KTH Computer Science
and Communication

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

TRITA-CSC 2009:03
ISSN 1653-5723
ISRN KTH/CSC/A--09/03--SE
ISBN 978-91-7415-218-0

KTH School of Computer Science and Communication
SE-100 44 Stockholm
SWEDEN

Akademisk avhandling som med tillstånd av Kungl Tekniska högskolan framlägges till offentlig granskning för avläggande av filosofie doktorsexamen i tal-och musikkommunikation med inriktning på musikakustik fredagen den 06 mars 2009 klockan 10.00 i F-3, Kungl Tekniska högskolan, Lindstedtsvägen 26, Stockholm.

Tryck: Universitetsservice US AB

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

This work was made possible by the Canadian Baxter & Alma Ricard Foundation which funded 3/4 of the research. The work has also benefited from the financial assistance of the COST Action-2103. A short term scientific mission in Belgium was conducive to the data collection of Paper III. Also, gracious travelling scholarships and grants by the *Röstfonden*, the Canadian-Scandinavian Foundation and the KTH Knut och Alice Wallenbergs foundation rendered possible the presentation of this dissertation's work in different conference and academic forums.

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 pep-up emails and small notes supported me during this last part of the thesis journey. Our collaboration has developed in a truly great friendship that I hope will continue on in the years to come.

A note of appreciation is also in order for *Stellan Hertegård* who enthusiastically jumped on board in the first year of a voice health prevention project and participated in the realisation of Study II. His help with patient recruitment for Study V was also appreciated. I discovered that it is very practical to collaborate with efficient doctors...especially when laryngitis intervention are needed!

Without statistical help I would often have been lost. *Joachim Westerlund* was a significant (p< !!!) help and, in different ways, contributed to many of the studies presented here. Formally, his participation as a co-author in study V brought new and refreshing perspectives to my work.

Ingrid Verduyckt has contagious energy, a love for Sweden and embarked on my mission to translate the VHI for singers (from French to Swedish). I thank her for her dynamic look on life and her contributions to Study V.

I also wish to warmly thank all the other people invovled in study V: *Folke Alin,Gary Graden,Christina Hörnell,Mats Nilsson,Åsa Olsson,Helena Olsson,Roland Rydel, Anna Stenlund-Tyrén and Staffan Wilen.*

Many thanks to *Saven Hitech* people *(Lennart* and *Mohammed*) for lending DSP cards, AKG 420 microphones and troubleshooting my hopeless Acer! *Hans Larsson* for those DSP cards and cables exchange meetings at Södra Stationen, *Maria Södersten* and *Eva Holmberg* for frank opinions, support and good advice and those wonderful concert invitations!; *Mattias Heldner* for statistical discussions and book tips; *Christine Ericsdotter* for meticulously perusing the thesis and for providing frank and to the point comments-almost in a surprising unSwedish way (and for simply being great company!); *Becky Hincks* who volunteeringly and enthusiastically performed the last read through of the work before printing time and for North American solidarity; *Anna Hjälmarsson* reminded me that not only the actual thesis work was important. I am grateful that she made sure I would keep myself in shape and in the process also offered me an understanding ear. Finally, *Giampiero Salvi* was also a great help in the last crunch time when layout fixes had to be performed quickly and efficiently. He was able to step aside his own principles on subscript text in LaTex and helped me realise my non-mathematical abbreviation goals!

I wish to thank everyone at the department of Speech, Music and Hearing for contributing a certain "je ne sais quoi" dynamism to the workplace and for the energy and involvement that makes TMH such a distinct and special place! I will never forget the Formant orchestra moments, the Christmas traditions, the many institution meetings with goodies and lottery winners and especially the positive and fun entourage that exists in this place (and yes, those innebandy moments will stay with me but hopefully not the bruises!). I am particularly indebted to the Music Acoustic group (past, present and in-passing) who kindly welcomed me in their group: *Anders Askenfelt*, for all those "bonus" points and for those often late evening discussions on artistry and musicality; *Anders Friberg* for statistical talks and echoing laughter in the corridors (and his beautiful jazz playing); *Clara Maître* for help with useful functions in Matlab; *Erik Jansson* for his friendliness and kindly lending me a violin case; *Erwin Schoonderwaldt* for writing the SMPread function in Matlab, figure plotting help and smart tips...even bike expert reparation when in moments of complete bike handicap! *Eva Björkner* for sharing the ups and downs of being a singer amongst engineers! *Johan Sundberg* for lending me his books and manuscripts to my little heart's content and for attending to my many questions, often straightening out moments of confusion; *Gaël Dubus* for ensuring a continuing French touch in the group; *Kahl Helmer* for being so entertaining and sending me inspiring MMS from various vacation destinations, for persistently coming by every lunch to drag me away from the computer and of course for helping out with these small but oh so important button device boxes; *Kjetil Falkenberg-Hansen* for database ideas, LaTex tricks and for opening my eyes to the scratching world. Always ready to help and on the top of new practical software, Kjetil designed a great thesis cover! Other than simply sharing the same room, it was great to splurge on sushi, go obsessive with StarWars and sit in the nose bleeders of the opera house together; *Marco Fabiani* for always having a smart Matlab function to suggest, for dragging me to a Matlab workshop and for sharing his Italian home, family and Alps with me....not to mention all those little "ts" that make you so you!; *Roberto Bresin* for always being so good humoured, for ensuring regular social group outings and for being a model in networking and business oriented research thinking; *Sofia Dahl*, for her woman power talks, her versatility and her social company in certain of Stockholm's lounges; *Svante Granqvist* for being a sensitive guy and understanding my moments of frustration, for his methodology and pedagogical rigours, for providing a "green" working environment and be he blessed for all his smart software!

