresti CJ, Blumstein E, Sharma G. Meaningful features of voice range profiles from patients with organic vocal fold pathology: a preliminary study. *J Voice*. 1996;10:269–283.

10. Hilgenheger C, Sarfati J, Reyt E, Sittel C, Eckel HE. Value of the relative phonetogram (RP) for the evaluation of organic dysphonias. *Rev Laryngol Otol Rhinol (Bord)*. 1999; 120:231–237.

11. Dejonckere PH. Efficacy of voice therapy assessed with the Voice Range Profile (Phonetogram). *Rev Laryngol Otol Rhinol (Bord)*. 2003;124:285–289.

12. Södersten M, Ternström S, Bohman M. Loud speech in realistic environmental noise: phonetogram data, perceptual voice quality, subjective ratings, and gender differences in healthy speakers. *J Voice*. 2005;19:29–46.

13. Ma E, Robertson J, Radford C, Vagne S, El-Halabi R, Yiu E. Reliability of speaking and maximum voice range measures in screening dysphonia. *J Voice*. May 4, 2006. In press [Epub ahead of print].

14. Klingholtz F. The voice of the singer in the phonetogram. *Laryngorhinootologie*. 1989;68:62–66.

15. Pedersen MF. A longitudinal pilot study on phonetograms/voice profiles in pre-pubertal choir boys. *Clin Otolaryngol*. 1993;18:488–491.

16. Awan SN. Superimposition of speaking voice characteristics and phonetograms in untrained and trained vocal groups. *J Voice*. 1993;7:30–37.

17. Sulter AM, Schutte HK, Miller DG. Differences in phonetogram features between male and female subjects with and without vocal training. *J Voice*. 1995;9:363–377.

18. Hunter EJ, Svec JG, Titze IR. Comparison of the produced and perceived voice range profiles in untrained and trained classical singers. *J Voice*. 2005. In press [Epub ahead of print].

19. Askenfelt A, Gauffin J, Sundberg J, Kitzing P. A comparison of contact microphone and electroglottograph for the measurement of vocal fundamental frequency. *J Speech Hear Res*. 1980;23:258–273.

20. Baken RJ. Electroglottography. *J Voice*. 1992;6:98–109.

21. Sundberg J. Chest wall vibrations in singers. *J Speech Hear Res*. 1983;25:329–340.

22. Stevens K, Kalikow D, Willemain T. A miniature accelerometer for detecting glottal waveforms and nasalization. *J Speech Hear Res*. 1975;18:594–599.

23. Lippmann RP. Detecting nasalization using a low-cost miniature accelerometer. *J Speech Hear Res*. 1981;24: 314–317.

24. Horii Y. An accelerometric measure as a physical correlate of perceived hypernasality in speech. *J Speech Hear Res*. 1983;26:476–480.

25. Horii Y. Some voice fundamental frequency characteristics of oral reading and spontaneous speech by hard-of-hearing young women. *J Speech Hear Res*. 1982;25: 608–610.

26. Coleman RF. Comparison of microphone and neck-mounted accelerometer monitoring of the performing voice. *J Voice*. 1988;2:200–205.

27. Cheyne II, HA. *Estimating glottal voicing source characteristics by measuring and modeling the acceleration of the skin on the neck* [doctoral thesis]. Massachusetts: Institute of Technology; 2002.

28. Švec JG, Titze IR, Popolo PS. Estimation of sound pressure levels of voice speech from skin vibration of the neck. *J Acoust Soc Am*. 2005;177:1386–1394.

29. Popolo PS, Svec JG, Hunter EJ, Starr AC, Rogge-Miller K, Nix J, Titze IR. Teachers' Guide to Voice Dosimetry. NCVS Memo No. 04 2004; NCVS Online Technical Memo.

30. Rothenberg M. Interpolating subglottal pressure from oral pressure. *J Speech Hear Disord*. 1982;47:218–224.

31. Bunch M, Chapman J. Taxonomy of singers used as subjects in scientific research. *J Voice*. 2000;14:363–369.

32. Titze IR, Sundberg J. Vocal intensity in speakers and singers. *J Acoust Soc Am*. 1992;91:2936–2946.

33. Fant G. Preliminaries to analysis of the human voice source. *KTH Speech Music Hear Quart Prog Stat*. 1982; Report 4:1–27.

34. Schutte H. *The efficiency of voice production* [doctoral dissertation]. The Netherlands: Groningen University; 1980.

35. Sjölander P, Sundberg J. Spectrum effects of subglottal pressure variation in professional baritone singers. *J Acoust Soc Am*. 2004;115:1270–1273.

Paper III

informa
healthcare

ORIGINAL ARTICLE

Not just sound: Supplementing the voice range profile with the singer's own perceptions of vocal challenges

ANICK LAMARCHE[1], STEN TERNSTRÖM[1] & STELLAN HERTEGÅRD[2]

[1]*Kungliga Tekniska Högskolan, School of Computer Science and Communication, Department of Speech, Music and Hearing, Stockholm, Sweden,* [2]*Karolinska Institute, Department of Logopedics and Phoniatrics, Ear Nose Throat Cochlea Clinic, Stockholm, Sweden*

Abstract
A commercial phonetograph was complemented with a response button, such that presses resulted in marked regions in the voice range profile (VRP). This study reports the VRP data of 16 healthy female professionally trained singers (7 mezzo-sopranos and 9 sopranos). Subjects pressed the button to indicate sensations of vocal instability or reduced control during phonation. Each press thereby marked potential areas of difficulty. A method is presented to quantify the consistency of button use for repeated tasks. The pattern of button presses was significantly consistent within subjects. As expected, the singers pressed at the extremes of VRP contours as well as at register transitions. These results and the potential of the method for the assessment of vocal problems of singers are discussed.

Key words: *Evaluation tool, self-perception, singing voice, voice assessment, Voice Handicap Index, phonetogram, voice range profile*

Introduction

Computerized phonetograms or voice range profiles (VRP) are now easily accessible and are often part of standard clinical equipment. Current VRP systems often augment the range data with additional metrics that describe voice quality (e.g. crest factor, jitter, and shimmer) (1), and they provide these data not only on the bounding contours but also over the interior of the VRP. The phonetograph is an appealing tool for voice assessment (2), as it provides a summarizing voice image in which the specific interactions are depicted between an observed entity on the one hand (usually, the ability to phonate), and level and frequency on the other. In clinical settings, the VRP is commonly used as an objective acoustic measure, in combination with other subjective measures. The perceptual aspect of the clinical evaluation of voice is two-sided: as perceived by the patient and as perceived by the therapist.

Firstly, the modern and widely accepted health definition proposed by the World Health Organization (WHO) states that:

Health is a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity (3).

This shift in health definition has strongly impacted on the health care system. Clinical approaches increasingly value and measure patient self-perception and experience as an integral part of the overall evaluation process. In voice clinic environments, instruments for self-reporting such as the Voice Handicap Index (VHI) introduced in 1997, are commonly included in protocols (4). Much attention is directed to the patient's own vocal experience of the reported problem. Some researchers have attempted to adapt the VHI approach to the specific reality/concerns of the professional singer (5,6). It is recognized that there is a need to assess patient self-perceptions and also to adapt this assessment in response to the particular needs of certain groups of patients. Furthermore, a survey of the literature reveals increasing interest in the relationship of subjective assessment to other

Correspondence: Anick Lamarche (MMus). Fax: +46-8-790-7854. E-mail: anick@nada.kth.se

measures (acoustic or perceptual) common in clinical practice (7).

Secondly, external perception of voice quality is crucial to voice evaluation. Instruments such as the Grade, Roughness, Breathiness, Asthenia, Strain (GRBAS) (8) and Consensus Auditory Perceptual Evaluation–Voice (CAPE–V) (9) protocols serve to objectify and standardize the clinician's perception in evaluation (10–12). Despite some unresolved issues in this area (13,14), training perception provides satisfactory and fairly robust results (15). However, clinicians mostly train their perceptual judgment with the motivation to recognize phonatory failures in spoken voice. The voice problems of professional singers are often very specific and not always detected by mainstream voice assessment protocols, which typically are designed for speech. To singers, voice effort and pitch are well known dimensions and their voice problems typically occur at certain combinations of intensity and pitch. Therefore, the level-versus-pitch map of the VRP should be well suited for isolating problematic phonation. However, vocal problems might be so subtle that, even if they are perceived by the singer, they do not show up in the acoustic features displayed in the VRP, or indeed in any acoustic dimension.

By tapping into the singer's own perceptions, it might be possible to augment the VRP with non-acoustic but singer-relevant information. One way of doing this could be to allow the singer to signal some aspect of his/her production in a way that is automatically registered in context by the phonetograph. Such perceptual data could help bridge the gaps that exist between the singer's experience of voice production and what can be perceived or measured by the clinician.

Here, the use of a simple push-button for combining subjective immediate self-perceptual information and objective vocal measurements was investigated. Singers were asked to press the button whenever they felt that they did not have adequate control of their voice and/or when they felt discomfort. Each press of the button was registered and displayed as a black mark at the corresponding point in the VRP. Expectations for this group of singers were that button presses would generally be located at VRP extremes and would be mostly incidental or transitory. This paper describes some of the issues encountered with this approach, and how the reliability and the validity of such marker data might be assessed.

Methods

Signal acquisition

A computerized phonetograph, *Phog* (Version 2.00.10, Hitech Development AB, Sweden) was used in combination with a digital signal processing (DSP) sound card (BlueWaves LSI-PC/C32 board). The phonetograph was modified by author ST and Svante Granqvist to record also presses of an external hand-held button.

Each down-press of the button generated a 73 ms pulse, regardless of how long the button was held down or of how hard it was pressed. The duration of the button press was discarded, as subjects in development trials would sometimes hold the button pressed for a second or two—for example, over the ends of tones with large drops in sound pressure level (SPL)—resulting in irrelevant smears in the data and ambiguities in the subsequent interpretation. As illustrated in Figure 1, the binary pulses were recorded in a vacant channel, in parallel with the phonetograph's fundamental frequency (F_0), SPL, and voice quality parameters. Only button presses that were made during phonation were mapped into the VRP display, since their position would otherwise be undefined.

All recordings were conducted in a sound-treated and isolated but not anechoic room (volume 45 m^3, ceiling height 3 m, reverberation time, $T_{30} = 0.1$ s, reverberation radius >1.2 m across the spectrum, and 0.5 m deep absorbents). Singers were asked to perform in a singing stance at 30 cm mouth-to-microphone distance. For a few subjects, this distance was increased to 1 m (see Discussion) when the singer's SPL would exceed the 120 dB limit of the phonetograph; for these, the calibration of the sound file was consequently altered by a factor of 10.5 dB. A condenser microphone (Brüel & Kjaer, model 4003, Denmark) was used with a preamplifier (Brüel & Kjaer, model 2812) and a line amplifier (Nyvalla-DSP Audio Interface Box). Singers used a single earphone piece (Bassonic-Champion 4939, USA) to hear prompting tones during tasking.

The *Phog* system's criterion for detecting voicing is not a level threshold but a phonation period-time stability threshold. The running standard deviation in period-time over seven consecutive cycles is computed, and if the standard deviation is small enough, voicing is detected. This threshold was set to 0.2% or 75 cents standard deviation, given that even with a large vibrato, F_0 does not change by as much as 75 cents in seven glottal cycles. For this reason, tones with vibrato were reliably tracked. The
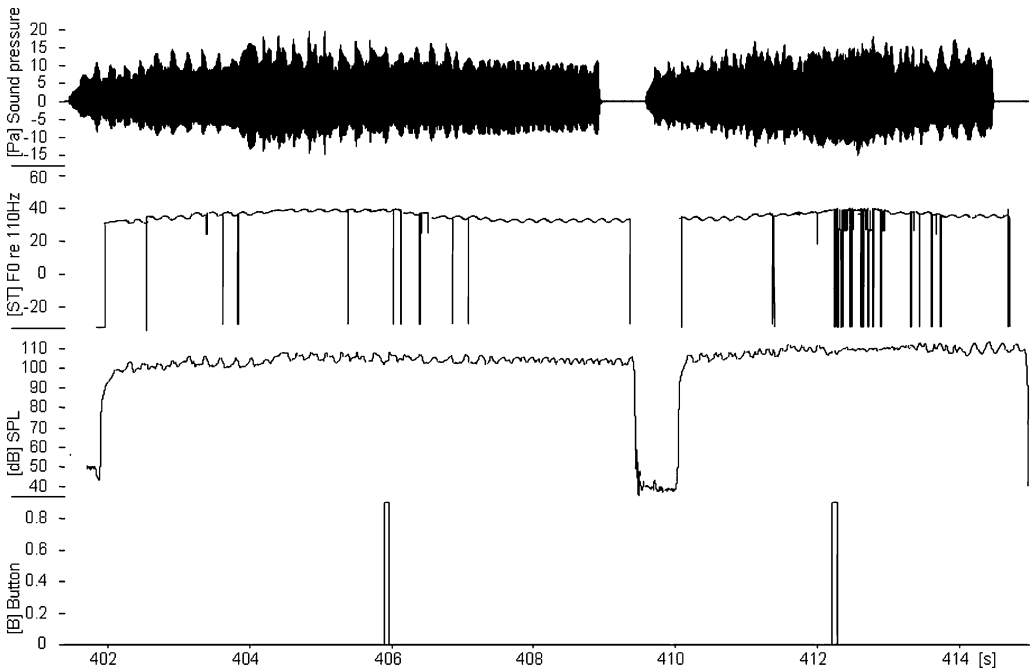
Figure 1. Signal file including, from the top, channels with audio, F_0, sound pressure level, and the button state.

phonetograph's resolution (cell size) in F_0 is one semitone, and in SPL it is one decibel. There was also a threshold for accumulated time: a cell was included in the VRP if it had been visited for a total of at least 25 milliseconds. This choice of time threshold meant that a single excursion of a vibrato cycle would be sufficiently long to be included for display and analysis.

Procedures and subjects

The subjects could communicate with the investigator by intercom, and visual contact through a window was possible. The subjects were not able to see the phonetograph display. This prevented them from being distracted by visual concurrent feedback as they performed the singing tasks and thereby enhanced an introvert locus of attention as they used the button. Singers were asked to perform the tasks on the phoneme /a:/. The three tasks of this experiment were as follows.

Task 1: A performance voice range profile was recorded. For this kind of VRP, subjects were asked to use a performance voice with their habitual vibrato at all times and to phonate as they deemed musically acceptable. The task was designed to resemble a typical vocalise, with a minor or major triad carrier. In a first step, subjects were asked to perform a *messa di voce* (a gradual rise and fall of musical dynamic on one stable frequency) on a comfortable tone in order to exercise their full performance mode dynamic range. Following this exercise, subjects sang the ascending and descending triad carrier in *pianissimo* as well as in *mezzo forte* and *fortissimo* (soft, medium, and loud) musical dynamics. Singers could break as they pleased and were given freedom in structuring their performance (phrasing, breathing, and pace). In order to test the consistency of behaviour, singers replicated Task 1 later in the procedure.

Task 2: A performance VRP was recorded for a discrete pitch exercise. A prompting pitch was played to the singer in an earphone. The singer was then asked to sing this tone in *mezzo forte*, *pianissimo* and *fortissimo* (medium, soft, and loud dynamics). The prompted intervals followed recommendations by Schutte and Seidner (16). The frequencies equivalent to the musical notes C-E-G-A in several octaves were tested across the singer's range. Again, performance voice only was required.

Task 3: Singers performed their best audition aria with lyrics. Concerning the use of the button, the singers were given the following instructions:

As you sing, press the button at any time you feel vocal instability or discomfort. Aim at communicating your sensations during your performance.

A total of 23 classical female singers were recorded; 5 subjects took part in an initial pilot phase of the experiment. All subjects had 4 years or more of professional training. Subjects 8 and 21 were excluded from analysis on the basis of vocal problems reported in a questionnaire. The analysis, excluding the pilot data, was thus conducted on the remaining 16 recordings of which there were 7 mezzo-sopranos and 9 sopranos. Finally, subject 7 needed to be excluded from the replicated task analysis due to technical difficulties. All 16 female participants were involved also in another gender-specific project, so for convenience this study reports results obtained with female singers only.

Button data validity

In this study, subjects were asked to perform two motor tasks simultaneously, one of them corresponding to a perceptual judgment of internal performance experience. It may be expected that such a combination of tasks and the request for explicit self-awareness might lead to various kinds of errors in the position of the button marks in the VRP.

Sources of errors in the button timing

Reaction time. Simply said, reaction time is the time taken between a stimulus and a movement. This time could also include a choice before the execution of the movement. Reaction time depends on nerve connections and signal pathways. Reaction times to sounds are similar to reaction times found for touch stimuli and are in the order of 140–160 ms (17,18). Despite differences in reactions to different stimuli, the time for motor preparation and response is constant for all types of reaction time tasks. Indeed, reaction time is linked to processing ('the space bar task') (19). Due to reaction delays, there is a possibility that when the subject presses the button, she is already near the end of the tone (SPL is descending rapidly) or she has already moved on to the next tone (F_0 may have changed), in which cases the mark will be erroneously placed.

While this study is not concerned with reaction time, its consequences need to be taken into account. The tasks were therefore executed at a pace that would allow the singer ample time to react. By asking the singer to phonate a minimum of 2–3 seconds per token, an average reaction time of 150 ms was accommodated into the task. A short training session prior to recording was included to decrease the singer's reaction time. This training session included either 'Ridente la calma' by Mozart or 'Somewhere over the rainbow' by Arlen. Initial phrases were sung in widely different keys in order to provoke some feeling of discomfort that would make the subject press the button.

Vibrato. The frequency and amplitude modulations incurred by the vibrato will introduce some uncertainty into the precise location of the button marks. Sources report different frequency extent values when it comes to the assessment of vibrato. On average a typical frequency swing can vary from 71 cents up to 128 cents (20,21). The level swing induced by vibrato was very variable, both from tone to tone and from singer to singer; according to this study's data observations, it could easily range from 0 to 5 dB. This seems to agree with previous reports on amplitude variation and vibrato (21). However, it is not practical to try to adapt the VRP to each singer's vibrato. The instant of pressing the button is probably quite unrelated to the vibrato cycle, so the vibrato will add a small random component to the SPL and F_0 co-ordinates of the button markers. This uncertainty needs to be borne in mind when examining the button marks in the VRP. In principle, the vibrato may be filtered out by technical means, but this would noticeably increase the response time of the device.

Post-task validation by the singers

As part of a post-recording questionnaire, singers rated how well the button markings on their VRPs reflected their performance experience and typical areas of vocal challenges. It seemed important to cross-check validity by giving the subjects themselves the chance to evaluate the instrument. Our aim was to find out if the display made sense to them with the experience still vivid in their minds. This type of post-task questionnaire included a definition of the VRP and was answered in writing. A visual presentation of the subjects' VRP for the repeated task was available to help the evaluation. When asked 'Do the button presses relate well to your own singing experience today?', visual analogue scale (a 10-cm line) ratings were very high, with 94% of the singers rating 7.5 and above.

These results suggest that, on the whole, the singers found the button marks to be consistent with their recollections of their performance. Similar results were obtained, with 91% answering in the affirmative to 'Are the highlighted portions of your VRP typical areas of vocal difficulty or/and limits?' These positive results suggest that singers viewed the

button as a possible way to communicate their perceptions, and provide some support for the validity of the button data. The use of the button was examined for each individual and was not generalized to the group.

Assessment of reliability

Previous studies report long-term and short-term subject variability in recording VRPs (22). Some degree of short-term variability in the overall VRP recordings as well as in the use of the button was therefore expected. If the new button data are to be useful, they need to be both valid and reproducible. When a subject repeats a task we expect the two sets of responses to be fairly similar. Since vocal difficulties can be transitory, especially in healthy singers, and because the singer's attention can wander, responses would never be identical. Visual inspection revealed similarities in most cases, but gave no quantification. However, if the button reports are reliable, repeating a task should give greater similarity to the first response than to the outcome of a random process with the same number of button presses. A procedure based on this requirement was devised to quantify the similarity of button presses in two VRPs, henceforth called VRP-A and VRP-B. It assessed only the similarity of the button data, not the similarity of VRPs A and B as a whole. For each recording, the *Phog* system saves both a file containing the VRP data matrix and a multitrack signal file containing the audio with a host of extracted parameters. For the custom analysis needed here, Matlab scripts were created to read the signal file and reconstruct the VRP matrix including the button data. Figure 2 depicts such a reconstruction. For reasons explained in Appendix A, a region around each button mark was constructed, as shown in Figure 3. The percentage of overlap of button regions in the VRPs A and B was then computed by simply counting coinciding and non-coinciding cells. This percentage of overlap was then used as a similarity score. For further detail concerning the four-step analysis elaborated for this study, see Appendix A.

Results

The similarity scores and *p*-values for 15 button-VRP pairs are shown in Table I. For 13 of the 15 subjects, the real button overlap percentages obtained were significantly higher than the ones obtained for the mean of 20 iterations of a randomized distribution of presses. The average similarity score was 19.3%.
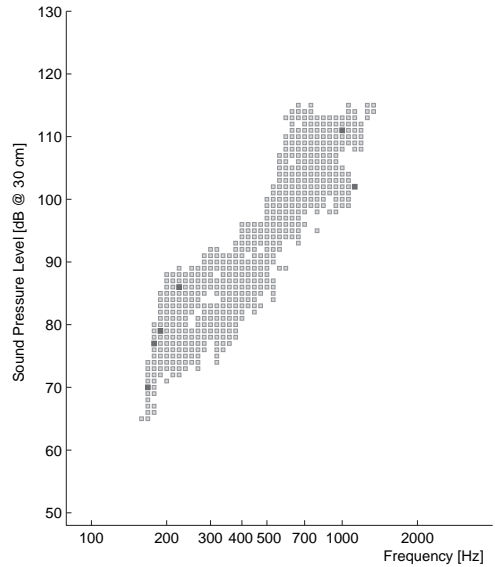

Figure 2. A voice range profile (VRP) which includes the button presses acquired during performance (darker cells).

Results for the similarity scores between Task 1 and Task 2 are given in Table II. When button presses were compared across different tasks, the
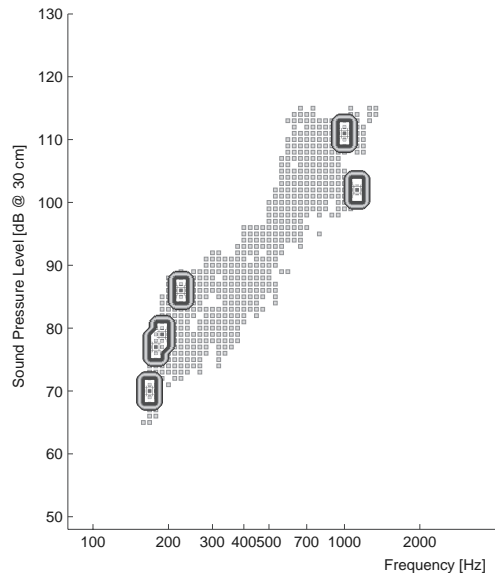

Figure 3. A Matlab reconstruction of a voice range profile (VRP) displaying button presses and button regions. As illustrated, healthy singers tend to press mostly at the boundaries of their voice: the extremes of the VRP.

Table I. Similarity of trials A and B for a repeated task (Task 1). Column 2 gives the similarity score: the percentage of overlap of button regions for trial A and B. Column 3 gives the mean overlap of 20 randomly redistributed A with B trials (an estimate of a Poisson distribution λ). Column 4 gives the probability, assuming the Poisson distribution, of the observed button overlap being an outcome of a random process. Bolded *p*-values are significant, indicating that the subject was replicating presses at higher than chance level.

Subject	Original overlap%	Random mean overlap%	*p*-value ($p < 0.05$)
6	28	14.3	**<0.001**
9	9	1.6	**<0.001**
10	10	6.8	**0.010**
11	30	14.1	**<0.001**
12	20	8.7	**<0.001**
13	9	4.8	**0.024**
14	15	5.1	**<0.001**
15	8	12.1	0.848
16	19	12.5	**0.030**
17	8	1.5	**<0.001**
18	20	7.3	**<0.001**
19	17	2.0	**<0.001**
20	35	28.9	0.110
22	8	3.6	**0.012**
23	51	8.0	**<0.001**

average similarity score dropped to 11.6%, and only 8 of 16 subjects were consistent with themselves across tasks. Real button overlap percentages were nearly always higher than the ones obtained for the mean of 20 iterations of a randomized distribution of

Table II. Similarity of different tasks (Task 1 and Task 2). Column 2 gives the similarity score: the percentage of overlap of button regions for Task 1 and Task 2. Column 3 gives the mean overlap of 20 randomly redistributed Task 1 with Task 2 (an estimate of a Poisson distribution λ). Column 4 gives the probability, assuming the Poisson distribution, of the observed button overlap being an outcome of a random process. Bolded *p*-values are significant.

Subject	Original overlap%	Random mean overlap%	*p*-value ($p < 0.05$)
6	26	16.2	**0.008**
7	15	11.0	0.092
9	0	0.7	0.489
10	23	9.4	**<0.001**
11	23	17.6	0.084
12	4	1.0	**0.004**
13	16	7.9	**0.003**
14	0	2.5	0.918
15	4	4.5	0.459
16	9	6.4	0.117
17	0	3.3	0.962
18	10	4.9	**0.012**
19	12	4.1	**<0.001**
20	23	13.2	**0.005**
22	7	4.4	0.076
23	13	1.9	**<0.001**

presses, but not all of these differences were statistically significant according to our criterion of non-randomness. An exception can be observed in Table II in the case of subjects 9, 14, and 17 where the similarity percentage reported was 0 and where the mean of a random distribution of presses held a higher percentage. In these cases, very few presses of the button were registered for Task 2, hence limiting the chances of an actual overlap with Task 1. When the presses from Task 1 were distributed randomly 20 times, possibilities for overlap became higher and, in turn, the *p*-values reported were amongst the highest in this analysis.

Discussion

With a microphone distance of 1 m, room acoustic effects can be of concern even in a heavily treated room. Because such a room (with isolated walls and ceiling) has a single dominating floor reflection, a random variation in SPL of about ± 0.8 dB at 30 cm and ± 2 dB at 100 cm can be predicted. This was closely confirmed by comparing the SPL-time contours in the *Phog* signal files, for real sessions reproduced at 30 and 100 cm using a high-quality loud-speaker in place of the singer. However, this random source of variation, although undesirable, can only make the similarity scores lower and not higher. Hence the similarity tests reported here are slightly more conservative than what would have been the case with a true anechoic room and a truly fixed microphone distance.

The validity of the button as a new device was investigated by testing the consistency of button presses for a singer. Overall, the similarity scores confirmed the subjective visual impression that the button information was not random, but was repeatable and therefore can be assumed to reflect actual difficulties experienced by the singers. Results for this group of subjects are encouraging as they attest to the applicability of the button-mediated responses as a new metric. For the replicated task, the singers demonstrated a significantly consistent behaviour in the use of the button. Expectations formulated at the onset of the experiment were met: button presses were in general located at VRP extremes, and, according to the singers' informal reports, presses were mostly related to momentary vocal difficulties (unprepared onset, phlegm on the folds, vocal limits, etc.). Because the subjects in this group were vocally healthy, yet asked to communicate problems during performance, button presses might be expected to be infrequent and/or range-specific. For a group of injured singers, one might expect the similarity of replicated tasks to be higher and more problem-specific.

When assessed across tasks, the similarity scores were lower but continued to support consistent behaviour for half of the subject group. There were no expectations that the button presses would be particularly reproducible across tasks, but the above-random effect for a part of the group is nonetheless worth mentioning. The reduction in similarity scores is not surprising since the performance VRP's overall shape for each task would be different. It follows that areas highlighted by button presses for one task might not even be registered in a different task. Consequently, if this method were used, say, for assessing pre-and post-therapy, the task would have to be the same, pre and post.

All in all, this new method could have clinical potential to document the performance experiences of singers. An advantage of this method is that the perceptual judgment is instantaneous and most likely intrinsically related to the experience of the moment (23). If the aim is to evaluate the singing voice and understand its failures in relation to stage performance and injury, then it is of interest to identify the phonatory conditions (pitch and effort level) that invoke a problem. For the singer who has trained his/her kinaesthetic sensitivity and is a vocal athlete, the instants of vocal problems caused by injury are very specific. The button, in this case could allow the depiction of those problematic areas and possibly assist effectively the evaluation and/or even the rehabilitation process.

It may be noted that, in soprano singing, fundamental frequency tuning to the first formant (F_0–F_1) practically eliminates sound pressure level (SPL) variability across vowels (24). Nonetheless, the phoneme /a:/ (and its variations) was used for all tasks with the exception of the aria excerpt. This decision was taken for the sake of comparability with other studies in the literature and to follow recommendations by Gramming and Sundberg (25).

Conclusion and future work

A VRP was augmented with a button to tap into the singer's perceptions as he/she performs. In using this button, singers met initial expectations that healthy singers would use the button to communicate transitory difficulties and press mostly at the extremes of their performance range. An attempt to quantify the reproducibility demonstrated that singers reproduced their use of the button at least more closely than would a random process. It seems reasonable to conclude that singers, to some degree, can communicate their perceptions with a button as they perform.

Our intention is to test the use of the button-augmented VRP with injured singing voices, as an integral part of the voice evaluation process. It remains to be seen if the use of the button in this case becomes more consistent and problem-specific. It appears that the combination of subjective self-perception and the objective VRP has the potential to offer a new layer of understanding of singing voice, in the research laboratory, the clinic, and the singing studio.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

1. Pabon JPH. Objective acoustic voice-quality parameters in the computer phonetogram. J Voice. 1991;5:203–16.
2. Holmberg EB, Ihre E, Södersten M. Phonetograms as a tool in the voice clinic: Changes across voice therapy for patients with vocal fatigue. Logoped Phoniatr Vocol. 2007;32:113–27.
3. World Health Organization. International Classification of Impairments, Disabilities and Handicaps. Geneva: World Health Organization; 1980.
4. Jacobson BH, Johnson A, Grywalski C, Silbergleit A, Jacobson G, Benninger MS, et al. The Voice Handicap Index (VHI): development and validation. Am J Speech Lang Pathol. 1997;6:66–70.
5. Morsomme D, Simon CH, Jamart J, Remacle M, Verduyckt I. A proposal to adapt the voice handicap index to the singing voice]. Rev Laryngol Otol Rhinol (Bord). 2005;126:305–13.
6. Rosen CA, Murry T. Voice handicap index in singers. J Voice. 2000;14:370–7.
7. Lehto L, Laaksonen L, Vilkman E, Alku P. Occupational voice complaints and objective acoustic measurements—do they correlate? Logoped Phoniatr Vocol. 2006;31:147–52.
8. Hammarberg B, Gauffin J. Perceptual and Acoustic Characteristics of Quality Differences in Pathological Voices as Related to Physiological Aspects. In: Fujimura O, Hirano M, editors. Vocal Fold Physiology, Voice Quality Control. San Diego: Singular Publishing Group; 1995. p. 283–303.
9. American Speech-Language-Hearing Association. Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V): ASHA Special Interest Division 3, Voice and Voice Disorders. 2002–2006. Available from: http://www.asha.org/NR/rdonlyres/C6 E5F616-972F-445A-AA40-7936BB49FCE3/ 0/CAPEVpro cedures.pdf (accessed May 2008).
10. Oates JM, Bain B, Davis P, Chapman J, Kenny D. Development of an auditory-perceptual rating instrument for the operatic singing voice. J Voice. 2006;20:71–81.
11. Kenny DT, Mitchell HF. Acoustic and perceptual appraisal of vocal gestures in the female classical voice. J Voice. 2006;20: 55–70.
12. Gelfer MP. Perceptual attributes of voice: Development and use of rating scales. J Voice. 1988;2:320–36.
13. Kreiman J, Gerratt BR, Kempster GB, Erman A, Berke GS. Perceptual evaluation of voice quality: review, tutorial and a framework for future research. J Speech Hear Res. 1996;36: 21–40.
14. Shrivastav R, Sapienza CM, Nandur V. Application of psychometric theory to the measurement of voice quality using rating scales. J Speech Lang Hear Res. 2005;48:23–35.
15. Hammarberg B, Fritzell B, Gauffin J, Sundberg J, Wedin L. Perceptual and acoustic correlates of abnormal voice qualities. Acta Otolaryngol. 1980;90:441–51.

16. Schutte HK, Seidner W. Recommendation by the Union of European Phoniatricians (UEP): standardizing voice area measurements/phonetography. Folia Phoniatr. 1994;35: 286–8.

17. Galton F. On instruments for (1) testing perception of differences of tint and for (2) determining reaction time. J Royal Anthr Inst. 1899;19:27–9.

18. Brebner JT, Welford AT. Introduction: an historical background sketch. In: Welford AT, editor. Reaction Times. New York: Academic Press; 1980. p. 1–23.

19. Miller JO, Low K. Motor processes in simple, go/no-go, and choice reaction time tasks: a psychophysiological analysis. J Exp Psychol Hum Percept Perform. 2001;27:266.

20. Prame E. Vibrato extent and intonation in professional Western lyric singing. J Acoust Soc Am. 1997;102:616–21.

21. Nawka T, Seidner W, Cebulla M. Experiments on intentionally influencing the singer's vibrato. In: Friberg A, Iwarsson J, Jansson E, Sundberg J, editors. SMAC93. Stockholm: Stockholm Music Acoustics Conference, Royal Swedish Academy of Music; 1993. p. 221–7.

22. Gramming P, Sundberg J, Akerlund L. Variability of phonetograms. Folia Phoniatr (Basel). 1991;43:79–92.

23. Darrow AA. Examining the validity of self-report: middle-level singers' ability to predict and assess their sight-singing skills. International Journal of Music Education. 2006;24:21–9.

24. Lamarche A, Ternström S. An Exploration of Skin Acceleration Level as a Measure of Phonatory Function in Singing. J Voice. 2008;22:10–22.

25. Gramming P, Sundberg J. Spectrum factors relevant to phonetogram measurement. J Acoust Soc Am. 1988;83: 2352–60.

Appendix A: *Details on the statistical approach*

Four main steps define the analysis.

1) *Button region definition.* To account for proximity without actual overlap, and also for vibrato-induced variations as described in the article (see Methods), a surrounding region was 'bled' around the co-ordinates of every button press. Based on previous reports and on observations of vibrato behaviour of the recorded singers, the region was chosen to range ± 2 dB in height and ± 1 semitone in width. As seen in Figure 3, each cell marked by button presses becomes the centre of a larger overlap rectangle of 5×3 cells. This region is somewhat analogous to a proximity weighting function as used in image correlation calculations.

2) *Overlap calculation.* Task 1 was replicated in broken practice style which meant that for each subject, there were two performance voice range profiles (VRPs) of Task 1 with button information available for comparison. The button regions in the VRP-A were overlapped with the button regions in VRP-B, and the percentage of total button region overlap was calculated.

3) *Randomization of the original button region obtained in A.* A high percentage of overlap is not in itself a good measure of similarity, since the degree of random overlap will be higher if the button marks are dense rather than sparse. Rather, we needed to know if the observed overlap in each case was higher than would be expected by chance. Therefore in a third step, the button regions in VRP-A were uniformly repositioned within the total VRP-A area at random, and the overlap with B was recalculated. This randomization was iterated 20 times for each pair of VRPs A and B. The average overlap and the standard deviation for the 20 iterations per subject were calculated in a final step.

4) *Cumulative distribution function.* For small to moderate amounts of overlap, the distribution of overlap outcomes of repeated random trials can be modelled by a Poisson distribution. The Poisson distribution is a discrete probability distribution returning only values greater than or equal to zero. In this context, the cell overlap is discrete: it happens in an integer number of cells. The parameter lambda (λ) is equivalent to the mean of the Poisson distribution. For each A-B pair of button maps, the average percentage overlap of the 20 random iterations was used as an estimate of the λ parameter of the Poisson distribution for that pair. Using the cumulative Poisson distribution function, the probability was calculated that the random overlap percentage for each subject was less than or equal to the real button overlap percentage. Significance was defined as alpha $=0.05$ for p, where p is the probability that the observed percentage of button overlap could have been the outcome of the simulated random process.

Similarity scores were computed also for *different* tasks, comparing button presses from Task 1 to those from Task 2.

Paper IV

Paper IV is in revision following a review

in *J Speech Lang Hear Res*

The layout is customised.

Not Just Sound II: an Investigation of Singer patient Self-Perceptions Mapped into the Voice Range Profile.

ANICK LAMARCHE[a], DOMINIQUE MORSOMME[b] STEN TERNSTRÖM [a]

[a] Kungliga Tekniska Högskolan, School of Computer Science and Communication, Department of Speech, Music and Hearing, [b] Université Catholique de Louvain

Abstract

Purpose: In aiming at higher specificity in clinical evaluations of the singing voice, singer perceptions were included and tested in conjunction with the voice range profile. **Method:** The use of a commercial phonetograph supplemented by a hand-held response button was clinically tested with 13 subjects presenting voice complaints. Singer patients were asked to press a button to indicate sensations of vocal discomfort or instability during phonation. Each press was registered at the actual position in the Voice Range Profile (VRP) so as to mark areas of difficulty. Consistency of button press behavior was assessed with a method developed previously. **Results:** In spite of their voice complaints, subjects did not press the button as much as healthy singers. Like healthy singers, the singer-patient group demonstrated consistent behavior but tended to press the button in completely different areas of the VRP space. The location of the presses was dominantly in the interior of the VRP and concentrated to a small fundamental frequency range. An extensive discussion examines carefully the reasons for such outcomes. **Conclusion:** The button augmented VRP could be a well needed resource for clinicians but requires further development and work.

Introduction

New perspectives on health definitions include both a growing awareness of the importance of the patient's self-perception of his/her problem, and the knowledge that a patient's treatment is very individually based. More and more effort is spent towards tailoring the clinical evaluation process to patient needs. Singers must meet high vocal demands, and have often high priority patient status in vocal clinics. They form a good example of a patient group with very specific needs. Several reports conclude that singers are at higher risk for voice disability and are prone to be impacted by these problems in a different way than are non-singers (Cohen, 2007; Morsomme, 2005; Phyland, 1999; Rosen, 2000).

Yet, there seems to be a discrepancy between what the literature reports on the one hand, and clinical approaches and practice concerning singer patients on the other. For example, it is only recently that voice-related quality-of-life instruments such as the Voice Handicap Index (VHI) have

been modified and adapted to meet the needs of singers (Cohen, 2007; Lamarche et al., in review; Morsomme, 2007; Murry, 2008). This discrepancy is perhaps related to the extent of resources available to clinicians. Clinicians are highly trained in recognizing, identifying and remediating phonatory failures in the spoken voice but might not be as well equipped in addressing the singing voice. Furthermore, since the voice problems of accomplished singers are often very specific and subtle, they might remain undetected by mainstream voice assessment protocols designed for speech (KayPENTAX, 2008). The Voice Range Profile (VRP) is often referred to as a tool that might be more sensitive to the subtleties of the singing voice, since it provides a map extending over the full vocal range. Thus the VRP, in comparison to isolated measures, is more likely to detect problems that occur only at certain effort levels and/or phonation frequencies.

In this light, a phonetograph augmented with a push-button was developed and tested with healthy singers (Lamarche, Ternström and Hertegård, 2008). This device superimposes subjective immediate self-perceptual information on the objective vocal measurements. By combining the self-perception of the singer during performance with a common clinical assessment, the VRP, the particular needs of the singer patient could be examined more closely. In a previous study (Lamarche *et al*. 2008), it was found that a singer's own perceptions could be used to produce a voice map containing acoustic as well as non-acoustic singer-relevant information. One outcome of this study was that healthy classical singers were consistent in using the button device. This led to the current question: will singer patients demonstrate more consistent button pressing than that found for healthy singers? Furthermore, will a particular pattern of button-device use emerge for this group or eventually even diagnosis groups? Indeed, it was expected that singer patients would press the button in the inner VRP areas rather than at VRP contour extreme limits as in the case of healthy singers.

The motivation for this paper was thus to examine how singer patients would use the button in the context of a VRP recording; and to explore how such VRP might further equip the clinician for the assessment of the singing voice status and facilitate the clinical evaluation process.

Methods

Signal Acquisition

A digital sound processing card (CAC Bullet II DSP) was used to run the computerized phonetograph, *Phog* (Version 2.00.10, Hitech Development AB, Sweden). The phonetograph was modified by author ST and Svante Granqvist to record simultaneously voice and presses of an external hand-held button.

The activation of the button device yielded pulses of 73 ms and this, regardless of how long the button was held down or of how hard it was pressed. The binary pulses were recorded in a vacant channel, in parallel with the phonetograph's fundamental frequency (F0), SPL, and voice quality parameters. Only button presses that were made during phonation were mapped into the VRP display, since their position would otherwise be undefined. For further detail, the reader is referred to Lamarche *et al*., 2008.

Recordings were conducted in a clinical environment. This imposed the choice of a headset cardioid microphone in order to reduce the influence of environmental, background noise and room reflections as well as allow more freedom to the singer while minimizing mouth to microphone distance changes (Cabrera, 2002; Lamarche *et al*., 2008). Prior to the recordings, a series of tests was performed on different kinds of headsets. It was found that the cardioid microphone recommended by the manufacturer for *Phog* and used by KayPentax (AKG model 420 headset) met the maximum level requirements (120-129 dB) for making VRPs of female classical singing. A low-noise microphone preamplifier was used (Line Audio Design model 2MP, Rinkaby, Sweden).