A kind and big thank you to all the singer participants involved in the experiments of this project. The work presented here completely depends on the quality and the expertise that the participants demonstrated and willingly shared with me. Your voices are golden and it was always a pleasure to spend some time in the recording studio with you!

I realise that many are left unmentioned and I wish to extend my appreciation to all those who have contributed in one way or another to this work (including smiles, jokes, hallway discussions, questions, listening, subtle comments and tips and company).

Although not as directly involved, many others have accompanied me along the way. I would like to mention my singing teachers: *Robert Miron, Jo-Anne Bentley* and *Rosemarie Landry*: they taught me so much about the human voice, singing and about the craft of music making.

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 Блгодар 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 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 squared of weight (in kg) to height (in meters). 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 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]

$\mathbf{F_0}$: 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 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 laryngeal 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 near the vocal folds (thyroid lamina) and the sternum bone (jugular notch).

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.

$\mathbf{VRP_{phys}}$ 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.

$\mathbf{VRP_{perf}}$: 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 clinical differences. 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 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 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 or a small open quotient, 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 VRP discussions.

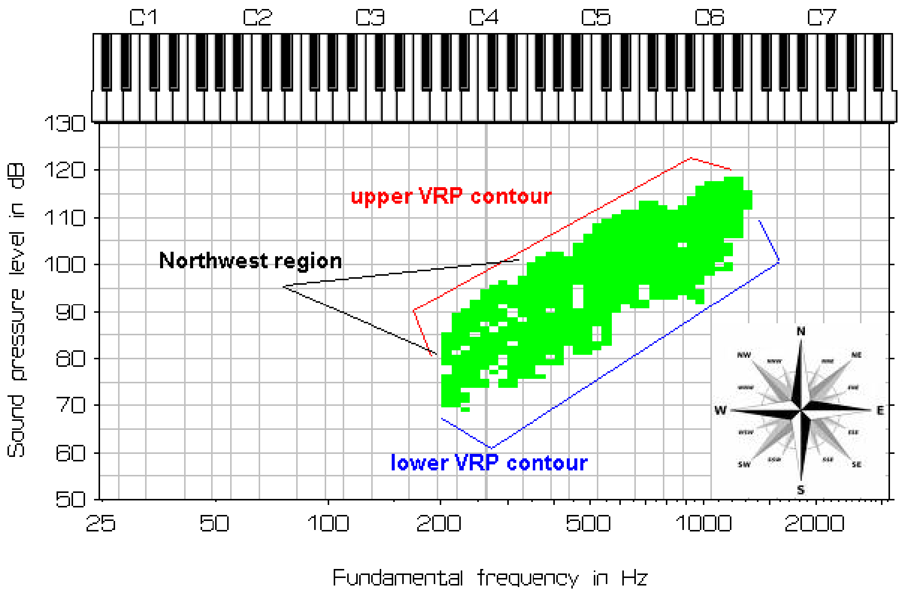


Figure 2.1: The VRP, a performance profile of a healthy lyric soprano,age 27 and active as a regional minor opera role singer. Note that, to accommodate the high SPL levels produced by female singers, the vertical axis was scaled differently than the usual VRP axis (40-120 dB).

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.

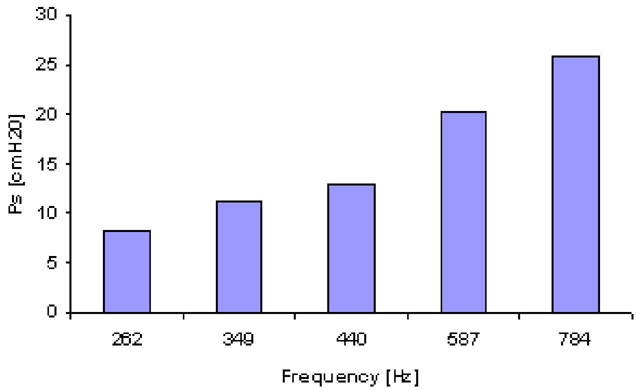


Figure 2.2: P_s as a function of frequency for a constant dynamic (*mf*). Results are collected for Western opera female singers (N=7), Lamarche 2006 (unpublished)

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_20 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