The phonetograph voicing thresholds were set to minimum 0.025 seconds for the accumulated time per cell, and maximum 75 cents for the F0-standard deviation over seven periods. The cell aspect ratio was 2/3 and the sampling rate was 16000 Hz. The calibration of SPL was performed for every subject. Microphone-to-mouth distance was measured from the front teeth to the boom and at this distance, a white noise calibration tone was played and measured at the microphone with a sound pressure meter (LA-210, Ono Sokki, Japan) using C-frequency-weighting. The mouth-to-microphone distance was then compensated for a distance of 30 cm using *Phog's* calibration settings.

Procedures

The investigator and the equipment were in the same room as the subject. By design, the patients could not see the computer screen. Subjects were asked to perform the tasks in a singing stance. Subjects used the phoneme /a:/ across all tasks. They were asked to use the button during each task to communicate feelings of discomfort or a loss of vocal control. Throughout the recording session, singers could break as they pleased, and were given freedom in structuring their performance (phrasing, breathing and pace). The following instructions were given to the patient-singers to guide their use of the button.

"As you sing, press the button whenever you feel vocal instability or discomfort.

Aim at communicating your sensations during your performance."

These instructions were formulated so as to encourage the singer to focus on the effects of phonation rather than on the audio feedback and voice quality aspects. The task protocol was as follows.

Task 1

The subject was asked to make a description of their personal warm-up routine in a spontaneous speech task. Such a theme was considered to be neutral in content and easily accessible to all subjects. This task also included a counting exercise in which the subject used soft, comfortable speech as well as loud public speaking voice. Subjects spoke in their native tongue (either French or Dutch). The total duration of this task was 3 minutes. Task 1 had mainly a training nature where subjects could acquaint themselves with the instructions and the task at hand.

Task 2

A physiological VRP was performed. In this task, voice quality was completely disregarded and the singer was encouraged simply to phonate as softly and as loudly as possible. A descending *glissando* (a slow frequency sweep) followed by an ascending *glissando* exercise was used to obtain the softest and loudest possible phonation across the subject's range. Total task duration was approximately 6-8 minutes.

Task 3

A performance VRP was recorded for a discrete pitch exercise. A performance VRP entails the performance aspects of voice such as vibrato, relative stage vocal dynamics, musicality etc. A prompting pitch was played to the singer. The singer was then asked to sing this pitch in a *messa di voce* exercise (an increasing-decreasing tone on a stable pitch). Prompted intervals equivalent to the musical notes C-E-G-A were tested across the singer's range. The duration of the task was 6 minutes. This task was central to the experiment and importantly addressed the singing voice as it is typically used by the artist. Since singer patient's are inclined to have vocal complaints related to their singing voice, it was most interesting to assess how they would use the button device during performance.

Task 4

An excerpt from a typical audition piece was recorded. This task was performed only with the approval of the speech language pathologist (SLP). The task duration was 1 minute. The performance of a song served to break-up the repetition of Task 3, and to assess if button pressing behavior would change with increased performance context.

Task 5

 The final task partly replicated Task 3. The regions previously marked by the button presses in Task 3 were used to select the parts of the task to be repeated. Subjects were oblivious to this decision. This procedure was chosen to minimize the use of the subject's voice. We felt it could be unethical with patients to subject them to additional loading in VRP areas where few or no button-press replications were to be expected. The task lasted 5 minutes.

Subjects

13 singers, 9 females and 4 males with a formal voice complaint were recorded. The group presented a variety of diagnoses: 6 cases of nodules, 5 of functional dysphonia, 1 of pharyngo laryngeal reflux and 1 complaint without diagnosis. All subjects had had vocal training. Singing levels and genres, however, were quite varied, extending from professional classical soloists and opera choristers to amateur gospel and ballad singers. Patients were recruited in a Belgian clinic. All subjects participating in this study were at the starting point of rehabilitation and had no prior knowledge or experience of

voice therapy. Data was collected at author DM's clinic, the Clinique Saint Luc, Centre d'Audio-Phonologie Saint Luc. Table 1 a) and b) presents a summary of the group demographics. An ethical vetting certificate was obtained from the *Regionala etikprövningsnämnden i Stockholm* (certificate 1358-31) and subjects all signed a consent form prior to the recording session. The VRP recording session was offered as a complement to regular evaluation.

Patients	Age	Sex	Singing Level	Diagnosis
1	34	F	Amateur gospel	Nodules
2	24	F	Amateur-pop	Nodules
3	23	F	Student pop-ballads	Nodules and Edema
4	24	F	Student classical	Fibrous Nodules
5	26	F	Professional classical soloist	Functional Dysphonia III
6	36	F	Professional opera chorister	Nodules
7	51	F	Amateur baroque chorister	Functional Dysphonia III
8	-	F	Amateur-jazz	Nodules
9	27	F	Amateur-pop	Functional Dysphonia
10	28	M	Professional opera chorister	Functional Dysphonia III
11	31	M	Amateur-ballads	Functional Dysphonia I
12	-	M	Amateur-classical	Pharyngo-Laryngeal Reflux
13	16	M	Student-classical	Vocal complaint without diagnosis

Table 1a) Singer-patient group data reporting age, sex, singing level and diagnosis (N=13). Functional dysphonia is reported accorded to the level scheme elaborated by Koufman (1982). Age data is missing for patients 8 and 12.

Patient	VHI-singer Total/120, F-E-P/40	DSI -5 to 5	Vital Capacity [l]	Psub [cm H₂0] Speech	Psub [cm H₂0] Singing	Max. Phonation [sec]
1	46, 12-18-16	-5.2	2.7	5	.	11
2	63, 15-20-28	1.5	2.7	14.4	16.1	14
3	20, 5-5-10	1.0	2.4	9.3	11.0	11
4	42, 8-15-19	3.5	2.1	10.6	20.0	9
5	38, 15-3-19	5.8	3.5	8.1	12.8	21
6	65, 19-24-22	3.8	3.5	12	14.0	17
7	54, 16-16-23	4.9	3.0	8.5	11.4	17
8	64, 16-27-21	-0.4	2.5	9.1	13.9	11
9	.	.	.	.	.	.
10	53, 16-20-17	-0.3	3.5	7.4	8.3	13
11	53, 17-19-17	-0.5	3.0	5.8	10.0	7
12	61, 20-25-16	.	.	.	.	.
13	.	.	.	.	.	.

Table 1b) Voice status evaluation parameters for singer patients (N=13). The Voice Handicap Index adapted for singers is reported for the total score and the functional (F), emotional (E) and physical (P) subscales. Scores for the Dysphonia Severity Index (DSI) are also included. Unfortunately, information for patients 9, 12 and 13 was not available.

Post-task validation by the subjects

A post recording questionnaire (Appendix A) was distributed to all the participants involved in the project. Subjects graded with visual analogue scales (VAS) their current vocal status as well as their impressions of their use of the button in performance.[1] In addition, subjects described their motivations for pressing the button. The descriptive portion of subject answers was categorized and the frequency distribution of key words was tabulated. The questionnaire contained a total of nine questions. This type of post-task questionnaire was accompanied by a definition of the VRP and was answered in writing.

---

[1] The VAS consists of an horizontal 100 mm line which is accompanied by binary pole anchors.

During this evaluation, the subject could visually refer to the outcome of Task 3 and 5. The purpose of the questionnaire was to verify instruction comprehension and also carefully document the patient's experience of using the button.

Analysis

With the same instructions as above mentioned, previous work (Lamarche *et al*., 2008) had shown that healthy singers used the button consistently. These results motivated the formulation of one of the current study's expectations: since singer-patients suffer from a vocal problem, discomfort and the loss of control can be expected to occur more systematically and thus lead to a more consistent button-pressing behavior than that observed in healthy singers. In order to test this, the analysis protocol of the first study was somewhat modified.

Similarity

The analysis technique developed in previous work (Lamarche *et al.,* 2008) was repeated here. Similarity percentage scores were calculated to assess the consistency of the subject's behavior when pressing the button. Figure 1 gives an example of these first analysis steps taken in the previous study. Button presses were mapped to the VRP and each press was attributed a surrounding region. The overlap of the regions between Task 3 and Task 5 was determined and the degree of overlap defined the similarity score (this was labelled "original overlap" in reference to the actual collected button presses). The original number of presses was then redistributed randomly over the VRP area of Task 3. The degree of overlap was reassessed between the randomized button presses and the presses obtained in Task 5 (this was labelled as the "random overlap"). The probability that the original overlap occurred at higher than chance level was tested against a Poisson distribution (a discrete probability distribution that gives the probability of a number of events occurring in a fixed period of time, or, in this case, a fixed area) with an alpha of 0.05.

Specificity

Not only was it important to assess how individuals were consistent in using the button but equally how often they pressed. Since presumably a vocal disorder would provoke a systematic problem, it was hypothesised that patients would make greater use of the button than healthy singers. The distribution of the rate of pressing and the specificity of the information were examined in terms of repetition detection and color mapping to enable visualization of this information. The rate of button presses was compiled by retrieving the occurrence of the button-device pulses synchronised with the F_0 and SPL channels. This was then displayed on the VRP by mapping the rate to a set of 3 discrete color variations on the display. The display varied from the lightest shade of grey for a low rate to a darkest shade of grey for the highest rate.

Figure 2 illustrates how overlapping button presses and regions are identified by darker shades of gray.

Pooling

Finally, as it was important to assess the differences between the healthy group and the subjects, accumulated button presses were pooled into one Figure. Total press data was accumulated for the healthy group, for the overall subject group and for the two main diagnosis subgroups: nodules and functional dysphonia.Both males and females were included in the accumulated data since the aim was to examine the button pressing and the possible overall location trends. Naturally, in a VRP assessment, sexes would have to be grouped separately for a proper evaluation.

Questionnaire

Figure 3 and Table 2 a) and b) show the questionnaire results. The questionnaire itself is reproduced in Appendix A. The questionnaire responses generally confirmed our observations with the exception of results obtained for question 4. This will be further addressed in the discussion. The central item, a question in which the subjects were asked to rate the correspondence of the button press display to their singing experience, the average VAS ratings were 75 % of the total line length (standard deviation was 4.6). Figure 3 demonstrates the mean and standard deviations for all 5 questions answered with VAS ratings. In general, the singers found the button markings to be consistent with their recollections of their performance and demonstrated good comprehension of the task instructions. Table 2 tabulates the collected qualitative material in terms of a) the type of difficulty experienced in the recording, and the definition and description of the main reason for pressing the button device; and b) the effect of vocal effort on performance. Because responses for question 7 of the questionnaire were scarce, this question was excluded from qualitative reports.

Results

Overall, the similarity results attested to the group's consistent behavior. In a second elicitation of the task, the button was pressed at higher levels than chance. The similarity scores and p-values for thirteen button-VRP pairs are shown in Table 3. For 8 of the 13 subjects, 62 % of the group, the original button overlap percentages obtained were significantly higher than the ones obtained for the mean of 20 iterations of a randomized distribution of presses. Significance was determined with an alpha of 0.05. The average similarity score for the subject group was 10.9 % while in a previous study, results for a healthy group of female singers were 19.3 %. The data from this previous study (Lamarche *et al.*, 2008) are included in Table 4 to facilitate comparison. The healthy group of singers used here for comparison, included 16 healthy female opera singers, 19 to 35 years of age and with a minimum of 4 years of training,

Original button overlap percentages were often higher than the ones obtained for the mean random overlap, but with our criterion for non-randomness, not all of these differences were statistically significant. This applied for both healthy and patient groups. Exceptions could be noted for subjects 3 and 4 where the similarity percentages were lower (0 and 1 respectively) than the mean percentage
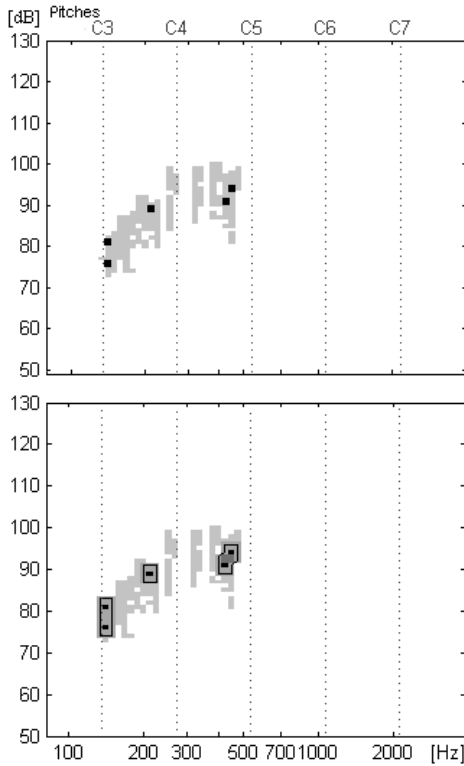
Figure 1. The performance VRP for one subject with the button presses (top) and including the zones mapped to each press (bottom). This VRP was obtained for patient 1, an amateur singer, singing mostly gospel and diagnosed with nodules. The vertical axis displays dB SPL @ 30 cm and the horizontal axis,logarithmic frequency in Hz.
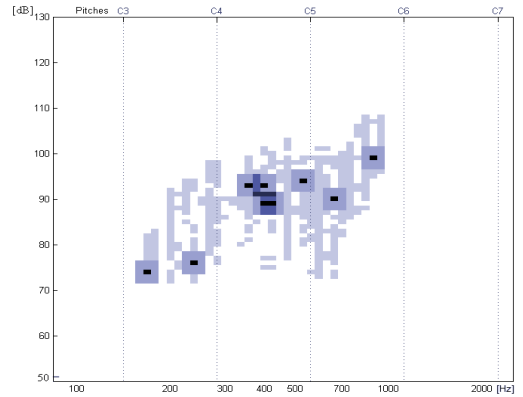

Figure 2. The merged VRPs for both Task 3 and 5, patient 4. The occurrence of button press zone overlap is depicted by darkening colors. In this way, the visualization of the button press overlap can help denote a particular area of concern. For this classical soprano diagnosed with nodules, the central region of the VRP seems especially problematic. Axes are defined as in Figure 1.
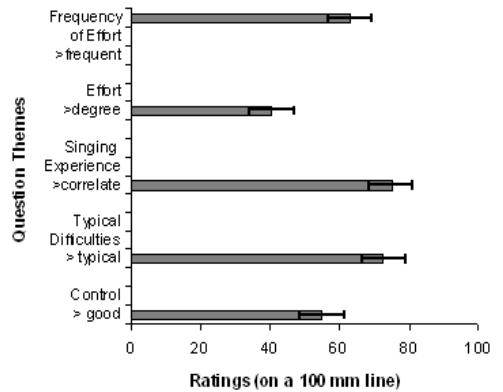
Figure 3. Post recording results of 5 questions as rated on a Visual Analogue Scale. Appendix A gives the details of the full questionnaire. Rating means and standard deviation are included for each theme. The main theme for each question is used as a category on the horizontal-axis and the semantic anchor that corresponded with the high extreme of the scale (in this case 100) is also included.

Thematic Categorization	Frequency of Answer	Description of the vocal difficulty
Functional	1	sudden interruption of phonation
	2	intermittent phonation
	1	voice transition zones management
Perceptual Auditive Qualities	1	clarity of the voice
	2	intonation precision
	1	the height of the sound
	1	mucus-or sensation of a veil
	1	lack of sonority in the soft voice
	1	less vocal power
Physical	1	increased pressure
	1	impacted respiration
	1	jaw tension
	1	tickling sensation in the throat
		Main reason for button pressing
Functional	5	lack of control
	3	no dynamic range flexibility in high tones
	1	from one register to another
	1	top notes
	1	hesitant onset for particular tones
Perceptual Auditive Qualities	1	cracking voice
Physical	2	forcing
	1	when feeling exceeding pressure
	1	improper breath support
	1	jaw compensation
	1	breath is insufficient
	2	larynx tension and height
Other	1	"…is just the thing that you cannot find words to explain when you have nodules"

Table 2 a) Singer-patient descriptions of systematic difficulties relevant to the button pressing and descriptions of the main reason for pressing the button. Descriptions are grouped into thematic categories. Frequency of appearance for each statement is tabulated.

Thematic Categorization	Frequency of Answer	Vocal Effort impact on Performance
Emotional	2	provokes lack of self-confidence
	2	fear that the sound will not happen
	4	need more concentration, stress is higher
	1	focus is on technique rather than performance
Functional	2	no vocal endurance
	1	Leave out certain high notes
Physical	1	mouth position differences

Table 2 b) Singer-patient responses concerning the impact of vocal effort on their performances. Statements are categorized per theme and frequency of appearance for each statement is included.

obtained for the random distribution of presses. Some exceptions had also previously been noted for subjects 9, 14 and 17 of the healthy group. These exceptions are associated with instances of insufficient sampling, i.e., where very few presses of the button were registered for either one of the tasks, thus increasing the risk of no overlap at all with the other task elicitation. When those limited presses were distributed randomly 20 times within the VRP space, the probability estimated as the mean overlap often became higher and therefore yielded the highest *p*-values for these cases.

Rate of Button Press

Healthy singers clearly had recourse to the button device more often than did singer patients. The bar graphs in Figure 4 a) and b) depict this unexpected trend. The overlap of button presses within a same task is depicted for the patient group in Figure 5 as well as parallel information retrieved from the previously studied healthy group (Lamarche *et al*., 2008). For both groups, results did not show any

signs of a warm-up effect; button press rates varied similarly in both elicitations and Task 3 and its replication, Task 5 yielded very similar VRP areas.

Pooling

Figure 6 shows the accumulated total of button presses of the 6 patients with nodules (all females) and Figure 7 displays the accumulated total of presses of the 5 patients (males and females) diagnosed with functional dysphonia. In a last step, button press information for Task 3 and 5 was merged to create a total density plot for the subjects and the healthy singers respectively. This type of plot is instrumental in depicting trends in a group's overall button press behavior. Figure 8 illustrates this information.

Subject	Original Overlap %	Mean Overlap %	p-values (α=0.05)
1	29	11.22	**<0.001**
2	9	7.66	0.242
3	0	2.03	0.869
4	1	5.60	0.975
5	7	0.31	**<0.001**
6	14	7.45	**0.010**
7	13	4.08	**<0.001**
8	0	0.0	------
9	5	1.51	**0.005**
10	24	6.29	**<0.001**
11	19	1.96	**<0.001**
12	7	5.66	0.211
13	14	0.47	**<0.001**

Table 3 Similarity results of trials for Task 3 and 5. Column 2 gives the "similarity score", the percentage of overlap of button regions for Task 3 and 5. The mean overlap with Task 5 from 20 iterations of presses from Task 3 randomly redistributed within the Task 3 contour (an estimate of a Poisson distribution parameter λ) is given in column 3. Column 4 gives the probability, assuming the Poisson distribution, of the observed button overlap being an outcome of a random process. Bolded *p*-values are significant at an alpha level of 0.05.

Subject	Original Overlap %	Rand Mean Overlap %	p-value (α=0.05)
1	28	14.3	**<0.001**
2	9	1.6	**<0.001**
3	10	6.8	**0.010**
4	30	14.1	**<0.001**
5	20	8.7	**<0.001**
6	9	4.8	**0.024**
7	15	5.1	**<0.001**
8	8	12.1	0.848
9	19	12.5	**0.030**
10	8	1.5	**<0.001**
11	20	7.3	**<0.001**
12	17	2.0	**<0.001**
13	35	28.9	0.110
14	8	3.6	**0.012**
15	51	8.0	**<0.001**

Table 4 The same analysis as in Table 3, but performed for a group of 15 healthy subjects. Bolded *p*-values are significant at an alpha level of 0.05.
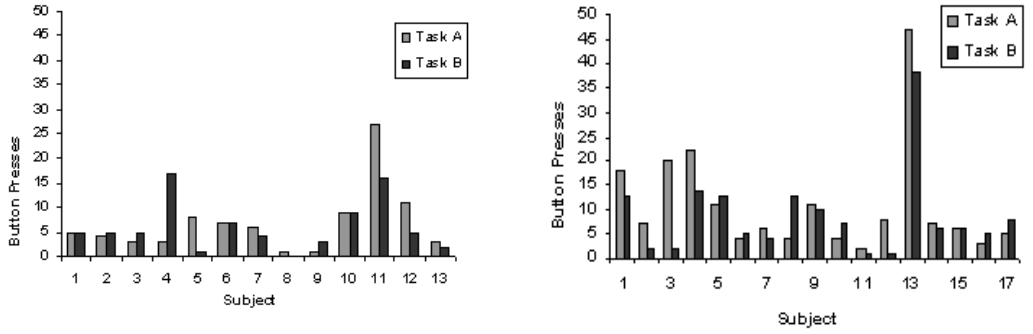
Figure 4a) Button press occurrences for singer patients in Task 3 and Task 5. **b)**. Results of button press occurrences for Task 3 and Task 5 for healthy singers.
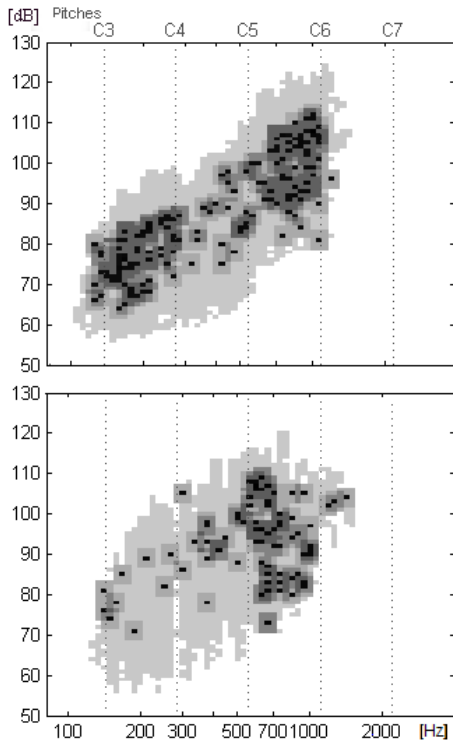

Figure 5. The accumulated button presses for a group of 15 female healthy singers (top) compared to a singer-patient group (bottom) of 9 females and 4 males. This averaging is performed only for Task 3 results. The darker the gray, the greater the button zone overlap. Axes as for Figure1.

Discussion

At the onset of this study it was expected that subjects would press more often, more consistently and depict different button pressing patterns than healthy singers. For the subject group, the locations of button presses within the VRP were expected to be concentrated in inner VRP areas as the fine motor control required in the messa di voce would be likely to pose problems in the presence of vocal disorder. In this way, the button presses could possibly elucidate new inner VRP areas of interest in concern to singer-patient voice assessment. Finally, it was also deemed interesting to examine the use of the button device and the location of press trends in relation to diagnosis groups; in this case nodules and functional dysphonia. However, due to the limited number of subjects, reliable and valid results could not be obtained.

Results partly differed from our expectations. The button press rate for singer patients was lower than that found for healthy singers.Furthermore, although conclusions of consistent behavior could be drawn for the singer-patient group, results were weaker than those found for the healthy group. Indeed, when tested across tasks, healthy singers had had comparable similarity scores to what was observed for the singer-patient group in the task replication. Some possible explanations for these somewhat weaker than expected results are suggested in what follows.

Indeed, when tested across tasks, healthy singers had had comparable similarity scores to what was observed for the singer-patient group in the task replication. Some possible explanations for these somewhat weaker than expected results are suggested in what follows.

Group Differences

The groups compared in this study represented very different populations of singers. The healthy singer group only included female professional opera singers, whereas the singer-patient group included both sexes and a variety of singing styles and levels. Professional classical singer-patient data proved to be practically impossible to collect within a reasonable timeframe and consequently, certain group criteria had to be relaxed or abandoned. The different nature of the groups could in itself be a possible underlying factor to the results obtained in this study.

The button enhanced VRP was not expected to create tasking issues given that the act of singing at a high level of performance also requires considerable multi-tasking. The singer must sing, attend to the accompaniment, play a role, dance and communicate with the audience. Consequently, adding the button device to VRP tasks for this population was not anticipated to cause task performance difficulties. However, the level of expertise of the singer might explain differences in the ability to manage multi-tasking such as required in a button VRP recording.
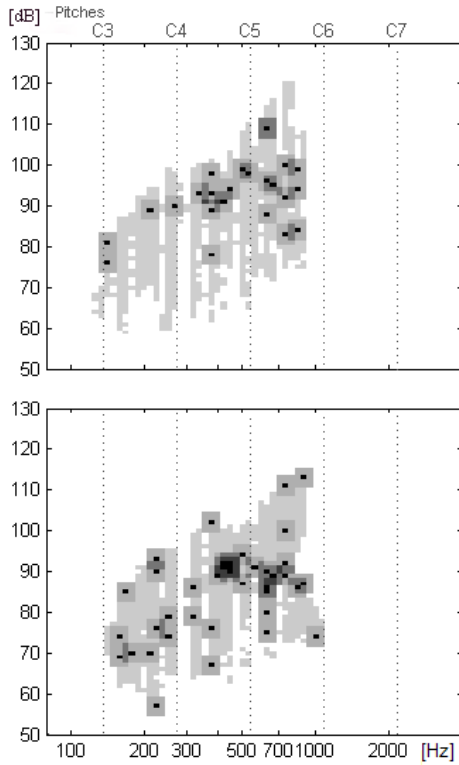
Figure 6. The accumulated button presses for 6 patients with nodules (all females) is compared for Tasks 3 (top) and 5 (bottom). Although pressing behavior changed a little between tasks, a tendency to press in the 400-450 Hz and 90 dB zone as well as just below 700 Hz in the range of 85-94 dB can be noted in both cases. Axes as for Figure 1.
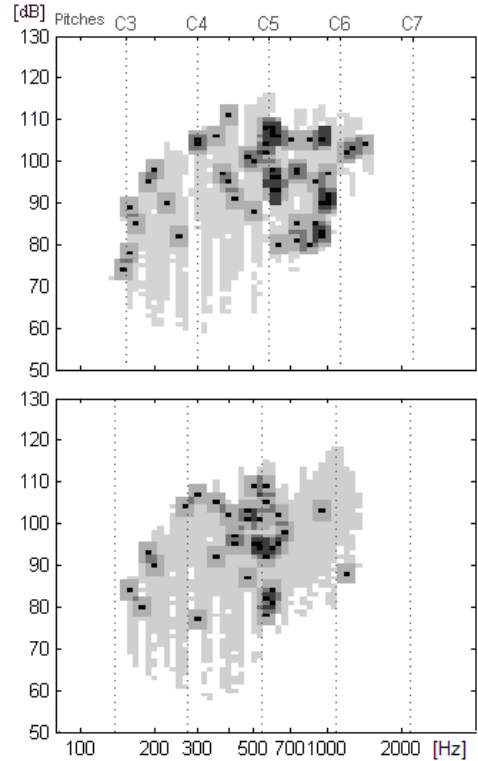
Figure 7. The accumulated button presses for 5 patients (3 females, 2 males) with functional dysphonia is compared for Tasks 3 (top) and 5 (bottom). Pressing occurs mostly at high SPL. Presses along the SPL extent for the frequency range 550-600 Hz was a recurring pattern for this group. Axes as for Figure 1.
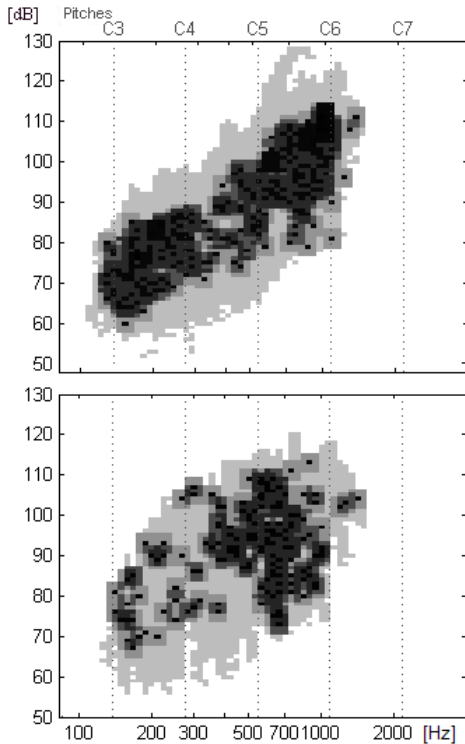
Figure 8. The total accumulated button press information for Tasks 3 and 5. The healthy group and the patient group exhibit clearly different patterns of button pressing. This outcome agrees with initial expectations that singer patients would press in inner VRP areas rather than at the VRP extremes. This trend is manifest despite the typical smoothing effects of averaging. Axes as for Figure 1.

The amateur or student singer may not have acquired a proper vocal gesture to enable their focus to shift to another task like the button pressing. Multi-tasking capabilities aside, the infrequent use of the button device observed for the patient-singer group could also be tied to the fact that these singers are not as aware of their vocal capabilities and limits as are professionals. Levels of self-confidence might also impact the use of the button and the frequency of pressing. It is more likely that professional singers were more self-confident in their vocal technique and performance and hence in using the button.

Load Effect

The proficiency of the singer might not be the only explanation for the observation of lower rates of button pressing and weaker consistent behavior. The greater difference between the two groups, the presence of a vocal complaint related to a voice disorder rather than level of training, can have greatly impacted the results of this study.

A certain vocal disorder load effect may come into play when singer patients are instructed to use the button device. The multi-tasking capabilities of the singer are then impaired by the presence of a disorder and the focus of the subject becomes more centred on the actual performance rather than on an internal analysis of events. The button tasks were perhaps requiring too much attention especially in the framework of a first voice status evaluation. Subject responses in the post-recording questionnaire seem to support this last interpretation. The most frequent responses to question 9 (see Table 2b) were emotional responses describing a heightened need for focus and concentration, loss of self-confidence and fear/insecurity related to phonation.

This load effect could have been potentially avoided with better training possibilities. A short training session had been designed in the context of the healthy group study but was not adopted for the singer-patient group experiment because of its additional vocal demands. Since another aim for these recordings was to collect singer-patient phonetographic data to compare to a recent normative study, tasks followed closely those used in the normative data collection (Lamarche *et al., in press*). We believed that instructing the subject in using the button for speech tasks and for the brief physiological VRP task, would provide sufficient experience in using the button device before the test. Healthy singer button press data in a previous study (Lamarche *et al*., 2008) had also demonstrated the singer's capability of generalising the button pressing across various tasks. For singer patients, however, the transition between the voice tasks and the performance tasks might be challenging enough to impede the generalisation of the button device training. Button pressing was scarce in the speech tasks (perhaps supporting the fact that these subjects had indeed *singing* voice complaints) while subjects pressed the button more often during the physiological VRP recordings. This task, due to its continuous and more rapid nature, was not intended for inclusion in the analysis but nonetheless could be used as an indicator that subjects understood the task at hand and could use the button device. Perhaps the load effect of the vocal disorder explicitly came into play when the singing voice was addressed.

With hindsight, and looking at past research in the field of motor theory, some explanation for the lack of generalisation can be found. Variable practice such as asking the subject to perform the button task in different voice modes typically diminishes training performance and rather improves long term learning (Titze & Verdolini, 2002).

Psychology of the singer

In this discussion of results, the subject's readiness to communicate problems must also be taken into account. Singers are often taught and drilled to hide imperfections, uncomfortable moments, errors and fatigue in order to keep a convincing and smooth performance. This is true for all stage artists, but for the singer, this becomes an intricate part of the act of singing. It is not uncommon, in the case of pathology or vocal problems and in contexts remote from the stage, for singers to shy away from revealing the problematic areas and to avoid or compensate for the problematic vocal gestures. This kind of behavior could consequently affect the way in which the subject presses the button. In a healthy state, when the singer is more self-confident and is relatively free from stress, button pressing could be more frequent without necessarily reflecting the totality of the singer's concerns. It would

follow that when the subject does push the button device, the importance of that button press is not negligible despite the lower rate of pressing.

In a similar line of thought, and in a more qualitative approach, it could be productive to address the subject's personality. For the singer, the impact of a vocal disorder is great and the button pressing could be heavily linked to the singer's coping strategy (i.e., either hiding and not pressing, or obsessing and pressing excessively).

Location of button presses

Expectations were met as to *where* in their VRP space subjects would press the button. This was a positive and central outcome of this study. Subjects did not press as often and as consistently as expected, but they did press the button in different regions of their VRP space in comparison to the healthy group. These different group tendencies were observable despite the effects of averaging. The total density plot averages for the healthy and the singer-patient group, Figure 8, clearly depicts the tendency of the singer patients to press in a latitudinal fashion, cutting through their VRP throughout the SPL extent of a 500 to 800 Hz frequency range. The healthy singers, on the other hand, tended to press at both extremes of their VRP. In Figure 8, the Gaussian (normal) character that results from the accumulation of any distributions has somewhat obscured this tendency to the extremes. However, we choose to present the data in this way, rather than to devise some scheme for VRP shape normalization. This outcome reveals the importance of examining the vocal dynamic capabilities of a voice and recording a full VRP rather than just a contour. The *messa di voce* exercise, for example, seems to be an ideal exercise for such an evaluation. The exercise indirectly gives information on the limitation and variation flexibility of the voice source. It would seem that problems or challenges are especially perceived along the full SPL extent of a voice in its higher range where control in general is more difficult and consequently extremely difficult in the case of a vocal disorder.

In Figure 6, the results representative of 6 subjects with nodules are illustrated. Nodules can be expected to impact the singing voice in terms of limitations of the upper range, onset delays in soft phonation (most particularly in the high range), reduction of the vocal endurance, a sense of increased effort and increased day-to-day voice variability. In the button press map, highlighted areas are mostly concentrated in the upper range and at the center-like part of the sound level extent. Three dominant areas can be noted: 1) a 450 to 550 Hz band in the top half of a 65 to 103 dB extent, 2) just below 700 Hz at the half point of a 72 to 110 dB extent and 3) the 850 Hz region at 10 to 15 db higher than the minimum of a 73 to 115 dB extent.

The functional dysphonia group exhibited a similar pattern (Figure 7). In both tasks, the subjects pressed mainly in the frequency region just above 523 Hz. The majority of button presses was found between 90 and 108 dB in Task 3 while somewhat lower, from 80 to 95 dB, in Task 5. In Task 3 there were also some highlighted regions in the higher range at approximately 880 Hz with similar SPL as found for the lower frequencies just mentioned. Button presses for the group with functional dysphonia were located slightly higher in frequency than what was observed for the group with nodules; however, with regard to SPL, the occurrence of these dominant button press regions was much the same. Interestingly, *mezzo forte* (an intermediate musical dynamic) in female singing is known to be approximately between 88-90 decibels. (Nawka et al., 1993, Lamarche et al, 2008) This

would mean that the dominant regions of button pressing corresponded approximately to the *mezzo forte* portion of the *messa di voce* exercise. Since this type of exercise requires very fine motricity involving gradual vocal fold adduction changes (Titze *et al.*,1999) subjects might feel particular control difficulties and even discomfort in the higher range of the voice (where the vocal folds are stretched and thin) as vocal fold mass must subtly, smoothly and gradually increase or decrease in order to regulate the resistance of the folds to the airflow. Furthermore, because the descending portion of a *messa di voce* can be anticipated to be particularly challenging, button presses were studied in relation to the audio signal. However, no specific timing trend with a particular ascending or descending gesture of the *messa di voce* could be discerned.

Further investigations including larger groups with better possibilities to categorize according to both diagnosis and voice category could assess more readily the kind of pressing pattern here observed. Clearly, singer-patient button presses reflect an extra dimension to the VRP recording since they convey information that is not necessarily available in the audio signal.

From the singer patient's point of view

Subject feedback concerning the relevance of button presses to their performance experience was generally affirmative. In the previous study, ratings like these were important for understanding the subject's experience of such a new device (Lamarche *et al.*, 2008) This time, the questionnaire was used both as a tool to validate instruction comprehension and to identify the main motivations leading to button pressing. Both VAS ratings and qualitative data shown in Figure 3 and tabulated in Table 2 a) and b) contain valuable material in relation to the button VRP data. Every subject answered in the negative when asked if the difficulties coinciding with the button presses were momentaneous, of an incidental nature. Informally. patients sometimes mentioned that it was difficult to choose *when* to press since the difficulties were felt constantly. This difficulty of choosing the right time to press the button could also explain the reduced rate of button presses and the low scores obtained for question 6. It can also be concluded that when subjects did press the button, the button press in question was likely to be related to a heightened sensation of difficulty.

Since each recording session was filmed in order to ensure good methodological procedure coherence and to document calibrations, subject comments in the in-between task time could be noted.

At times, the subject reported not pressing in specific places, or else wanting to immediately explain the presses. It was clear that there was a general tendency to forget to press the button even though the button device instructions were repeated at the onset of each task. These observations are consistent with the notion of a possible vocal disorder load. A subtle visual reminder placed in the vision field of the subject might be a small but helpful addition to the protocol and reduce memory load.

From the clinician's point of view

This type of enhanced VRP was developed and tested with the intention of improving evaluative possibilities for the clinician working with the singer patient. It was deemed important to integrate a clinical angle to the results presented above and to the overall experiment process.

Clinically, the button markings were perceived as instrumental in depicting precise frequency areas that pose problems to the patient. Not only could the clinician assess physiological and acoustical voice aspects in one visual assessment but the somaesthetic dimension could now be included as well and in this way give a more complete image of the voice. Furthermore, the button augmented VRP was found to lead easily to constructive dialogue with the patients and in particular facilitate the verbalization of the patient's self-perception of the problem. Finally the button presses in the VRP were viewed as a tool to further objectify the communication of patient data among treatment partners (speech language therapist, osteopath, laryngologist etc.).

The button augmented VRP can help the clinician in his/her challenge to summarize multiple information channels. It makes a link between what is observed, what is heard and what is not.

Pedagogical possibilities

Since the button is used concurrently with the singing task, it allows for an uninterrupted performance, leaving areas marked for later discussion and evaluation. In a motor-theory of learning perspective, this non-invasive way of noting areas of particular difficulty could support improved long term learning, as the button VRP tracks the subject's ratings of their own performance and can then be used as augmented feedback to increase the effectiveness of feedback in a learning situation (Titze & Verdolini, 2002). Moreover, this kind of feedback would typically be used as terminal feedback and therefore be an optimal type of biofeedback in a learning situation. The button-VRP could be further developed to offer interesting interactive possibilities where the pedagogue or speech therapist also button marks the VRP (in a different color scheme) as the student or patient performs. Perception is very individual yet often at the center of a successful vocal education or rehabilitation. The button in this way could indicate clearly where perceptions diverge and where they meet, and in this way give rise to educative discussions and analyses.

Conclusion and Future Work

As inferred by Titze (1992), the VRP and its features can be very useful in voice habilitation and rehabilitation. An attempt to further enhance the VRP as a clinical tool by integrating the singer's perception via button presses can give greater value to the already speedy and easily accessible VRP voice assessment information. Initial results are promising. The singer-patient group showed a comparable consistent button pressing behavior to the healthy group. A more consistent behavior however, was not noted. It had been expected that singer patients would actively use the button device since vocal effort was assumed to be generally higher, yet the opposite behavior was observed. Finally, in comparison to the healthy group, an overall button pressing trend for the singer-patient group was clearly identified: healthy singers press at the extreme limits of the VRP while singer patients tend to press in a narrow band of the higher frequency range in the middle of the SPL extent for that band.

From the qualitative point of view, the use of the augmented VRP was a positive experience for both the patient and the clinician. Further work in finely tuning the button VRP is however needed. More specific models of the source of error such as vibrato and response delay time need to be

implemented in order to achieve a deeper analysis level of button pressing similarities and information. As a proof of concept type of study, this work offers interesting possibilities for clinical applications as well as research.

The button-VRP could be an important documenting tool for pre-post vocal status. It could help objectify the assessment of change in voice status which is not necessarily only related to vocal range. Furthermore the assessment of pre and post interventions like surgery or rehabilitative therapy could be greatly supported by these button presses as they could directly attract attention to problem areas and in this way refine the course of the rehabilitative process. On the same lines, this kind of button enhanced VRP could even be of diagnostic interest. The button markings in the VRP could potentially map out interesting frequencies for closer laryngoscopic examination.

Acknowledgments

References

Cabrera D, Davis P, Barnes J, Jacobs M, Bell D. Recording the operatic voice for acoustic analysis *Acoustics Australia / Australian Acoustical Society.* 2002; 30(3):103-108.

Cohen Seth M, Jacobson Barbara H, Garrett C Gaelyn, et al. Creation and validation of the Singing Voice Handicap Index. *The Annals of otology, rhinology, and laryngology.* 2007; 116 (6):420-6.

KayPENTAX. Voice range profile, model 4326. Retrieved 2008-08-06, (Last updated, September 15, 2008) from: **http://www.kayelemetrics.com/Product%20Info/CSL%20Options/4326/4326.htm**

Koufman JA, Blalock PD. Classification and approach to patients with functional voice disorders. *The Annals of otology, rhinology, and laryngology.* 1982; 91(4 Pt 1):372-7.

Lamarche A, Ternström S. An Exploration of Skin Acceleration Level as a Measure of Phonatory Function in Singing. *Journal of Voice.* 2008; 22(1):10-22.

Lamarche A, Ternström S, Hertegård S. Not just sound: supplementing the Voice Range Profile with the singer's own perceptions of vocal challenges. *Journal of Logopedics, phoniatrics, vocology.* 2008; 13:1-8. DOI:10.1080/14015430802239759

Morsomme D, Gaspar M, Jamart J, Remacle M, Verduyckt I. Adaptation du VoiceHandicap Index à la voix chantée. *Revue de laryngologie-otologie-rhinologie.* 2007; 128(5): 305-314.