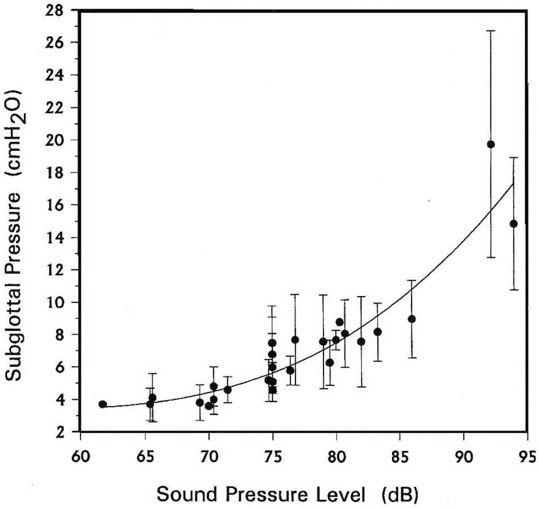


Figure 2.3: Mean P_s as a function of mean SPL summarised from 11 various reports of non-singer investigations (N ranging from 1 to 25). Figure taken from Baken [6]. The curve is a third-order regression fit to the mean P_s values. Error bars indicate the standard deviation for the group averages.

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 (± 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 psychoacoustic 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, an 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 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 throughout 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].

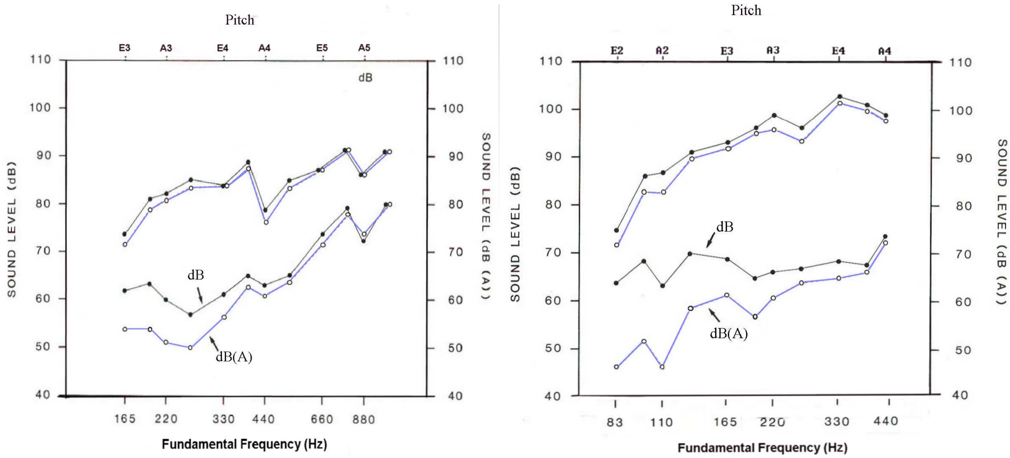
3.3 Standardisation

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 (± 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 ≈ 6 % (± 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

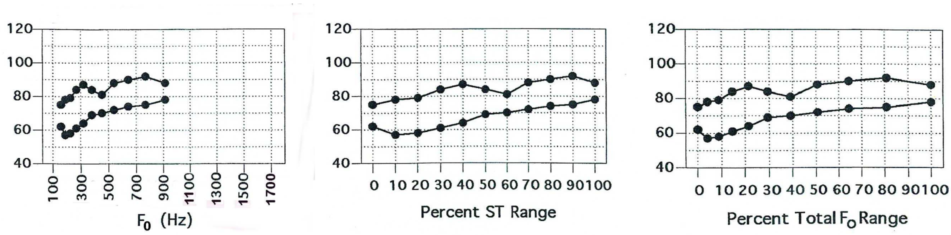


Figure 3.2: Taken from Baken, these figures depict well the visual transformation that occurs due to the normalisation of the VRP frequency scale. The data illustrated here is average data for female non-singers.

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 laryngeal mechanism transitions. The authors departed from their observations of laryngeal mechanism manifestations in the VRP. They claimed that two 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

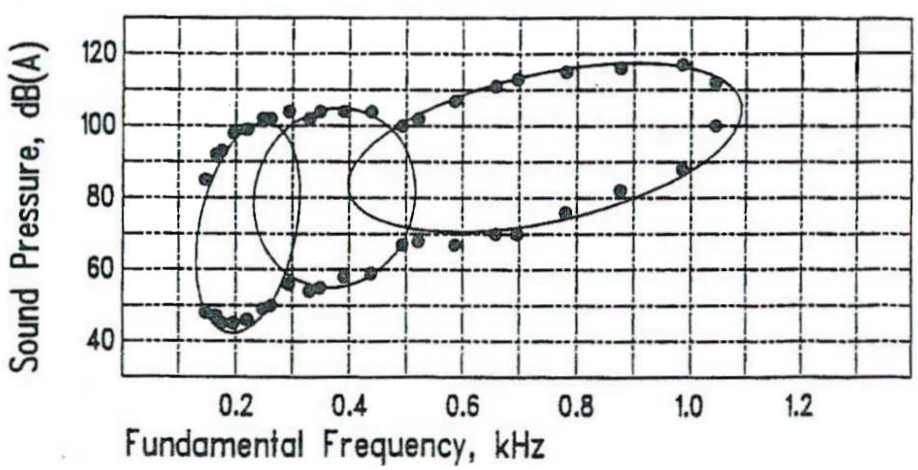


Figure 3.3: The fitting of ellipses in an attempt to apporixmate VRP contour points. In this example, three subareas are defined: the lower ellipse designates chest voice or M1, the middle ellipse, a transitory space or voix mixte and the upper ellipse, head voice or M2. Reprinted with permission [1].