Morsomme D, Simon C H, Jamart J, Remacle M, Verduyckt I. [A proposal to adapt the Voice Handicap Index to the singing voice]. *Revue de laryngologie - otologie - rhinologie.* 2005; 126: 305-313.

Murry T, Zschommler A, Prokop J. Voice Handicap in Singers. *Journal of Voice.* 2008. Jan 10. Doi:10.1016/j.jvoice.2007.10.013

Nawka, T, Seidner W, Cebulla M. Experiments on intentionally influencing the singer's vibrato *Proceedings of the Stockholm Music Acoustics Conference,* 1993 :221-228

Phyland DJ, Oates J, Greenwood KM. Self-reported voice problems among three groups of professional singers. *Journal of Voice.* 1999; 13(4):602-11.

Rosen C A, Murry T. Voice handicap index in singers. *Journal of Voice.* 2000; 14:370-7.

Schutte H K, Seidner W. Recommendation by the Union of European Phoniatricians (UEP): Standardizing Voice Area Measurements/Phonetography. *Folia Phoniatrica.*1994; 35:286-88.

Titze I.R., Verdolini, K. (in preperation) <u>Vocology</u>. 2002

Titze I R, Long R, Shirley, G I, Ramig, L O, Caroll, L M, Riley, W D. *Messa di voce*: An investigation of the symmetry of crescendo and decrescendo in a singing exercise. Journal of the Acoustical Society of America, 1999: 105(5) : 2933-2940.

Titze I R. Acoustic interpretation of the voice range profile (phonetogram). *Journal of Speech and Hearing Research.* 1992; 35: 21–34.

Appendix A

Questionnaire post enregistrement /   Post recording Questionnaire

Cocher sur la ligne le grade perçu. Toujours faire le lien avec votre état vocal actuel.
Mark the line according to your perception. Always refer to your current voice status

1. **Impression globale de votre contrôle vocal tel que perçu aujourd'hui.**
   Indicate your overall impression of vocal control today.

   Mauvais/Bad _____ **Bon**/Good

2. **Les zones identifiées pendant l'enregistrement représentent-elles vos difficultés habituelles ?**
   Are the highlighted portions of your phonetogram typical areas of difficulty or/and limits?

Atypique /Atypical _____ **Typique**/Typical

3. **Selon vous, les pressions exercées sur le bouton traduisent-elles correctement   votre expérience vocale de ce jour ?**
   Do the button presses relate well to your singing experience today?

   Non/No _____ Oui/Yes

4. **Les pressions exercées sur le bouton correspondent à quelle sorte de difficultés vocales?**
   With which kind of vocal difficulties do the button presses coincide?

   Momentanées/Incidental   _____   **Systematiques/**Systematic _____
   (Ex: phlegm on the vocal folds)

   Lesquelles/ Which ones

   _____
   _____
   _____

5. Pouvez-vous brièvement décrire la raison principale qui vous a poussé à utiliser/activer le bouton?
Can you explain the main reason for your use of the button?

_____
_____
_____
_____

6. **Évaluer de façon générale votre effort vocal en lien avec votre utilisation du bouton.**
Grade your general vocal effort at the times of button presses.

Aucun/None _____ E**xtrême**/Extreme

7. **Spécifiez en quelques mots clés ce qui décrit le mieux votre effort/inconfort vocal ?**
Which words best describe your vocal effort?

_____
_____
_____

8. **L'effort vocal ou l'inconfort vocal sont-ils fréquents lors de vos prestations vocales ?**
Is this type of effort frequent in your performance?

Non/No _____ **Oui**/Yes

9. **Si oui, comment influent-ils sur vos prestations vocales?**
In case of the affirmative, how does it influence your performance?

_____
_____
_____

Paper V

Paper V is in review in *Logoped Phonaitr Vocol*

The layout is customised.

The Swedish version of the Voice Handicap Index adapted for Singers

ANICK LAMARCHE[a], JOACHIM WESTERLUND[b], INGRID VERDUYCKT[c], STEN TERNSTRÖM[a]

[a] Kungliga Tekniska Högskolan, School of Computer Science and Communication,

Department of Speech, Music and Hearing, [b] Stockholms Universitet, Institute of psychology

[c] Université catholique de Louvain, Cliniques Universitaires Saint Luc

Abstract

Objective: The recent Belgian adaptation for singers of the Voice Handicap Index (VHI) was translated and readapted in Swedish. This study's aim was to evaluate the validity and reliability of a Swedish version. **Method:** In a parallel group design, 96 healthy singers and 30 singer-patients with various diagnoses completed a Swedish version of the singer adapted VHI. A prospective evaluation of the Swedish voice health status instrument was carried out. In average, delays between test-retest were between 14 to 16 days. Validity and reliability as well as the internal coherence and group differences were assessed. **Results:** The singer-patient group scored significantly higher than the control group. Reliability was confirmed by high Cronbach's alpha (>.78) for test-retest scores as well as each subscales. In particular, test-retest stability in both groups was confirmed by high values for Cronbach's alpha (>.8). For both the control and patient groups, test and retest scores compared closely to previously reports with respect to overall scores. Retest results were slightly lower than initial test scores. **Conclusions:** The Swedish translation of the adapted VHI for singers (Röst Handicap Index för sångare or RHI-s) is valid and reliable and shows sensitivity to the singer's concerns. It can be considered a useful tool in the clinical assessment of Swedish healthy or pathological singers.

Keywords: *Voice Handicap Index, singers, validity, reliability, Swedish, voice disorders, singing levels, singing genre, self-perception*

Introduction

According to the World Health Organization, health is described as a "state of complete physical, mental and social well-being and not merely the absence of disease or infirmity"(1). This definition embeds the importance of the patient's self-perception of his/her problem. Indeed, this widely accepted definition has strongly impacted the health care system and the way patient assessment is carried out.

Clinical approaches increasingly value and measure patient self-perception and experience as an integral part of the overall evaluation process. In voice clinic environments instruments for self-report are commonly included in protocols. An example is the Voice Handicap Index (VHI), a Likert or summation scale type of questionnaire introduced in 1997 (2). This questionnaire was constructed with a three subscale structure: a functional scale, a physiological scale and an emotional scale.

Correspondence: Anick Lamarche (M.Mus), Fax: 08-790-7854. E-mail: anick@nada.kth.se
(In review)

Each subscale contains 10 items and is worth 40 points. The total possible points are 120. High VHI scores typically correspond to a severe impact of voice disorder (2). With this instrument, attention is directed to the patient's own vocal experience of the reported problem and thus the patient's perception becomes a key element in the rehabilitation process.

Since 2002, when the Agency of Healthcare Research and Quality officially recognized the VHI to be a "reliable and valid diagnostic tool", the VHI has become the most commonly used self-perceptual test in voice clinics (6). In 2004, a thorough study of Franic *et al.* reported on the quality of the questionnaire.

"(the) VHI was the only Voice Quality of Life (VQOL) instrument to meet study consistency and reproducibility standards for the overall instrument score (alpha = 0,95 and r = 0,92,respectively)" (7).

Many validated translations of the original VHI now exist in Polish, Hebrew, German, French, Chinese and Swedish. The equivalence of translations across Europe and their comparison to the original American version was recently deemed valid (8) and good internal coherence was found. Moreover, the usefulness of the three subscale structure was confirmed for all VHI versions. Certain adaptations might, however, be necessary to optimise the use of the VHI with specific population groups (9-12). Singers have high vocal demands and therefore form an exemplary group in need of this kind of specific attention. Singers are generally well in tune with their vocal health. As others have advanced, singers often address vocal problems promptly and are most likely to seek consultation rapidly (14). Furthermore, the vocal concerns of this population might not even relate to the speaking voice (13). In 2000, Rosen and Murry, while determining the degree of handicap reported by singers, demonstrated that this population's VHI scores were lower than the general incoming voice

patient. At that time, they reached the conclusions that a low score did not necessarily indicate a weak handicap and rather pointed out the specific needs of this group. In their analysis, Rosen & Murry identified only five items of the VHI (on a total of 30) that could have some relevance to singing voice problems (14).

In 2003, a five study project was undertaken in Belgium. The aim was the construction of a VHI adapted for singers. This work, performed in French, extracted pertinent items from the original VHI and sought to include new items that would truly reflect the patient's concern (11,15). More recently, Cohen *et al.* also introduced a singer specific instrument with the creation of the Singing Voice Handicap Index (SVHI). This new index was constructed based on the combination of symptoms reported by singers and clinical experience. The index resembled the original VHI with its 5-point Likert scale yet, differed in that it discarded the physical, functional and emotional subscales and rather, adopted a single scale structure (the results of a principal component analysis motivated this choice). The final version contained 36 items for a total of 140 points (12). Murry *et al.* also followed in the same direction, including three singer-relevant items in the original VHI-10 to meet the needs of singers (a reduced version of the VHI including 10 items only) (15).

The aim of the current investigation was to translate and adapt in Swedish the VHI for singers (originally in French) and test its validity and reliability with the Swedish singing population. A large part of the work was inspired by the last reworked version produced by Morsomme *et al.* (11, 15) but at the same time, it sought to combine certain aspects of the work accomplished by Cohen *et al.* (12). This work's driving hypothesis was that the Swedish singer adapted VHI (RHI-s) adequately evaluates the voice handicap of the Swedish singer.

Methods

Translation and adaptation (French to Swedish)

The first step in this work was to translate in Swedish the work accomplished in French by Morsomme *et al.* (11,15). For this step, author I V, a Belgian logoped (or speech language pathologist (SLP)) educated in Sweden, and author A L, independently translated the Belgian version to Swedish. These two translations were compared and modified upon unanimous decision. The differences between both translations were principally matters of style and formulation rather than content. Thereafter, a lay French native, well versed in Swedish, performed a retranslation of the work to French. This process was somewhat more challenging, as the individual involved in this part of the work had no prior knowledge in the area of voice. At the same time, this ensured full objectivity in the translation process and exposed immediately any unclear aspects of the questionnaire items. Each item of the instrument was discussed individually and points of divergence from the Belgian version were addressed in a group editing context. The editing group, composed of two Swedish logopeds (or SLP) singing voice specialists and the first author, worked in including corrections as well as certain adaptations idiomatic to the Swedish language. All initial 30 items and subscale structure were preserved. This final version was submitted to five judges: a professional singer, a choir singer, an opera student, a voice pedagogue, and finally a logoped. According to judge responses, minor adjustments were made and a final version of the Swedish version of the VHI adapted for singers (Röst Handikap Index för sångare or RHI-s) was achieved. The RHI-s was made available online in the form of an interactive Portable Document Format (PDF) form. See appendix A to view the form used in this study.

The confidentiality of the information was preserved in that information was not stored on internet but rather, subjects submitted filled forms by electronic or regular mail.

Participants

An ethical vetting was obtained by the Regional Ethical Review Board in Stockholm ("*Regionala etikprövningsnämnden i Stockholm*", certificate 1358-31). A total of 126 volunteer subjects were recruited for the project. Subjects were divided into two groups**:** (1) A healthy vocal group, with no report of vocal problem history, counted 96 singers (36 males, 60 females) with a mean age of 39 , standard deviation (SD) of 11 years (range 20-63). This group was made up of professional, amateur and student level singers of all genres (the classical genre was dominant). These participants were recruited randomly in different choirs, the Royal Opera House and different music schools in Stockholm. This group was the control group for this study. Control group data is organized in Table 1a).(2) A patient group, including 30 patients (11 males, 19 females), with a mean age of 35 and SD of 11 years (range 19-58), recruited in clinics across Sweden. The singing level distribution of the patient group was fairly comparable to the control group's distribution. Choir singers formed a large part of the control group (69%) whereas vocal soloists formed the majority in the patient group (67%). The dominant singing genre (classical in the control group) became "other" for the singer-patient group (the latter category included genres such as afro, pop, jazz, blues, rock and soul). The classical genre remained however, largely represented; 27% of the patients defined themselves as classical artists. The Broadway/Musical genre also increased from 1% in the healthy group to 13% in the singer-patient group. Table 1b) contains the data for the different singer-patient subgroups. Patients recruited had a wide variety of diagnoses. Diagnoses are tabulated in table 2.

classes. In the case of the test phase, singers met with the investigator and received brief oral and written instructions. The test form was filled within five to ten minutes. With respect to the retest, singers were contacted again either by email, telephone or through choir directors, teachers or repetition coaches. The procedure was quite similar for the patient group. Singers were contacted in clinical environments (phoniatric or logopedic) where they were asked to fill the form in the waiting room. Most patients were later contacted either by Internet or by telephone for the retest phase.

Healthy Singers		Females			Males				
		Students	Amateurs	Professionals	Students	Amateurs	Professionals	**Total**	%
Choir	Classical	—	14	26	17	12		69	72
Solo	Classical	9	—	2	2	—	2	15	16
	Broadway / Musical	—	—	1	—	—	—	1	1
	Gospel	2	—	2	—	—		4	4
	Other	1	—	2	—	—	2	5	5
Band	Other	—	1	—	1	—	—	2	2
	Total	12	15	33	3	17	16	96	100

Table 1a. The control group according to gender, singing level, context and genre. Percentages are provided to facilitate between groups comparisons. Note the large representation of choir singing in this group.

Singer-Patients		Females			Males				
		Students	Amateurs	Professionals	Students	Amateurs	Professionals	**Total**	%
Choir	Classical	—	1	1	—	—	1	3	10
Solo	Classical	—	—	3	—	—	2	5	17
	Broadway / Musical	1	1	—	1	—	1	4	13
	Gospel	—	1	—	1	—	—	2	7
	Other	2	1	3	—	1	2	9	30
Band	Other	—	1	4	—	—	2	7	23
	Total	3	5	11	2	1	8	30	100

Table 1b. The singer-patient group according to gender, singing level, context and genre. Percentages are provided to facilitate between groups comparisons. Note the large representation of soloists as well as the great number of singer-patients reporting Contemporary Commercial Music styles. Here the category "Other" includes styles such as afro, jazz, soul, blues and soul.

Subject selection criteria were defined as 1) Swedish speaking, 2) regular singing practice, 3) never have answered the RHI-s, 4) not have benefitted of SLP therapy between the test and the retest and 5) be at least 16 years of age.

Procedure

Singers who were part of the healthy group were contacted at choir rehearsals, during opera intermissions, in lesson time and during master

Diagnosis	Number of Cases	Mean Score	Standard Deviation
Edema	8	55.1	19.5
Dysphonia	6	51.1	10.5
Nodules	4	62.0	23.9
Chronic Laryngitis	3	44.5	17.0
Polyp	2	57.0	10.2
Sulcus	1	33.0	-
Hemmorhage	1	80.0	-
Cyst	1	60.0	-
Dysarthria	1	48.0	-
Vocal Fold Scarring	1	71.0	-
Larynx disease	1	77.0	-
Benign vocal folds	1	34	-

Table 2. Singer-patient reported diagnoses. The average RHI-s total score and the standard deviation for each diagnosis group are given in columns three and four.

Some patients received a copy of the form in a pre-stamped envelope and were asked to complete the form in two weeks time. A reminder to the effect of the retest delay time was included on the Test form patients received. Moreover, most patients were additionally reminded by telephone or electronic mail.

In some cases, the follow-up medical session was planned to correspond to the delay interval and patients could do the retest in the clinic. The amount of time allotted between test and retest measures was considered very critical. When measuring the same thing twice, the time elapsed between measurements will strongly impact the correlation of the measurements. In the case of scales such as the VHI, the Grade, Rough, Breathy, Asthenic and

Strained (GRBAS) scale and the Voice Symptom Scale (VOISS), 14 to 16 days delay is strongly recommended in the literature (17, 18, 19). It was of high importance that the delay time would be long enough so that subjects would not remember their test responses yet short enough so that the voice status would not change significantly. Morsomme *et al.* obtained an average of 13.85 days (SD 3.08 days) (15) and Cohen *et al.* reported an average of 17.3 days (12). In this study, the test-retest delay time for the healthy group was on average 14.8, SD 3.3 days and for the singer-patient group, the mean was somewhat higher; 16.1 SD 3.8 days.

Statistical Analysis

The data analysis was designed to assess if the RHI-s correctly measured the impact of the voice disorder on the singer. The distribution of the data, validity and reliability as well as variable effects were assessed with the statistical software S.P.S.S. (SPSS Inc, Chicago, Illinois, version 15.0.1, 2006). An alpha level of 0.05 (two-tailed) was used for all statistical analysis.

Validity

Validity testing refers to the degree to which the test actually measures what it claims to measure. Test validity is also the extent to which inferences, conclusions, and decisions made on the basis of test scores are appropriate and meaningful.

Construct validity was assessed, for the patient group, by a Pearson's r correlation measure. The patient's VAS ratings of the overall severity of the voice problem were compared to the patient's RHI-s scores.

Discriminant Validity was determined by comparing mean scores between the control and the patient group. This comparison was performed by means of and independent samples *t*-test.

Reliability

Reliability testing mainly refers to the degree to which a test is consistent and stable in measuring what it is intended to measure. In short, a test is reliable if it is consistent within itself and across time.

Test-retest reliability was determined by a Pearson's correlation yet, for the sake of data comparison, Spearman's rho was also included. This analysis, like a matched *t*-test, served to assess the test-retest stability over time. The delay time ranged approximately between 14 to 16 days as described above. Correlations were observed for the total of subjects as well as the healthy and patient groups respectively. Test and retest mean differences were calculated and critical differences (defined as 2 standard deviations from the mean) were defined.

Internal consistency was measured with the Cronbach's alpha. The higher the score, the more reliable the generated scale. The internal consistency was determined separately for both test and retest as well as for controls and patients. Internal consistency for the three subscales was also investigated. For this purpose, a Cronbach's alpha was obtained for each subscale in the test, as well as in the retest, and for controls and patients respectively.

Factor analysis was performed in order to further verify the pertinence of the RHI-s subscale structure. In such an analysis, the group population should be roughly 5 times greater than the items tested. In our case, a total of 126 singers amounted to 4.2 times greater than the 30 items tested.

Age, sex, singing genre, singing level, and singing context were all variables that could potentially influence RHI-s scores. Such effects were controlled for each of these variables.

Results

Figure 1 a-b) depict the RHI-s mean score

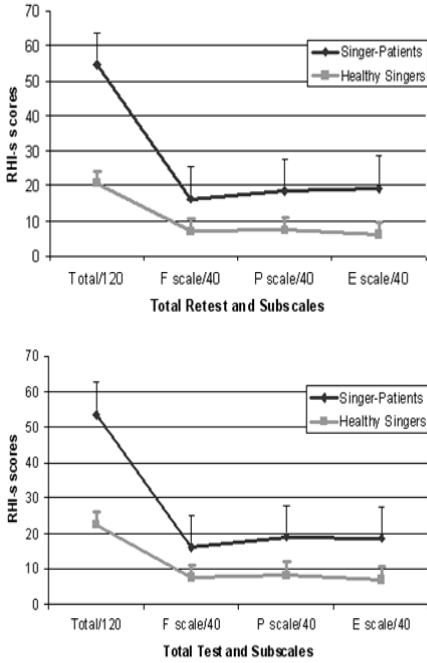trends for both the control and the patient group.

Figure 1 a-b. Singer-patient and healthy singer RHI-s test scores and b) for the RHI-s retest scores. The total points (over 120 possible points) and the functional (F), physiological (P) and the emotional (E) subscales scores (over 40 possible points respectively) are illustrated. Positive standard error bars depict the statistical variation present in the data.

Validity

The correlations found between VAS scores and the RHI-s scores were high and thereby supported the construct validity of the RHI-s; r =.74 (p <.001, N = 30) for the test and r =.84 (p <0 .001, N = 30) for the retest.

To assess the discriminant validity of the instrument, mean RHI-s scores were compared between groups. Both test and retest averages differed significantly. The independent sample *t*-test results were; (*t*=-10.794, df =124, *p*<.001, d= 2.28) for the RHI-s test score and (*t*=-10.777, df =124, *p*<.001, d=2.28) for the Retest score.

The data was examined to investigate further the test's capacity of differentiating between a healthy singer and a singer-patient. In this investigation, all patients had a score equal or higher than 31. The possibility of defining 31 as a cut-off score was explored. Such cut-off resulted in .1 sensitivity and .76 specificity. Figure 2 illustrates the receiver operating characteristic curve for the test results of both groups. Healthy singers in average had a total score of 22 (on 120), SD 13; 76 out of 96 participants had lower scores than 31. On the other hand, singer-patients *all* scored above 30, the group average was 54, SD 18.
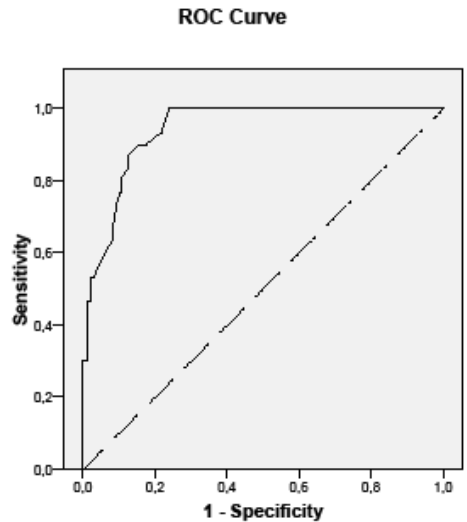
Figure 2. The receiver operating characteristic curve for the test results of both control and patient groups. For a cut-off score of 31 points, a 100% sensitivity yielded a 76% or 1-.24 sensitivity.

Reliability

Test-retest stability was confirmed by high correlations values. When both groups were pooled together, test-retest scores yielded a

Pearson's r correlation of .91 (.89).[1] When groups were studied separately, a correlation of .81 (.81) was found for the healthy group and a correlation of .85 (.86) for the patient group. The mean difference scores were quite low and singer-patient scores changed the least from test to retest. Table 3 contains the mean differences found for total and subscale scores, the confidence intervals for those means as well as two-times the standard deviations from the means. This last information was considered useful in determining the clinical significance of the RHI-s. Total score differences greater than 16 and changes of more than 7 points on the physiological and emotional subscales and 6 points on the functional subscales were considered significant in terms of voice status change.

Group		N	Mean	CI 95%		2*SD
Control	Total	96	1.93	-0.03	3.89	16
	Physiological	96	0.93	-1.03	2.89	7
	Functional	96	0.20	-1.76	2.16	6
	Emotional	96	0.89	-1.07	2.85	7
Patient	Total	30	0.50	-1.46	2.46	20
	Physiological	30	0.63	-1.33	2.59	9
	Functional	30	0.27	-1.69	2.23	10
	Emotional	30	0.70	-1.26	2.66	9

Table 3. Mean differences between test and retest RHI-s total and subscale scores for control and singer-patient groups. The 95 % confidence interval is given (column 4-5) for each of the mean differences. Two standard deviations defined the critical limit and thus the definition of a clinically significant change.

Internal consistency was very high. Because this analysis considers item scores, the distribution for each item was assessed. Item 3 and 27 were particularly skewed. To address this aspect of the data distribution, a square root data transformation was performed on all items and the Cronbach's alpha was reassessed. However, only negligible differences were observed. Table 4 indicates the Cronbach's alpha obtained for the RHI-test and the RHI-

---

[1] Most consulted reports seem to show a preference for Spearman's rho. For the sake of comparison, this correlation is also provided in parentheses. However, since data distributions were normal, the authors did not see any reason why not to use Pearson's r.

retest *untransformed* item scores for pooled groups as well as separate groups.

	Control	Patient	Merged Group
RHI-s Test	0.919	0.943	0.959
RHI-s Retest	0.938	0.941	0.965

Table 4. Cronbach's alpha for distinct groups and merged groups, and for test and retest instances respectively. These results support a reliable internal coherence of the RHI-s.

The Cronbach's alpha was also examined for each subscale. Table 5 lists the Cronbach's alpha obtained per subscale and per test or retest for pooled groups as well as separated groups.

Subscales	Control		Patients		Merged Groups	
	Test	Retest	Test	Retest	Test	Retest
Physiological	0.835	0.896	0.836	0.834	0.904	0.924
Functional	0.852	0.804	0.783	0.80	0.865	0.8
Emotional	0.820	0.834	0.890	0.893	0.904	0.919

Table 5. Cronbach's alpha results for subscale score analysis for distinct groups as well as pooled groups.

The square root transformed RHI-s scores were subjected to PCA. This revealed the presence of four components with eigenvalues exceeding 1, explaining 50.4%, 6.9, 4.4 and 3.4% of the variance, respectively. Oblimin rotation was performed to aid in the interpretation of these two components. The rotated solution revealed the presence of a simple structure, in which factor one comprised both physiological and functional factors; factor two was best defined by performance related items; factor 3 identified emotional items that tended to be oriented towards the concern of other's perceptions and factor 4 grouped mainly self-perception emotional items, see table 6.

For each participant and for each factor, a factor score was computed. A 2 (Group; patient/control) X 4 (Factor; 1/2/3/4) mixed analysis of variance showed a significant difference between the two groups ($F_{1, 124}$ =

38.88, p < .001). However, a significant interaction effect Group X Factor ($F_{3,\,372}$ = 18.12, p < .001) revealed that patients differed from controls only when factor 1 and 4 were considered, see figure 3.
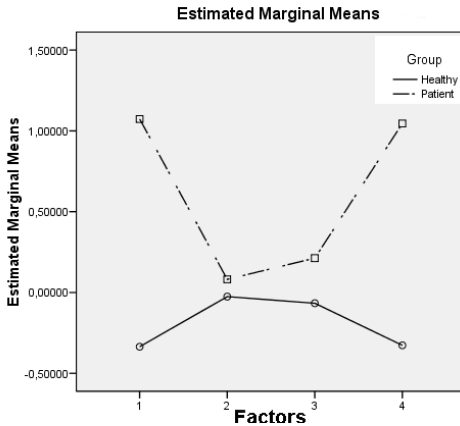

Figure 3. The estimated marginal means for the score results associated to the PCA determined four factors. Factor 1 (mostly related to functional and physiological items) and factor 4 (grouping internal emotional items) are clearly most apt in significantly distinguishing between controls and singer-patients.

In order to examine sex-differences, a 2 (Sex) X 2 (Group; patient/control) between-subjects analysis of variance was performed. A significant Sex X Group interaction ($F_{1,\,122}$ = 5.79, p = .018) revealed that the difference between the two groups (patient/control) on the RHI-s was greater for female than for male participants. Figure 4 illustrates this effect.

Discussion

In this investigation we tested the hypothesis that the Swedish singer adaption of the VHI (RHI-s) is able to successfully evaluate the level of voice handicap of a singer. Given that the patient's subjective perception of his or her voice disorder is practically as important as an effective physical treatment programme (22), it follows that tailoring an instrument such as the VHI to meet the needs of singers is beneficial for the patient as well as for the clinician. In this way, subjective perception of the patients (singers) is better assessed and patient assistance can be improved.
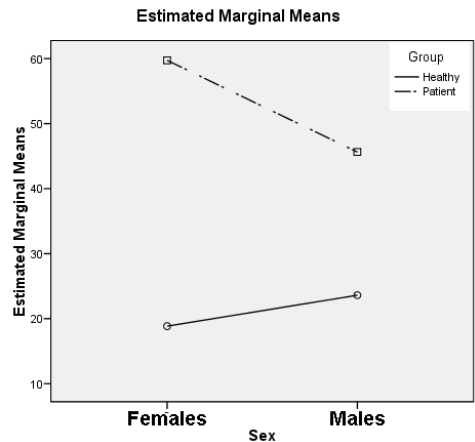
Figure 4. The estimated marginal means between sexes. This figure plots the results of an ANOVA investigation of the effect of sex on RHI-s scores. An effect was found for the total RHI-s scores only. Females scored differently according to their voice status whereas this effect was negligible for men.

Subjects

Subject groups might appear small. According to the Swedish union of artists, there are approximately 500 000 registered choir singers in Sweden; 2 choirs are officially recognised as professional choirs (personal communication with *Sverigeskörförbund* and *Rikskonserter*). In short, choir singers alone make up 6 % of the total Swedish population (this statistic is being reviewed in 2008). Unfortunately no statistics could be found for professional soloists or other groups of singers. It is also very likely that this above mentioned statistics is slightly

Questionnaire Items	Components			
	1	2	3	4
12. F: My voice is unstable (worsens with phonation or during the course of a song).	.861			
19. F: I do not succeed in obtaining a good voice even after a warm-up.	.803			
18. P: My voice is breathy.	.768			
26. P: My voice feels rough and dry.	.749			
02. P: It feels like I must force to be able to sing.	.747			
11. F: I cannot reach high pitches when I sing.	.732			
01. F: I have difficulties in changing from one voice register to another.	.702			
15. E: I am dissatisfied with my voice colour (timbre, focus and core).	.701			
20. P: It happens that it is difficult to know whether my voice will sound clear or not.	.690			
21. P: I usually need to force my voice to sing.	.660			
17. P: The sound of my voice varies during the course of one performance, rehearsal or concert	.633			
07. E: I am worried by my voice difficulties	.453			.585
13. P: My speaking voice is worst after singing.	.505			
14. P: I feel a discomfort or pain in the larynx when I sing.	.486			
04. P: I have difficulties to complete my musical phrases.	.414	.373		
05. F: My voice is not easily heard over an accompaniment (piano, orchestra or instrumental group).		.799		
22. F: I have difficulties to express my feelings when I sing.		.780		
16. F: I have difficulties in adapting my voice to different situations of vocal performances (room, distance, environment, audience size, theme and atmosphere).	.490	.492		
06. F: I feel rejected from music projects because of my voice		.471		
23. E: I find that others do not understand my voice difficulties.			.579	
03. F: I avoid singing with an accompaniment (piano, orchestra, instrumental group).			.502	
25. E: I feel devalued and belittled because of my voice			.421	
09. E: My mood is negatively affected by my voice difficulties.		.354		.810
08. F: I lose income due to my voice difficulties.				.777
30. E: It happens that I feel discouraged when I think of my voice problems.				.648
28. E: I think about my voice difficulties even when I do not sing.	.480			.596
24. E: Singing makes me tense and stressed.				.594
27. E: I become anxious with the thought of singing.			.428	.585
29. E: It happens that I refuse to sing.				.571
10. P: My voice gives way intermittently.	.362			.485

Table 6. The pattern matrix. This table contains the results of the PCA analysis which defined a preferable structure of four components (or factors). In the event that an item scores twice, the larger number is retained for group categorisation. For the sake of comprehension the Swedish items have been informally translated to English. Factor 1 tends to group most functional (F) and physiological (P) items while Factor 2 and 3 group factors that best related to performance issues and the perceptions of others. Factor 4 mostly groups self-perceived emotional items. Some exceptions like items 8 are difficult to explain.

Correspondence: Anick Lamarche (M.Mus), Fax: 08-790-7854. E-mail: anick@nada.kth.se
(In review)

optimistic since choristers are likely to participate in more than one choir. With a total of 126 singer subjects, this study is a rather well suited representation of the Swedish singing population.

It follows that there was a high choir representation in subjects (see table 1a-b). This cultural aspect of Swedish music involvement brought forth some interesting subject comments in relation to the items of the RHI-s. For example, choristers repeatedly perceived the items 3, 5 and 16 as irrelevant. These items addressed the audibility of the voice with accompaniment as well as the performance context flexibility of the voice. Indeed, these elements have soloist characteristics. Nonetheless, they can also be generalised to choir singing (accompaniment could also mean other singers and performance context could mean changes from a dry rehearsal room to a reverberate church). Some modifications to finely tune these details could have been made in the RHI-s but on the other hand, the variable associated to the context of singing showed no particular effect on overall or even subscale results. Consequently, modifications would not contribute meaningfully to end results.

Some other comments were noted in the test-retest process and were traced mostly to item 11. Here, the item wording was questioned; "kan inte"-not able to-, was thought to have a too strong connotation and "har svårt att"-have difficulties to was suggested instead. Consequently, the semantic weight of "kan inte" could have limited RHI-s responses concerning that item.

The homogeneity of the subject distribution was difficult to control since the patient data collection was not directly performed by the investigators. However, the sex and singing level proportions were quite similar in both groups. Great distribution differences concerned mostly the singing genre and singing context variables. While the subjects of the control group were predominantly classical choir singers, patients mostly sang styles like pop, rock, soul, jazz and afro and were active as soloists. The same patient trend can be noted in Cohen *et al.* (12). Surprisingly, the asymmetries in population demographics do not have influential weight in the results. In fact, no other effect than an effect of sex was observed. These results were somewhat puzzling in that a professional singer who suffers from a voice disorder would be expected to be strongly impacted by voice anomalies (namely in terms of income, career, reputation etc). That results did not support this kind of expectation, can suggest that the general impact of voice disorder on any singer is fairly equivalent in that the act of singing is hindered.

Furthermore, the dominating demographical characteristic of the patient group is certainly not negligible. Indeed, here lies very important information for building better prevention programmes and tailoring voice care to singer-patients. Such patient data indicates the need for further research on singing genres belonging to the Contemporary Commercial Music (CCM) and warrants further attention.

When variables such as sex, age, genre, level and context of singing were examined, the only effect that could be found was an effect of sex. Female scores differed distinctly between controls and patients whereas such differences could not be reported for control or singer-patient males. This difference was observed only for total RHI-s scores and not at the subscale level. Among many other variable effects, Morsomme *et al.* reported a similar effect. They found that overall and subscale scores were higher for female subjects (15). In the clinical reality, there tends to be a higher prevalence of female voice complaints yet, due to this study's balanced distribution between the sexes, this should not have impacted seriously the results presented here.

Data Collection

The collection of test-retest data typically poses

Correspondence: Anick Lamarche (M.Mus), Fax: 08-790-7854. E-mail: anick@nada.kth.se
(In review)

many challenges and in this light, the RHI-s was made as accessible as possible to both subject groups. Hard copies were distributed and an electronic version of the questionnaire was made available on internet. Surprisingly, the internet response rate was very low and hard copies remained the best option for tests and retests. The electronic version of the questionnaire was perhaps not as convenient as intended since it required Adobe Reader. Moreover, in view of keeping a high level of confidentiality, questionnaire answers could not be downloaded or stored and had to be sent by email. To facilitate score computation of the internet RHI-s, the VAS scale was converted to a series of radio button placed on a 10 cm line. This modification was inspired by the fact that the VAS scale has been formally converted to an 11-point Likert scale numbered from 0 to 10, where patients are asked to place an "X" on the number line(23).

In spite of the intent to facilitate the subject's involvement, the steps involved were perhaps too cumbersome. Data collection difficulties could not be avoided in the retest phase, especially in what entailed patient data collection. This follows what others have previously reported (i.e., Cohen *et al.* report a loss of 50% of the data). A solution to such data losses would be greatly valuable for future investigation like this one.

Reliability and Validity of the Scales

The inclusion of the VAS scale followed the example of the work elaborated by Cohen *et al.* (12). Such an addition to the RHI-s could allow a more rigorous evaluation of internal coherence. In the event that patients scored highly on the RHI-s it was important that they would also highly scale the overall impact of voice pathology. Here, we found that correlations between the RHI-s and VAS results were very high. The questionnaire seems to successfully capture the subject's perceptions regarding a voice disorder. The

addition of the VAS scale was felt to be only pertinent to the patient group testing. It was assumed that the control group, in the absence of vocal health problems, would respond in very low scoring of the scale and therefore results obtained for patients would be sufficient to assess the internal coherence of the instrument.

A critical difference of 16 points and 6 points for the functional subscales and 7 points for the physiological and emotional subscales were delimited. This falls in line with previous VHI reports (2, 4). These critical differences can help in elucidating the voice status changes from the time of diagnosis to the time of therapy completion.

A small learning effect was observed. Retest scores were generally lower (here it seems like subjects learned from the first administration and adjusted their answers accordingly on the second administration). Morsomme *et al.* reported a similar observation (15). This trend was however so minimal that it did not affect significantly the reliability of the test-retest which was quite high.

All in all it was observed that results of validity and reliability were generally higher for the patient group. Such an observation only further denotes the robustness of the RHI-s as an instrument for the evaluation of voice disorder impact on the singer.

Cronbach's alpha results:

In 1978, Nunnally recommended an alpha coefficient of .7 as an acceptable reliability coefficient. Yet, in 1994, he adjusted is recommendation to a minimum of .5) (20-21). The analysis work of Morsomme *et al.* based itself on Mesbah *et al.* where a .7 minimum is recommended (15). Recommendation for a decisive minimum value is perhaps difficult in the sense that the Cronbach's alpha is a relative value which depends on the objective of the analysis.

Three American studies on the psychometric

evaluation of disease specific quality of life instruments in voice disorders report high internal consistency reliability and test/retest reliability criteria (ranging between .9 and .95)(8). In our case, for a group comparison level by subscale, a minimum of $\alpha = .5$ was certainly too modest and an $\alpha = .7$ or higher alpha value was considered more justifiable. Morsomme *et al.* reported an $\alpha > .80$ for all subscales and Cohen *et al.*, an alpha of .97 for the total score (the SVHI does not contain any subscales).

In our study all of the Cronbach's alpha results (for total scores as well as for subscale scores) were very high and fell in line with the proposed American standards for reliability (see Table 5 and 6).

An important note needs mention in respect to subscales and Cronbach's alpha results. Despite the high alpha obtained for the subscales, a PCA was performed to further investigate the usefulness of the three subscale structure. Cohen *et al.* (12) discarded this structure entirely upon results of a principal component analysis. They however, did not expose the results of this analysis. In the current study, the results of the PCA suggested that perhaps another categorical structure would best fit the current instrument. Unexpectedly, this finding did not corroborate with the reliability coefficients found for those subscales.

The results of the PCA were most interesting as they demonstrated that the items of the RHI-s might not necessarily divide according to the subscale structure that has been used until now. 4 factors rather than 3 seem to best explain the categorisation of items. When these factor groups were further investigated, the factors that seem the most pertinent in differentiating patients from controls were factors 1 and 4. These two factors grouped mostly items that were concerned with 1) physiological and functional concepts and 2) self-perception or internal related emotional items. Factors 2 and 3 appeared to group items mostly related to

performance or interpretation (aesthetics) issues and external emotional items (pointing to the perceptions of others). That factor 2 and 3 do not seem to weigh significantly in the discrimination of singer-patients and controls simply underlines that the items grouped by these factors reflect a common concern of the singer *regardless* of the vocal state. A change of vocal state does not seem to accrue the individual's reaction to these items.

Cut-off score

Since the RHI-s is one of many measures included in the clinical voice evaluation of a patient, the low specificity results (.76) obtained in setting a differentiating cut-off score is not terribly consequential. With a score of 31, it was demonstrated that the RHI-s could classify singer-patients with 100% accuracy (sensitivity was 0.1). However, this classification also entailed a certain number of true non positives. To be more accurate, one fourth of the healthy population could not be correctly identified. In the framework of the clinical evaluation, these true non positives would most likely be identified by other measure outcomes and so the ability of the RHI-s to correctly identify singer-patients remains an interesting result. Cut-offs, like the one proposed here, are never without error but they can prove nonetheless useful in roughly gagging the voice status of a patient.

Conclusions

When given the RHI-s, singers with voice problems obtain significantly different scores than healthy singers. Some item scores, related to physiology, function and internal emotional states seem to be most important in defining the singer-patient's voice handicap. Variables such as voice classification, singing context and level of performance as well as singing genre had unexpectedly little impact on the healthy and the singer-patients scoring of the RHI-s.

In contrast, an effect of sex was observed; greater differences were found for female

healthy and singer-patients than for men. Moreover, this study brought forth interesting and important details concerning the singer-patient population. High representations of soloist of classical as well as other CCM styles corroborated recent findings of other similar studies. It can be concluded that the RHI-s is a valid and a reliable instrument and deserves to be included in the overall voice assessment of the singer-patient.

Acknowledgements

References

1. Organization World Health. International Classification of Impairments, Disabilities and Handicaps. Geneva: World Health Organization, 1980.
2. Jacobson BH, Johnson A, Grywalski C, et al.. The Voice Handicap Index (VHI) Development and Validation. *Am J Speech Lang Pathol*. 1997; 6:66-70.
3. Fulljames N, Harris S. Voice outcome measures: correlations with patients' assessment of their condition and the effectiveness of voice therapy. *Logoped Phoniatr Vocol*. 2006; 31:23-35.

4. Wheeler KM, Collins SP, Sapienza CM. The relationship between VHI scores and specific acoustic measures of mildly disorded voice production. *JVoice*.2006; 20:308-317.
5. Woisard V, Bodin S, Yardeni E, Puech M. The VHI: correlation between subjective patient response and quantitative assessment of voice. *JVoice*. 2007; 21: 623-31.
6. Biddle AK, Watson LR, Hooper CR, Lohr KN, Sutton SF. Criteria for determining disability in speech-language disorders. *Tech. Rep. 52*, Agency of health care research and quality, 2002.
7. Franic DM, Bramlett RE, Bothe AC. Psychometric Evaluation of Disease Specific Quality of Life Instruments in Voice Disorders. *JVoice*. 2004; 19(2):300-315.
8. Verdonck-de-Leeuw IM, Kruik J, DeBodt M, Guimaraes I, Holmberg EB, Nawka T, Rosen CA, Schindler A, Whurr R, Woisard, V. Validation of the Voice Handicap Index by assessing equivalence of European translations. *Folia Phoniatr*. 2008; 60(4): 173-178.
9. Schuster M, Lohscheller J, Hoppe U, Kummer P, Eysholdt U, Rosanowski F. Voice handicap of laryngectomees with tracheoesophageal speech. *Folia Phoniatr*. 2004; 56:62-7.
10. Zur KB, Cotton S, Kelchner L, Baker S, Weinrich B, Lee L. Pediatric Voice Handicap Index (pVHI): a new tool for evaluating pediatric dysphonia. *Int J Pediatr Otorhinolaryngol*. 2007;71(1):77-82.
11. Morsomme D, Simon C~H, Jamart J, Remacle M, Verduyckt I. [A proposal to adapt the voice handicap index to the singing voice]. *Rev Laryngol Otol Rhinol (Bord)*. 2005;126:305-313.
12. Cohen SM, Jacobson BH, Garrett CG, *et al.*. Creation and validation of the Singing Voice Handicap Index. *Ann Otol Rhinol*

Laryngol. 2007;116 (6):420-6.