himself reports a lack of reliability in the computerised method's detections of mechanism 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 el-

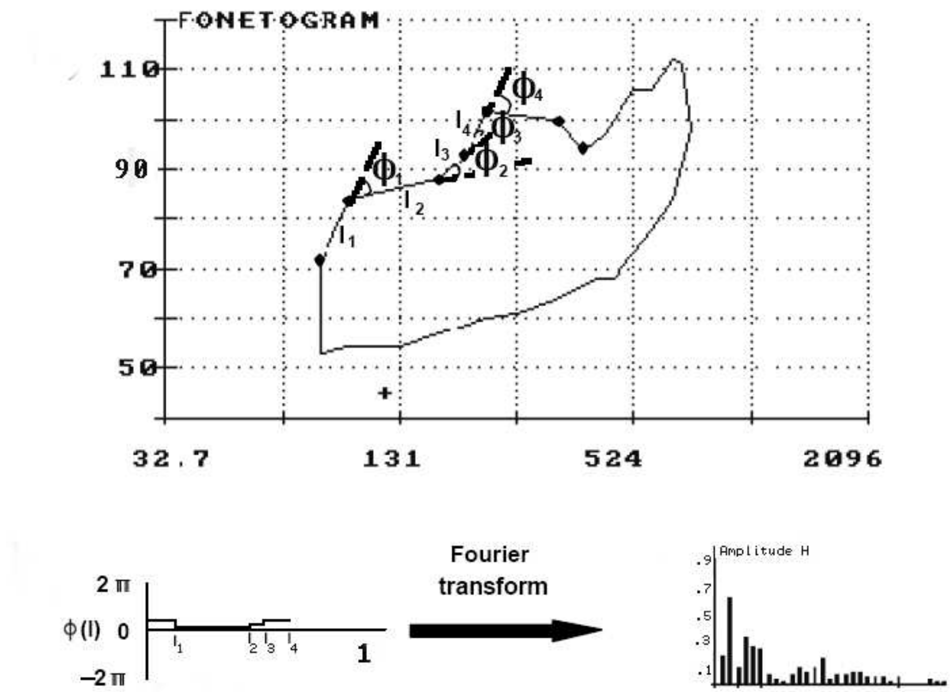


Figure 3.4: The Fourier descriptor approach in three steps, a) line lengths are calculated from lowest loud phonation to highest and angles between these line segments are calculated; b) line lengths and angles are plotted according to new axes; and c) The Fourier descriptors (general shape contributors 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].

liptical method: there is a dependence of the method on data acquisition. In this 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.

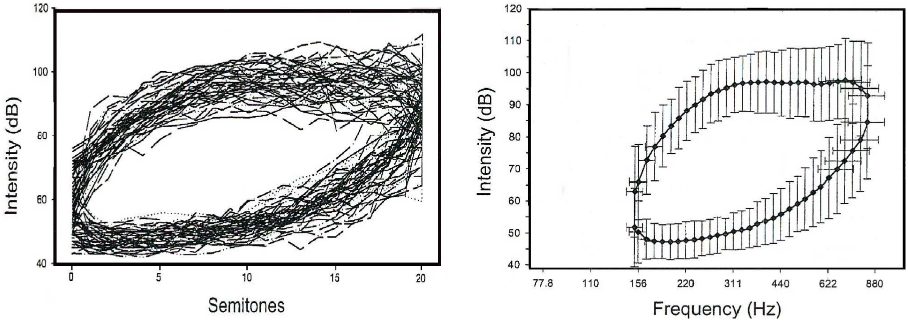
The idea was further developed by Pabon *et al.*[115]. In an automated setting, the number of times that a cell (1 ST by 1 dB wide) falls inside a phonation contour is tracked and this count is visualised by a colour map (the darkest showing the highest accumulation and the palest, the lowest). At the 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



(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.

any VRP shape can be brought to a common uniform base. This base can then be 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