13. Hogikyan ND, Rosen CA. A review of outcome measurements for voice disorders. *Otolaryngol Head Neck Surg*. 2002; 126(5): 562-572.

14.   Rosen CA, Murry T. Voice handicap index in singers *JVoice.* 2000;14:370-7.

15.   Morsomme D, Gaspar M, Jamart J, Remacle M, Verduyckt I. [Adaptation du Voice Handicap Index à la voix chantée] Voice Handicap Index adapted to the singing voice Rev *Laryngol Otol Rhinol* 2007;128(5):305-314.

16.   Murry T, Zschommler A, Prokop J. Voice Handicap in Singers. *JVoice.* 2008 [Epub ahead of print].

17.   Hogikyan ND, Sethuraman G. Validation of an instrument to measure voice-related quality of life (V-RQOL). *JVoice*.1999; 13(4):557-69.

18.   DeBodt MS, Wuyts Fl, Van De Heyning PH, Croux C. Test-Retest Study of the GRBAS scale: Influence of experience and professional background on perceptual rating of voice quality. *JVoice*. 1997;11:74-80.

19.   Wuyts Fl, De Bodt MS, Van De Heyning PH. Is the reliability of a visual analogue scale higher than an ordinal scale? An experiment with the GRBAS scale for the perceptual evaluation of dysphonia. *JVoice*. 1999;13:508-517.

20.   Nunnally JC. Psychometric theory , New York: McGraw-Hill 1978.

21.   Nunnally JC, Bernstein AH. Psychometric Theory (3rd ed.). New   York: McGraw-Hill   Series   in Psychology.1994.

22.   Giovanni A. Le bilan d'une dysphonie. Etat actuel et perspectives *Solal.* 2004: 245.

23.   Ma E, Yieu E. Voice activity and participation profile: assessing the impact of voice disorders on daily activities *J Speech Lang Hear Res*. 2001; 44:511-524.

Appendix A: The RHI-s form

DIAGNOS:

RÖSTHANDIKAPPINDEX (RHI) anpassat för sång

Anick Lamarche baserad på Morsomme et al 2005- [A proposal to adapt the voice handicap index to the singing voice]Rev Laryngol Otol Rhinol (Bord).

Här ser Du ett antal påståenden som många andra har använt för att beskriva sin röst och hur rösten påverkar deras liv.
Kryssa i rutan för det svar som visar hur ofta Du upplever samma sak.
Om inte frågan är aktuell för Dig: kryssa i rutan för "Aldrig".

		Aldrig	Nästan Alldrig	Ibland	Nästan Alltid	Alltid
1	F: Jag har svårigheter att byta från ett register till ett annat.	○	○	○	○	○
2	P: Det känns som att jag måste anstränga mig för att sjunga.	○	○	○	○	○
3	F: Jag undviker att sjunga med ackompanjemang (piano, orkester, instrumentalgrupp).	○	○	○	○	○
4	P: Jag har svårt att genomföra mina fraseringar.	○	○	○	○	○
5	F: Min röst har svårt att göra sig hörd över ackompanjemang (piano, orkester, instrumentalgrupp).	○	○	○	○	○
6	F: Jag känner mig utanför i musikprojekt på grund av min röst.	○	○	○	○	○
7	E: Jag känner mig oroad av röstsvårigheterna.	○	○	○	○	○
8	F: Jag förlorar inkomst på grund av röstsvårigheterna.	○	○	○	○	○
9	E: Mitt humör påverkas negativt av röstsvårigheterna.	○	○	○	○	○
10	P: Min röst försvinner nästan emellanåt.	○	○	○	○	○
11	P: Jag kan inte nå mina höga toner när jag sjunger.	○	○	○	○	○
12	F: Min röst är instabil (försämras under en övning eller en sång).	○	○	○	○	○
13	P: Min talröst är sämre efter att jag har sjungit.	○	○	○	○	○
14	P: Jag känner ett obehag eller en smärta i struphuvudet när jag sjunger.	○	○	○	○	○
15	E: Jag är missnöjd med min röstfärg (timbre, täthet, kärna).	○	○	○	○	○
16	F: Jag har svårt att anpassa min röst efter olika förutsättningar för vokala prestationer (lokal, avstånd, miljö, publikstorlek, tema och atmosfär).	○	○	○	○	○
17	P: Min röstklang varierar (utan min vilja) under en och samma föreställning, repetition eller konsert.	○	○	○	○	○
18	P: Jag har läckage på min röst (för mycket luft i rösten).	○	○	○	○	○
19	F: Jag lyckas inte få en "bra röst" ens efter uppsjungning.	○	○	○	○	○
20	P: Det händer att det inte går att förutsäga om min röst kommer att låta klar eller inte.	○	○	○	○	○
21	P: Jag brukar behöva ta i mycket för att sjunga.	○	○	○	○	○
22	F: Jag har svårigheter att förmedla mina känslor när jag sjunger.	○	○	○	○	○
23	E: Jag tycker att andra inte förstår mina röstsvårigheter i sång.	○	○	○	○	○
24	E: Att sjunga gör mig spänd och stressad.	○	○	○	○	○
25	E: Jag känner mig nedvärderad och förminskad på grund av min röst.	○	○	○	○	○
26	P: Min röst känns klangfattig och torr.	○	○	○	○	○
27	E: Jag får ångest av tanken på att sjunga.	○	○	○	○	○
28	E: Jag tänker på mina röstsvårigheter även när jag inte sjunger.	○	○	○	○	○
29	E: Det händer att jag avstår att sjunga.	○	○	○	○	○
30	E: Det händer att jag känner mig uppgiven när jag tänker på mina röstsvårigheter.	○	○	○	○	○

Länk till internet blanketten: http://www.speech.kth.se/~anick/interact1.pdf

Var snäll och skriv tydligt

Namn

Födelsedag (åå-mm-dd)

Telefon nummer/mobil

Epost Adress:
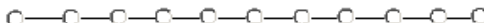
Kön

Datum (åå-mm-dd)

Sång Genre: (ett val bara)

Annat:

1. Nivå: (ett val bara)

2. Vilket sammanhang: (ett val bara)

Om Du upplever ett röstproblem just nu:
Hur svårt är ditt röstproblem?   Markera linjen nedan med ett streck för att indikera ditt röstläge idag.

Inget röstproblem

Svårt röstproblem

Om Du inte har Outlook kommer rutor upp för att hjälpa Dig att skicka iväg svaret.Följ noga menyn som visar sig när man klickar på "submit by Email". Spara först Ditt svar enligt menyns förslag som en .xml-fil och välj sedan vilket slags email-server Du vill använda (för yahoo, hotmail, thunderbird och gmail, klickar du på alternativ 2 i menyn).

*Det är viktigt att samma blankett fylls i igen om två veckor
och skickas in per mejl eller post.
Skriv genast datumet i din almanacka .
De data som samlas in är annars oanvändbara.*

Lycka till och tusen tack för Din medverkan!!

Kontakt Person Anick Lamarche
KTH-Tal Musik och Hörsel
Lindstedtsvägen 24 10044 Stockholm
anick@media.kth.se, Mobil: 07 68213018

Appendix A

Proposal for SRP, VRP$_{\text{phys}}$ and VRP$_{\text{perf}}$ acquisitions

These suggestions are based on the task designs presented in Chapter 4, section 4.4 of this doctoral dissertation. Calibration procedures are also described in this last Chapter. It is strongly suggested that the recordings be acquired with a linear or C weighted SPL measure.

A.1 Prior to the recording

- Ask the subjects to warm-up prior to the recording (Notation of the time and length of the warm-up as well as the time of recording can be useful in the event of future recordings).

- Instructions are given both, in written form and verbally (Key words and the order of the procedure can give the subject a framework and dissipate nervousness or anxiety).

- A quick orientation to the VRP with visual feedback should be included.

A.2 SRP

- Speak with the subjects freely inside the recording studio.

- Tasks should be performed in the subjects' native tongue

- Ask the subjects to describe a typical vocal warm up (using the same type of voice that was used in the conversation prior). One minute of speech is recorded. Contextual instructions are needed. An example is "pretend you are discussing this with a fellow singer in the hallway while waiting for a practice room" (It is useful to give a few more theme alternatives in the event

the singers should run out of things to say, keep these themes as emotionaly neutral as possible.)

- Ask the subjects to count from 20 to 40 in a soft voice (no whisper) yet as if they are putting a baby to sleep

- Ask the subjects to count from 40 to 60 as if speaking on the telephone

- The subjects count from 60 to 80 as if holding a seminar for a group of 50 persons or more. (Here, the singers can be called to visualise a typical seminar room).

- If the calling voice is of interest in the investigation, a short phrase can be used to call out. An example is: "Heh wait for me!".

Research has demonstrated that many factors can influence this type of speech based recording (the task material, the subject's emotional state and the emotional content of the task, the environment, the possible expectations of the investigator, and finally, the absence of an interlocutor). This is namely one of the reasons for which a reading task is not suggested (articulation and speech behaviour tend to be more posed and unatural).

A.3   VRP$_{\textbf{phys}}$

- Vibrato is not to be included and voice quality is disregarded

- An explanation of the motivation for these two exclusions can help the singers understand the nature of the task. The singers should be made aware of the interactivity of this task. It is useful to explain that several attempts are performed until the best possible complete VRP is obtained. Breaths can be taken whenever necessary. This also yields resting instances.

- The subjects choose a comfortable pitch and dynamic and phonate on a sustained [a]. (The investigator should take note of the selected pitch).

- This is repeated with the instruction to this time reduce the comfortable dynamic to a bare minimum ("barely any sound at all"). The singers need to understand that stable phonation is not expected. Demonstrations can be useful.

- From this point, the singers should descend in discrete pitch steps (chromatic scales are efficient) maintaining the same dynamic. The singers should be encouraged to phonate as low as possible and be reminded to sing as softtly as possible. A glissando exercise can be initially used to then return to discrete steps. The lowest pitch is repeated 3 times for reliability purposes.

- The pitch noted previously should be played to the singers and the same exercise is repeated for the higher part of the voice. It is useful to return to the same comfortable dynamic and make a crescendo to the very soft dynamic before starting the ascending steps.

- A sweeping type of phonation can help secure higher pitches (singers will tend to stop phonating near their typical tessitura limits). The singers are instructed to phonate on a short glissando and hold the last pitch. This ought to be repeated until the extreme high pitch is obtained 3 times.

- The same procedure is repeated in the extreme loud dynamic. Many demonstrations are useful here as well. The singers should be shown that register breaks are the goal. When the exercise is begun with glissandi exercises, the areas of laryngeal mechanism transitions are best detected. In order to help the singer initiate phonation in the desired voice mechanism, phrases like "No way!" can be used. The singers could state the phrase once and then repeat it, sustaining the last word. From this sustained word, a glissando could be initiated without changing the phonation. The bottom pitch should be repeated several times.

- Instead of returning immediately to the comfortable pitch to address the higher voice, it can be interesting to have the singers first make an ascending glissando. Voice breaks are sometimes easier to "catch" on an ascending task.

- Throughout the VRP$_{phys}$ recording it is necessary to remind the singer of the sound level goal (as little/much voice as possible). Contexts such as a baseball game, or a fast attraction ride or even winning the lotterie can help the singer think of a voice use that is typically "loud" yet excludes singing. Conversely, many demsontrations can be needed to bring the singers as close to a phonation threshold level as possible.

A.4   VRP$_{perf}$

- The recording only includes typical singing voice. It is best to avoid visual feedback in order to help the singers musicaly perform rather than compete with the screen.

- Ask the singers to perform according to "what is musically acceptable to them as an artist" and to use a dynamic range that is proper for stage performance with acompaniement (vibrato should now be included).

- Ask the singers to visualise their favourite performance venue and a reasonable size audience. (Singers are used to receive such instructions and to perform these types of visualisations during the course of training since practice rooms are typically small and lend themselves badly to stage realities).

- Vocalise instructions should be made available in notation as well as shortly demonstrated. The singers are instructed to perform this exercise as musically as possible (including phrasing, intent and stamina). A general pace is conducted at the onset of the vocalise only.

- The singers sing a comfortable pitch and dynamic. A descrescendo is performed to attain a stage soft dynamic. And the vocalise is performed in descending-ascending order. The singers should be reminded to respect their tessitura and to end the vocalise according to the lowest and highest pitches they would perform on stage.

- The exercise is repeated in a loud dynamic. It can be useful to remind the singers to visualise that they are accompanied by an orchestra. The procedure order mentioned above should be respected.