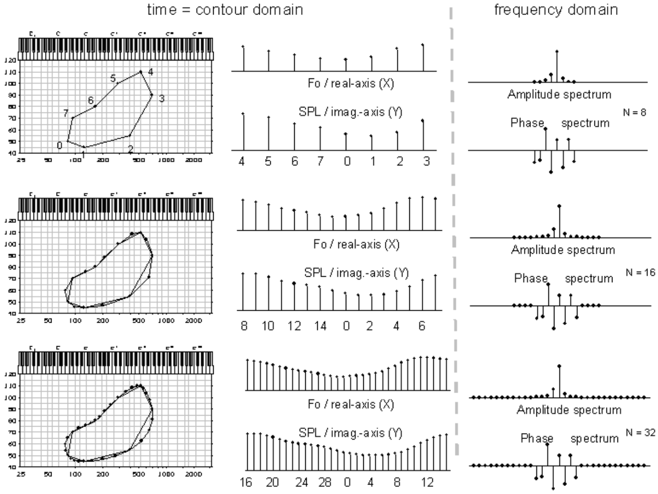


Figure 3.6: Interpolation by zero-padding. The figure displays an important part of the contour averaging by DFT. When contour points in the contour domain are numerous, the contour spectrum (mirroring the same number of points) will contain more high frequency components (+/-). Since these high-frequency 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 uniformily resampled to a predefined number of points. The figure is reproduced from Pabon *et al.* [114].

the study of non-singers. Data published on the VRP of singers is relatively scarce 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.

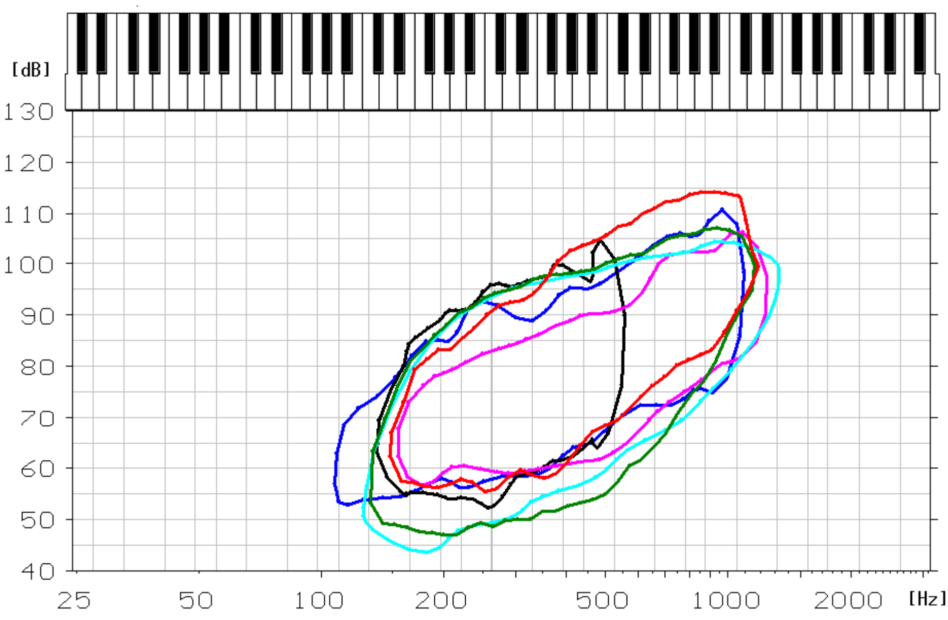


Figure 3.7: Singing voice VRP data compared (female voices only). **Sulter et al. in cyan** , N=42 trained females, Pabon in green, N=23 female singing students, **Åkerlund *et al.*in blue**, N=10 female professional singers,**LeBorgne & Weinrich in pink**, N=17 female graduate singing students, **Hacki 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.

$\mathbf{VRP_{phys}}$

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$_{\mathrm{perf}}$)

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$_{\mathrm{perf}}$ which will be henceforth adopted throughout the thesis. The VRP$_{\mathrm{perf}}$ is here viewed in a similar way as the SRP. It is a context-based and behaviour-dependent recording. The VRP$_{\mathrm{perf}}$ 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 $1\mu W$	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_0varies 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_0range, minimal/maximial 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.

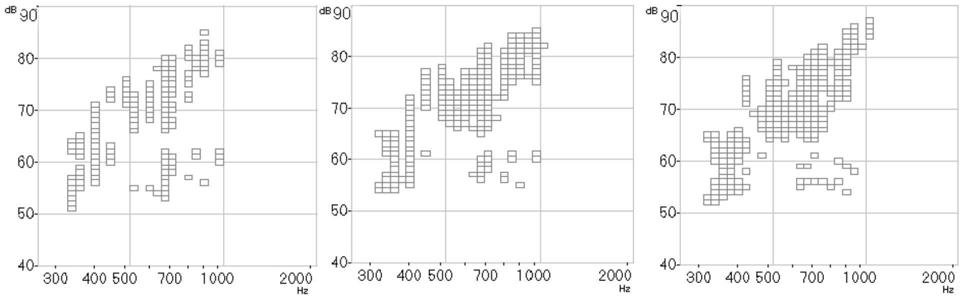
4.1 Voice Measurements

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

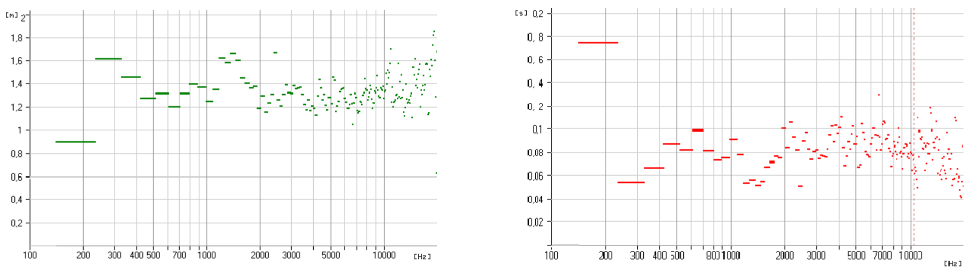


(a) no vibrato (b) ± 45 cents vibrato (c) ± 90 cents vibrato

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 first 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	

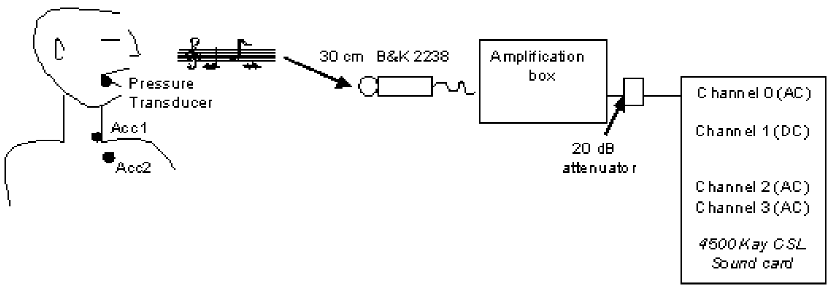


Figure 4.3: Experiment Set-up for the audio, aerodynamic and accelerometric measurements reported in Paper II. Acc1 indicates the position of the accelerometers at the jugular notch and Acc2 the sternum bone.

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 vocal folds [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").

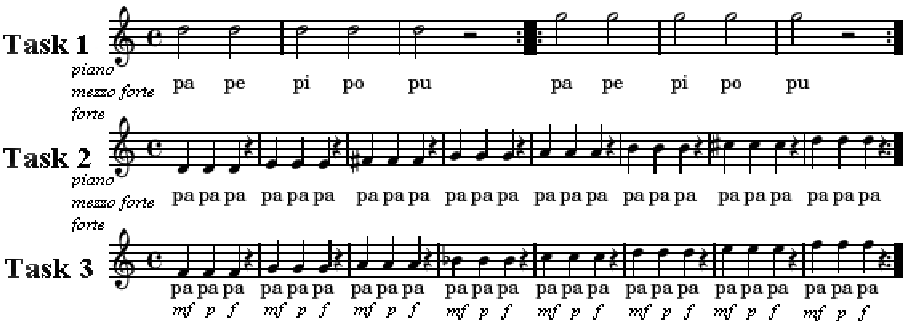
The definition and the motivation of the VRP_{phys} were explained to each subject. Various examples of maximal and minimal phonation with no concern for voice quality were given, and the non-singing aspect of the exercise was stressed. Subjects first started at a mid-range pitch, in an *as soft as possible* 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, ± 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 $\mathrm{VRP}_{\mathrm{perf}}$ differs considerably from a physiological $\mathrm{VRP}_{\mathrm{phys}}$. 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.

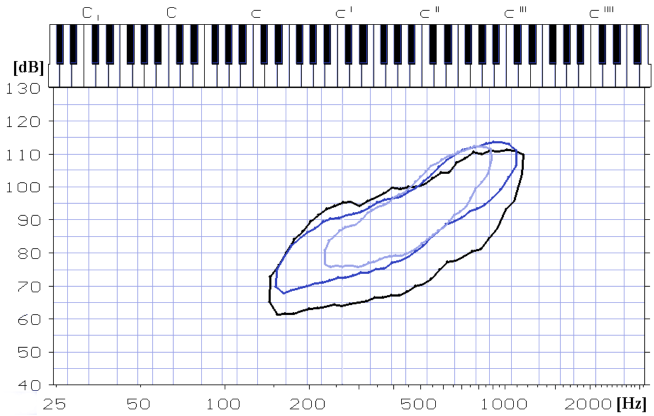


Figure 5.1: A physiological contour (black) average compared to two types of VRP_{perf} (the vocalise and the aria) averaged contours (shades of blue). This figure appears in Paper I as Figure 10.

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 $\mathrm{VRP}_{\mathrm{perf}}$. When task effects were examined, no significant tasking effect could be detected in the $\mathrm{VRP}_{\mathrm{perf}}$ recordings, yet $\mathrm{VRP}_{\mathrm{phys}}$ results, especially in concern to the lower curve, demonstrated potential tasking effects.

Two noteworthy VRP metrics for the singing voice were introduced: $\mathrm{Percent}_{\geq 90\mathrm{dB}}$ (the percentage of the voice area equal and/or above 90 dB) and the $\mathrm{SPL}_{\mathrm{ext}}$ (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.