Appendix B

VRPs of Professional Female Classical Vocalists

Figure B.1: Physiological VRP (*glissando* task). From the top left, soprano subjects 15, 16, 18, 19, 20 and 22.
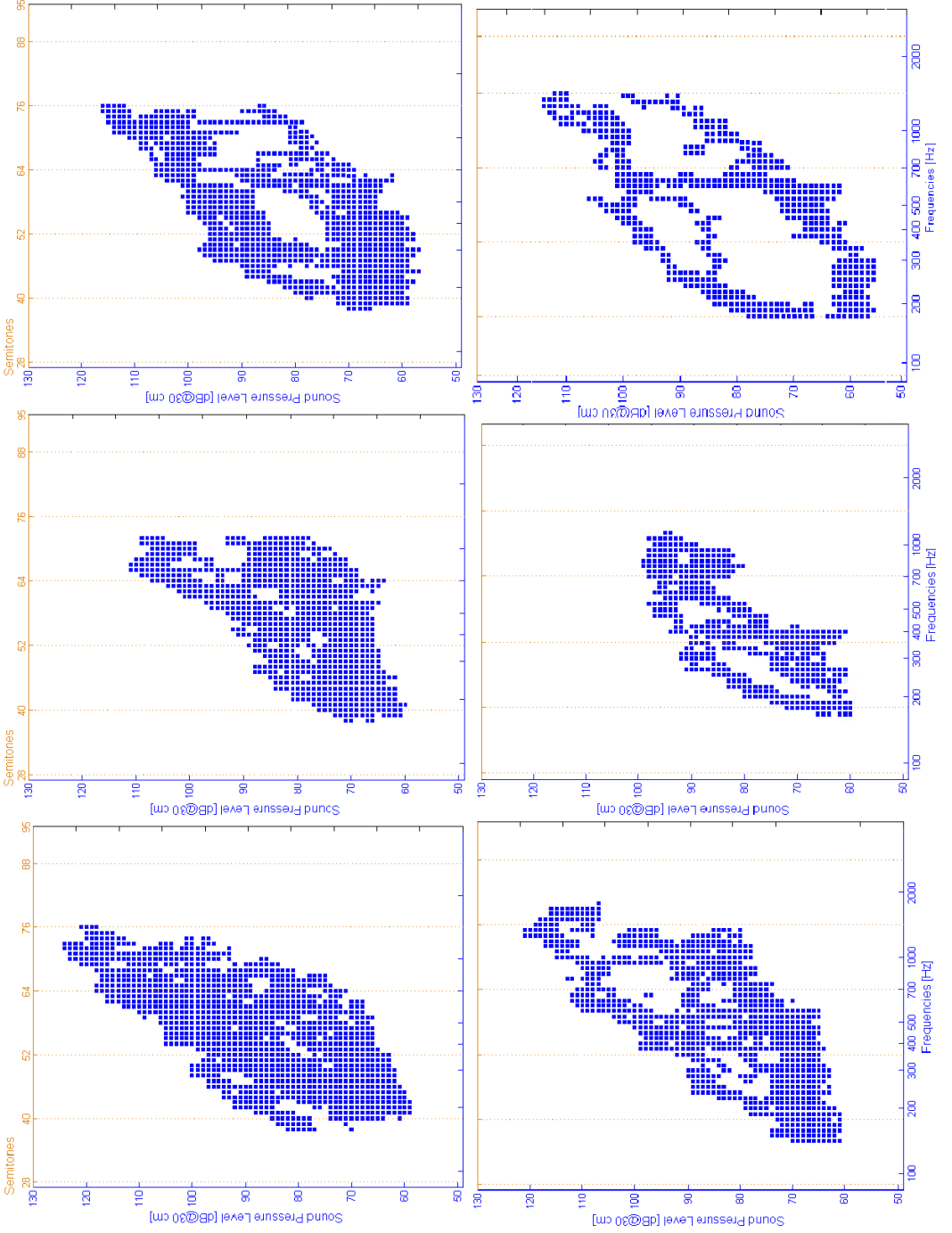
Figure B.1: Physiological VRP (*glissando* task). From the top left, soprano subjects 23 and 25.
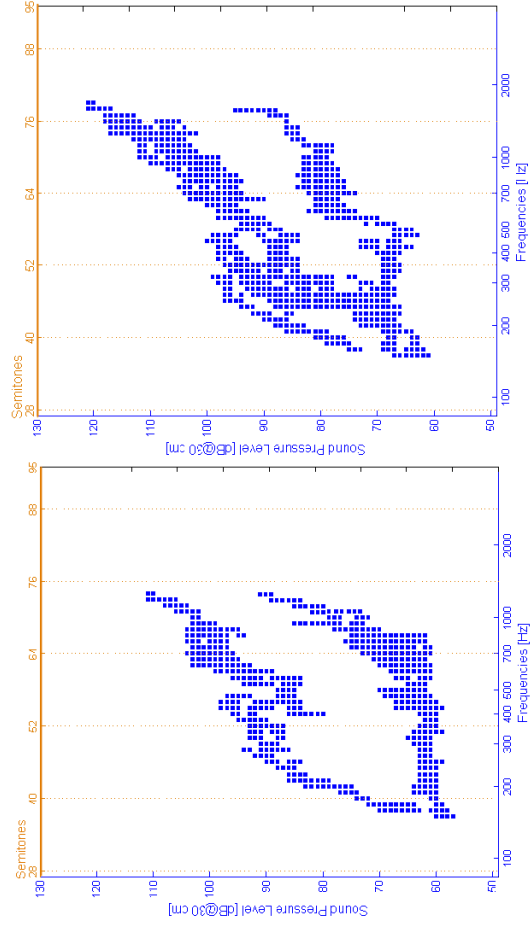
Figure B.2: Performance VRP based on the vocalise task. From the top left, soprano subjects 1, 2, 3, 6, 8 and 9.
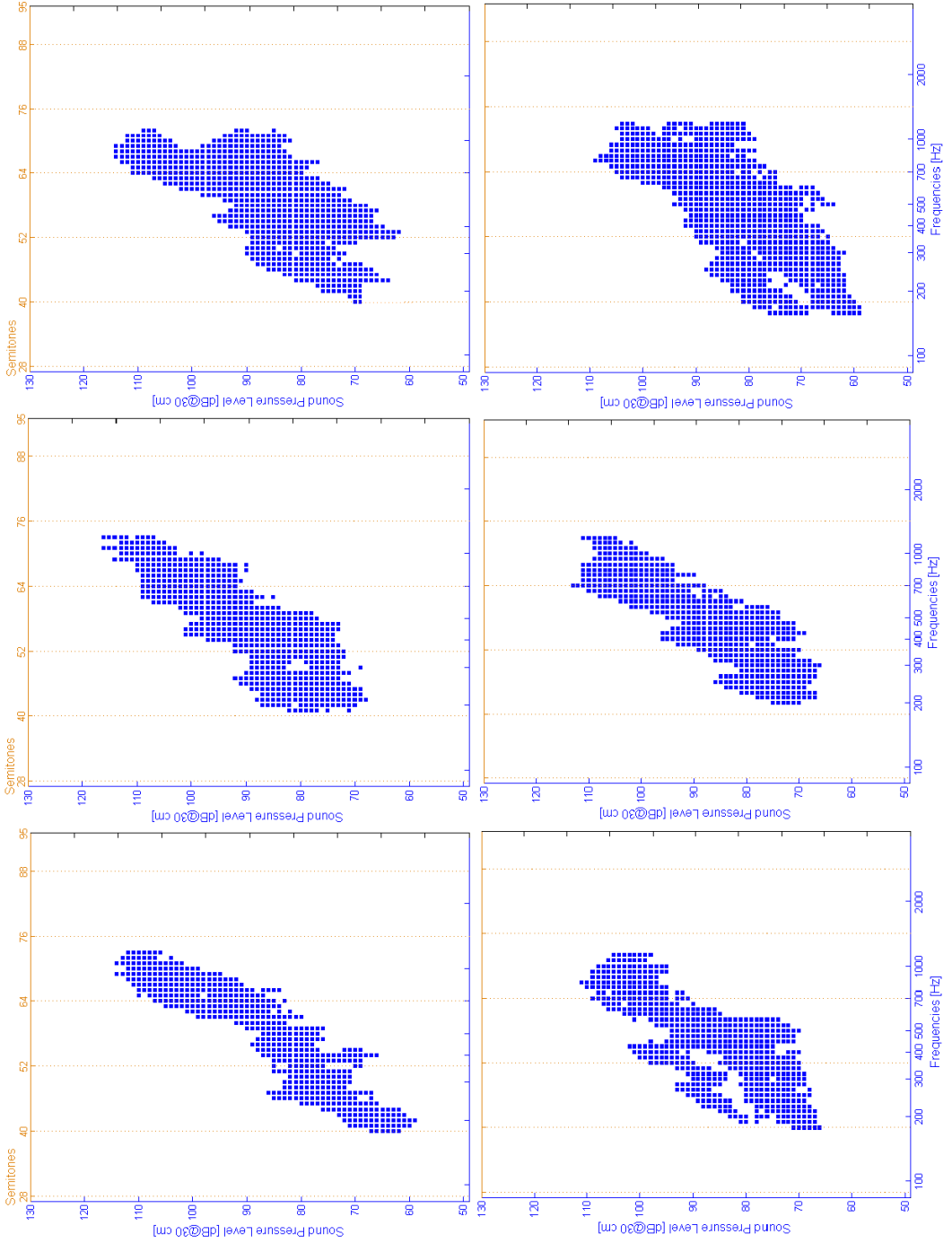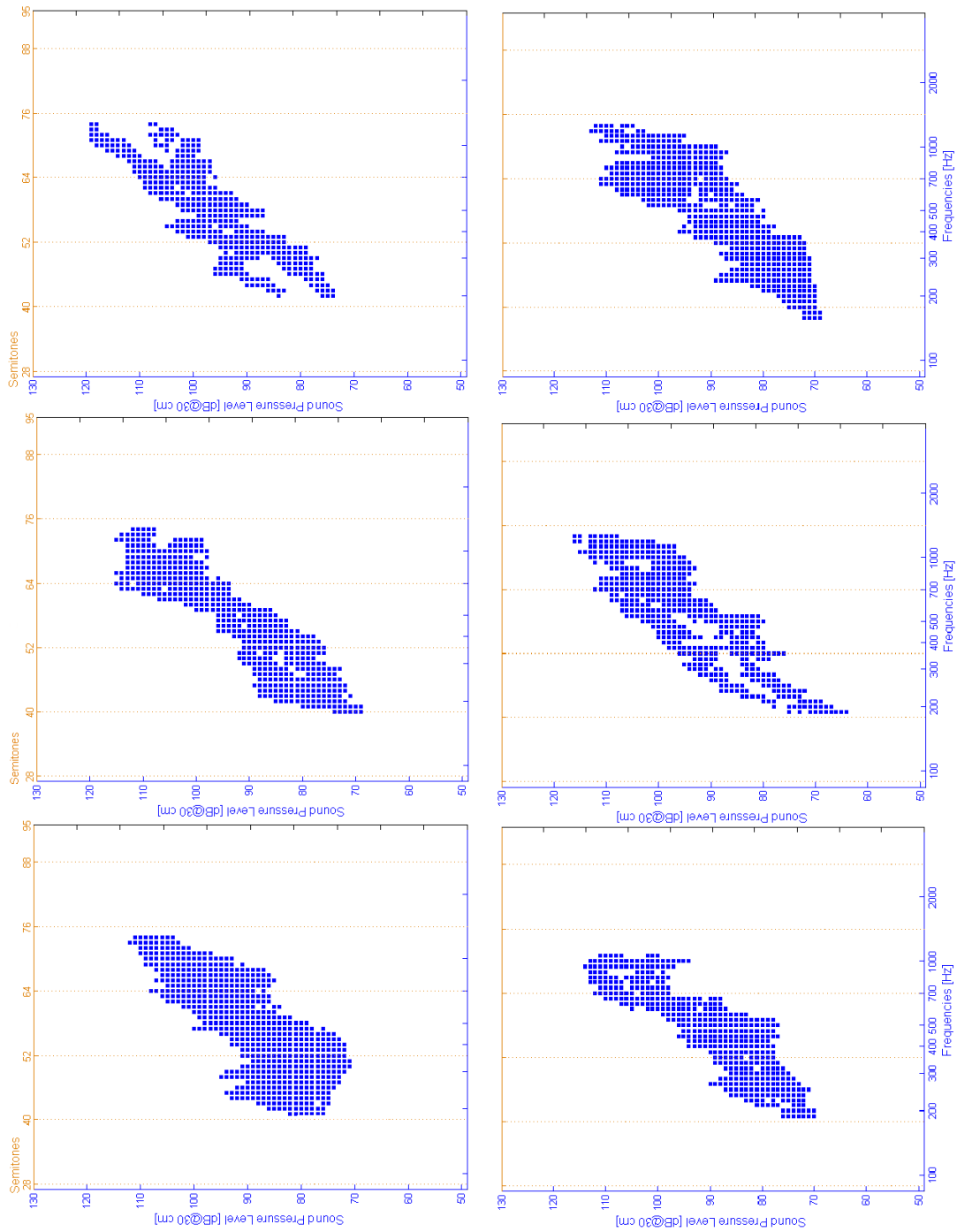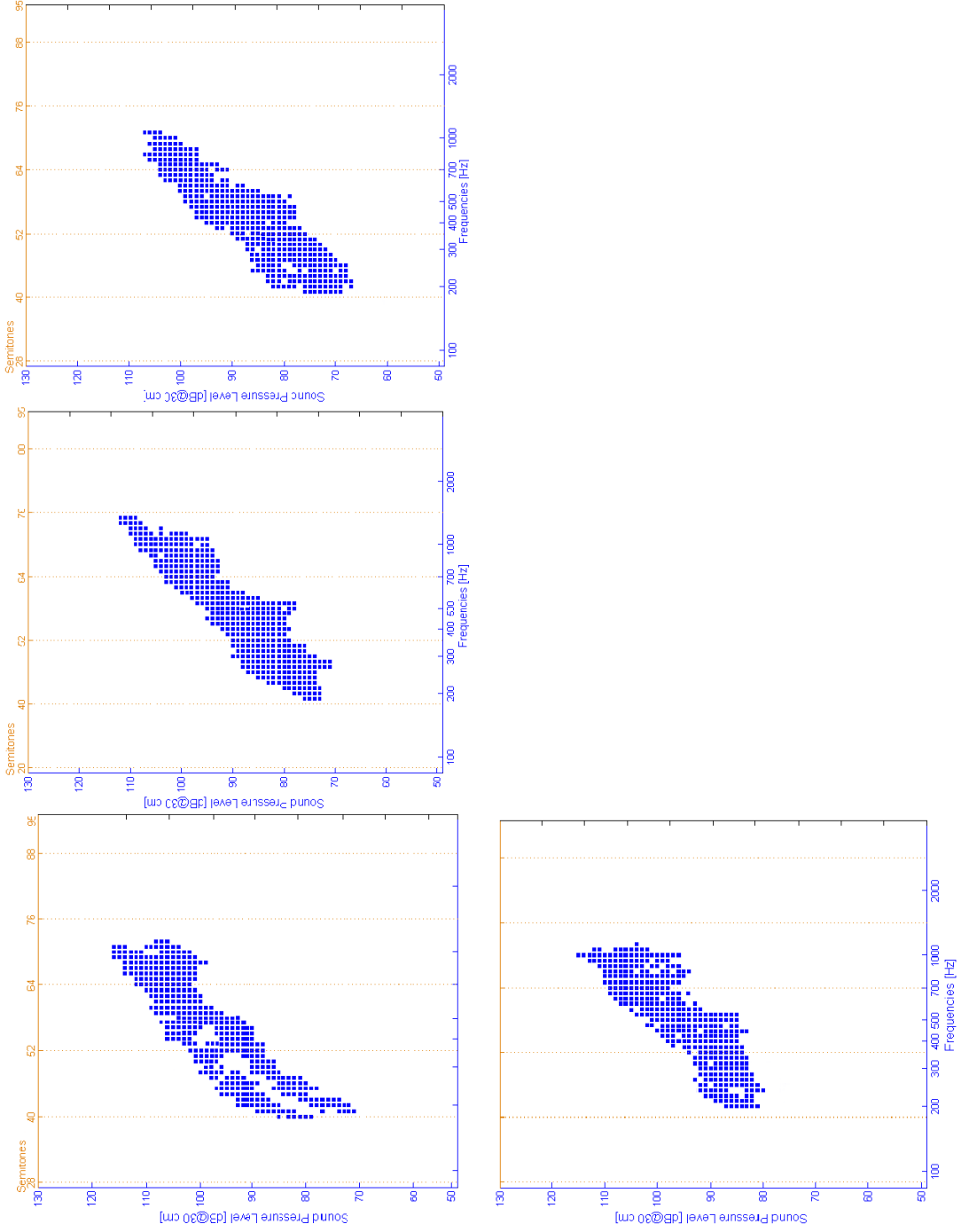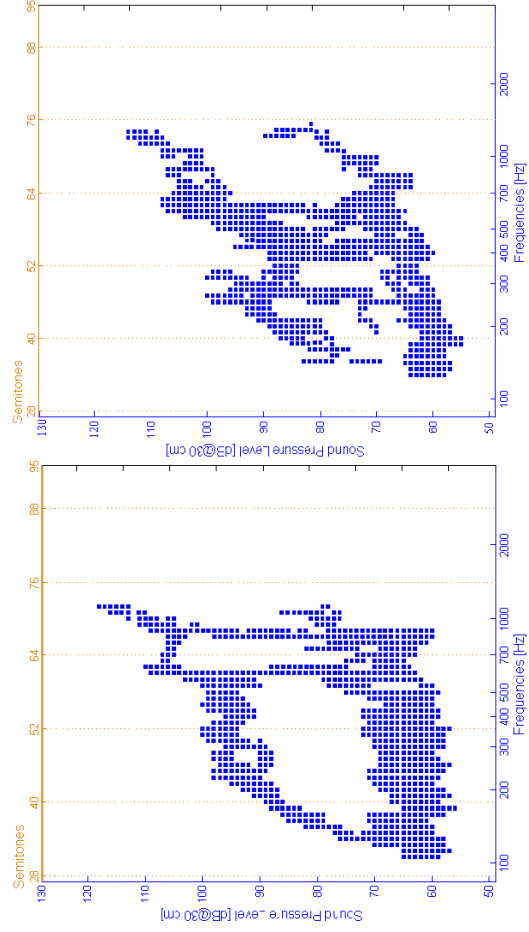
Figure B.3: Performance VRP based on the vocalise task. From the top left, soprano subjects 11,14,15,16,18 and 19.

Figure B.4: Performance VRP based on the vocalise task. From the top left, soprano subjects 20, 22, 23 and 25.

Figure B.5: Physiological VRP (*glissando* task). From the top left, mezzo-soprano subjects 21 and 30.

Figure B.6: Performance VRP based on the vocalise task. From the top left, mezzo-soprano subjects 4, 5, 7, 10, 12 and 13.
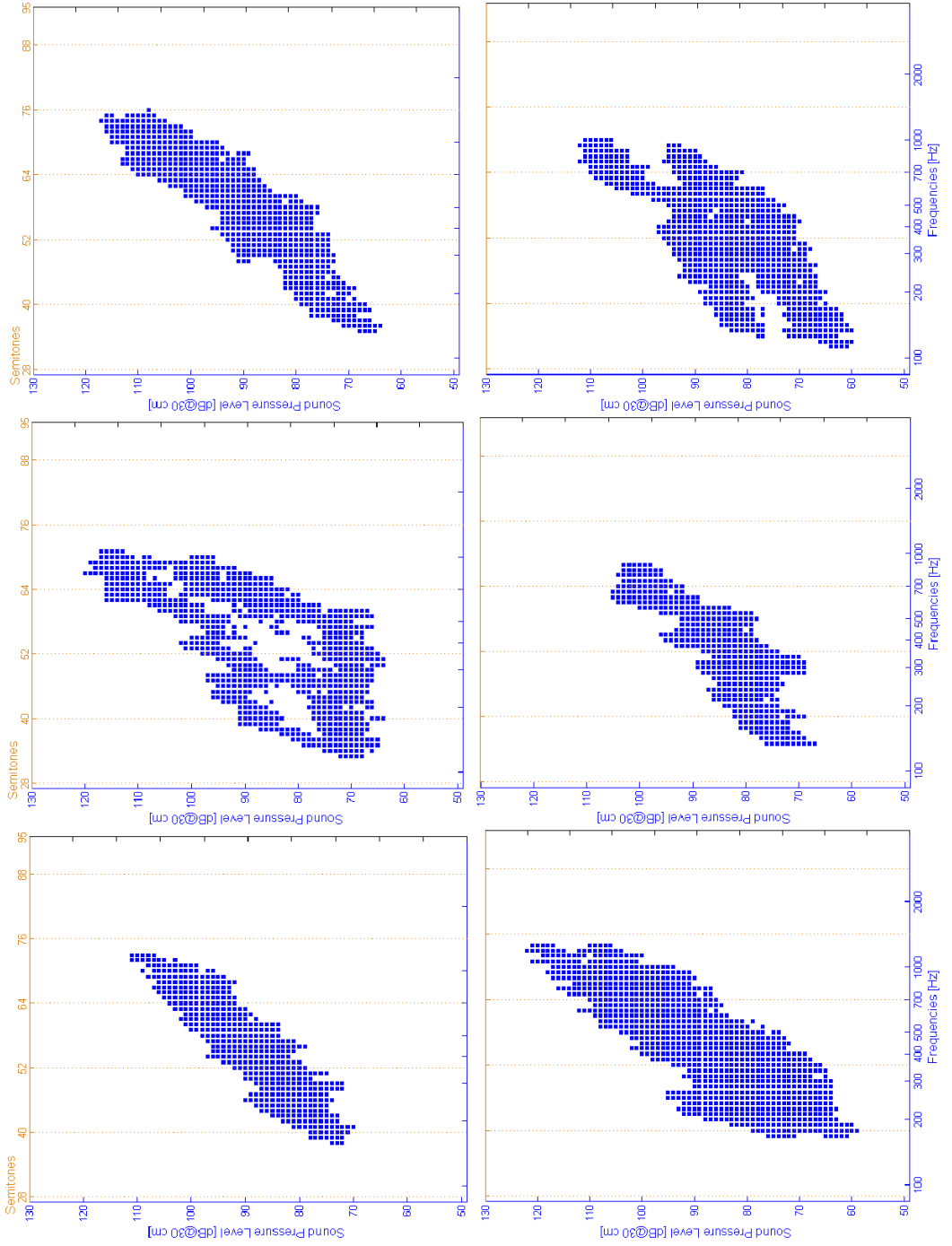
Figure B.7: Performance VRP based on the vocalise task. From the top left, mezzo-soprano subjects 21 and 30.
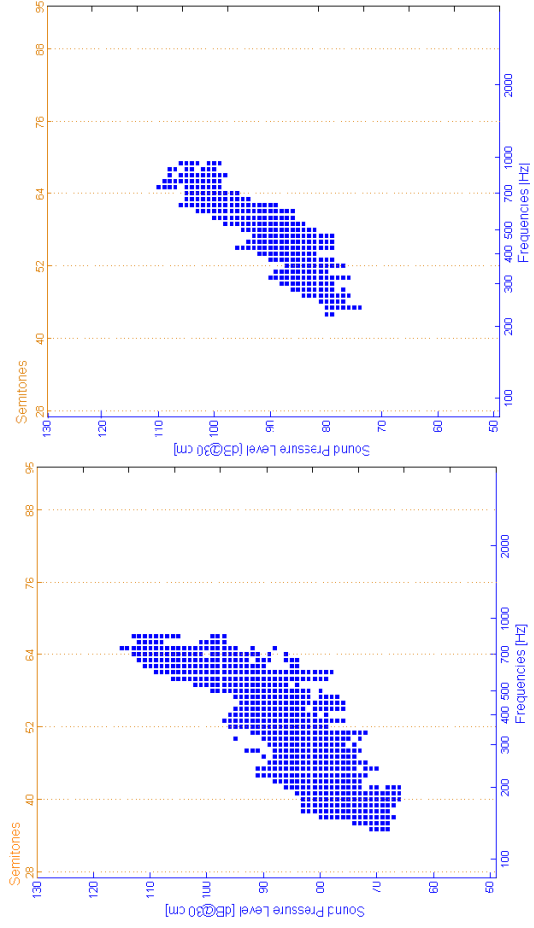
Figure B.8: Physiological VRP (the glissando task). From the top left, contralto subjects 17, 24, 26, 27, 28 and 29.
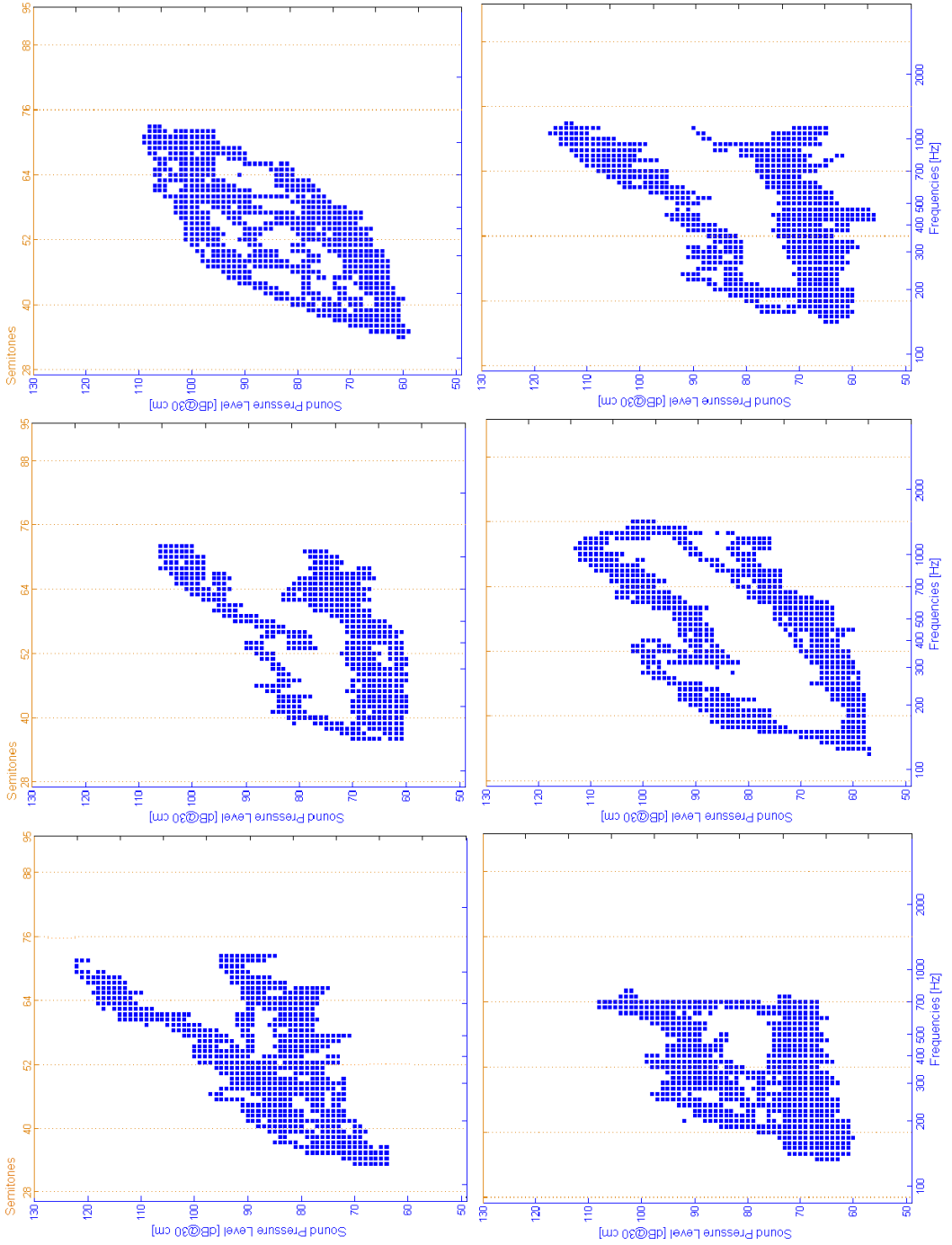
Figure B.9: Performance VRP based on the vocalise task. From the top left, contralto subjects 17,24, 26, 27, 28 and 29.