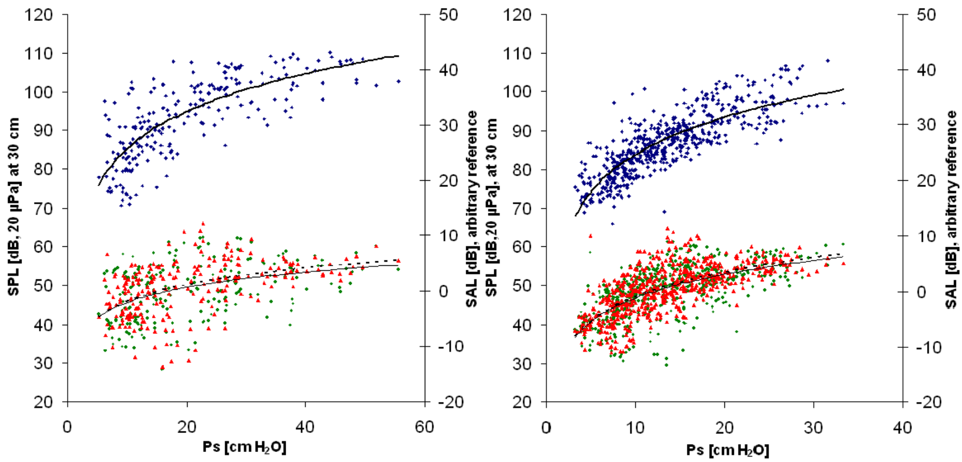
Results

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 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.

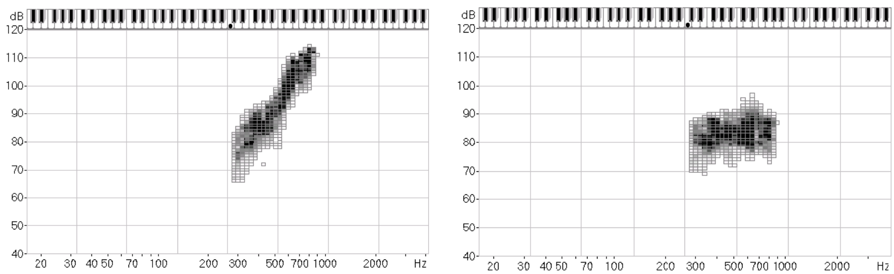
Secondly, in the singing voice, SAL is capable of displaying more immediate information whereas SPL includes a variety of factors (dominated by F_0) that impact voice amplification. While F_0 is the factor that explains most of the SPL variation, the same is not true for SAL. Although changes were small, 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$_{\mathrm{N}}$ (measured at the jogular notch) and SAL$_{\mathrm{S}}$ according to the division of the dataset of Study II into two statistical designs. The regression outcome for a) $YSPL$=$14ln(x)$+ 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$_{\mathrm{N}}$ with green triangles and SAL$_{\mathrm{S}}$ 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

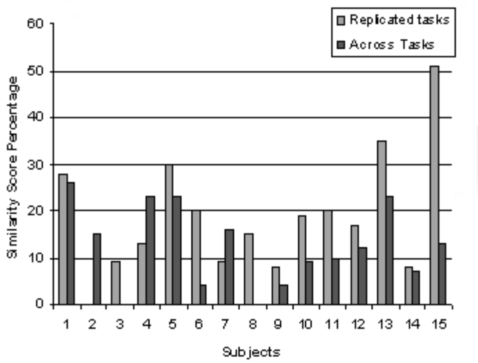


Figure 5.4: Similarity scores obtained for a replicated task and across different tasks. For the majority of subjects, similarity scores lowered when the button pressing behaviour was observed across different tasks.

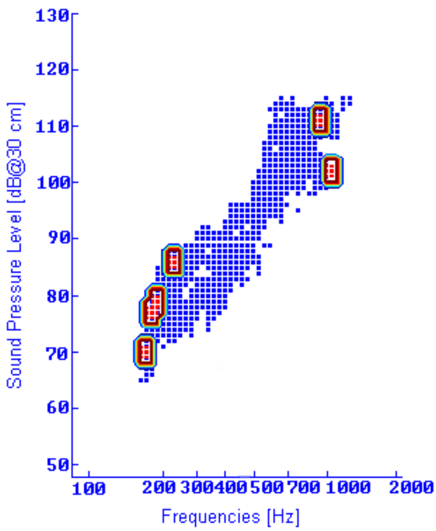


Figure 5.5: A Matlab reconstructed VRP displaying button presses and button regions for a healthy professional soprano. The tendency to press at the periphery of the VRP was a common observation for all 16 singers.

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.

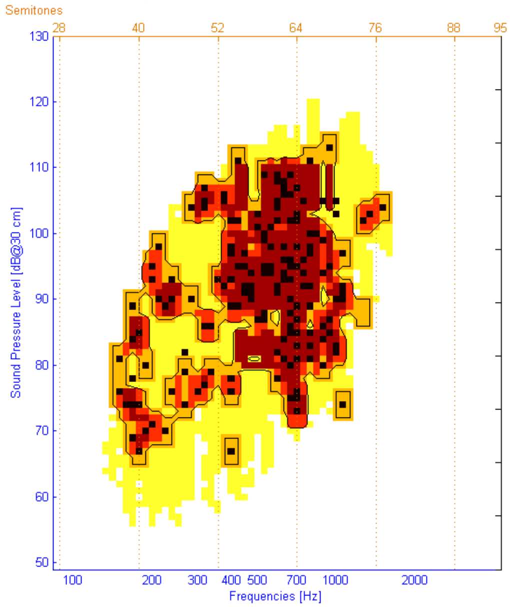
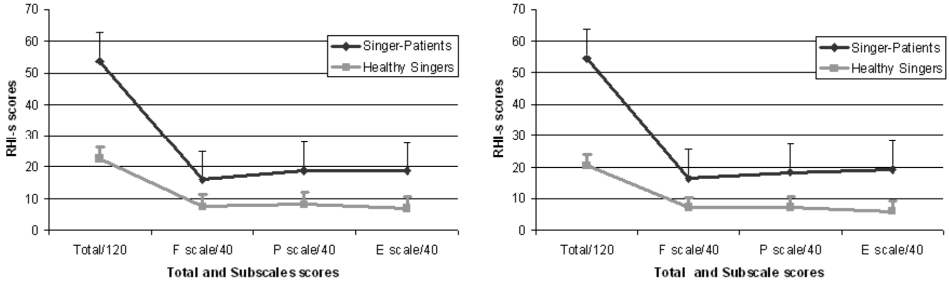


Figure 5.6: Total accumulated button presses for singer patients. The yellow-orange hue identifies areas of single button presses whereas the darker hue gradations underscore the amount of overlap obtained from comparing the initial task to its replication. This figure appears as part of Figure 8 (lower 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, $\mathrm{VRP}_{\mathrm{perf}}$ 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 $\mathrm{VRP}_{\mathrm{perf}}$. The latter is not advocated as a replacement for the $\mathrm{VRP}_{\mathrm{phys}}$ 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

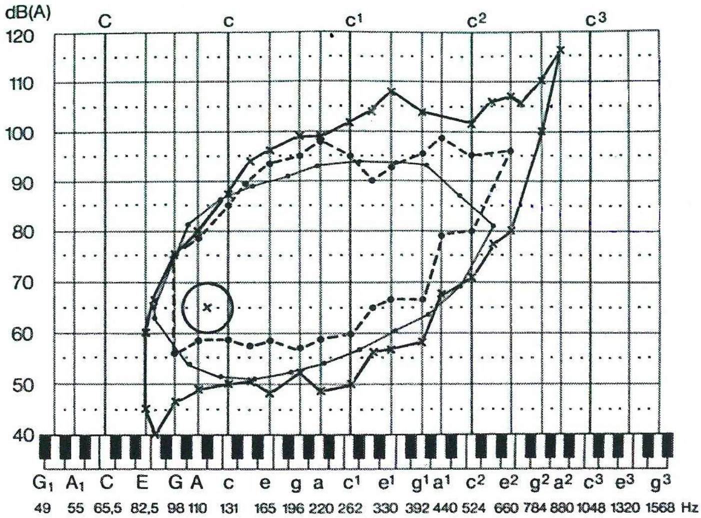


Figure 6.1: A Baritone singer with a singing voice problem: particular dynamic restrictions between mechanisms 1 and 2. This is an example taken from Schultz-Coulon which clearly demonstrates the need to do both a VRP_{phys} (the crossed symbols) and a "musical range of phonation" VRP (the dotted line) 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 represtents a normative reference.

should clearly illustrate the voice transitions between vocal mechanisms ("register") 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 physiological 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 <u>the adaptation of common clinical tools to the concern and the perception of the singers</u>. 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 <u>the creation and documentation of normative references for the singing population</u>. 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 $\mathrm{VRP_{phys}}$ 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 $\mathrm{VRP}_{\mathrm{perf}}$.
  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 $\mathrm{VRP}_{\mathrm{perf}}$ 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, $\mathrm{P_s}$ is more strongly related to SPL than it is to SAL. More research is needed to understand the SAL-$\mathrm{P_s}$ 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 Durchschnittsstimfeld. *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 Berücksichtigung 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. Investigating voix mixte: A scientific challenge towards a renewed vocal pedagogy. In M. Marin & J. Ross (Eds.) K. Maimets-Volt, R. Parncutt, editor, *Proceedings 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 capacity 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 phonetographic 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 computer 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 (loratadine). *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 / stimmfeldmessung). 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 Stimmfeldes. *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 44).

[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 fortissimo 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 untrained 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 systemic 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 pressure 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 articulation. *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).

Appendices

Appendix A

Proposal for SRP, VRP_{phys} and VRP_{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$_{\mathbf{phys}}$

- 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$_{\mathrm{phys}}$ 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$_{\mathbf{perf}}$

- 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.

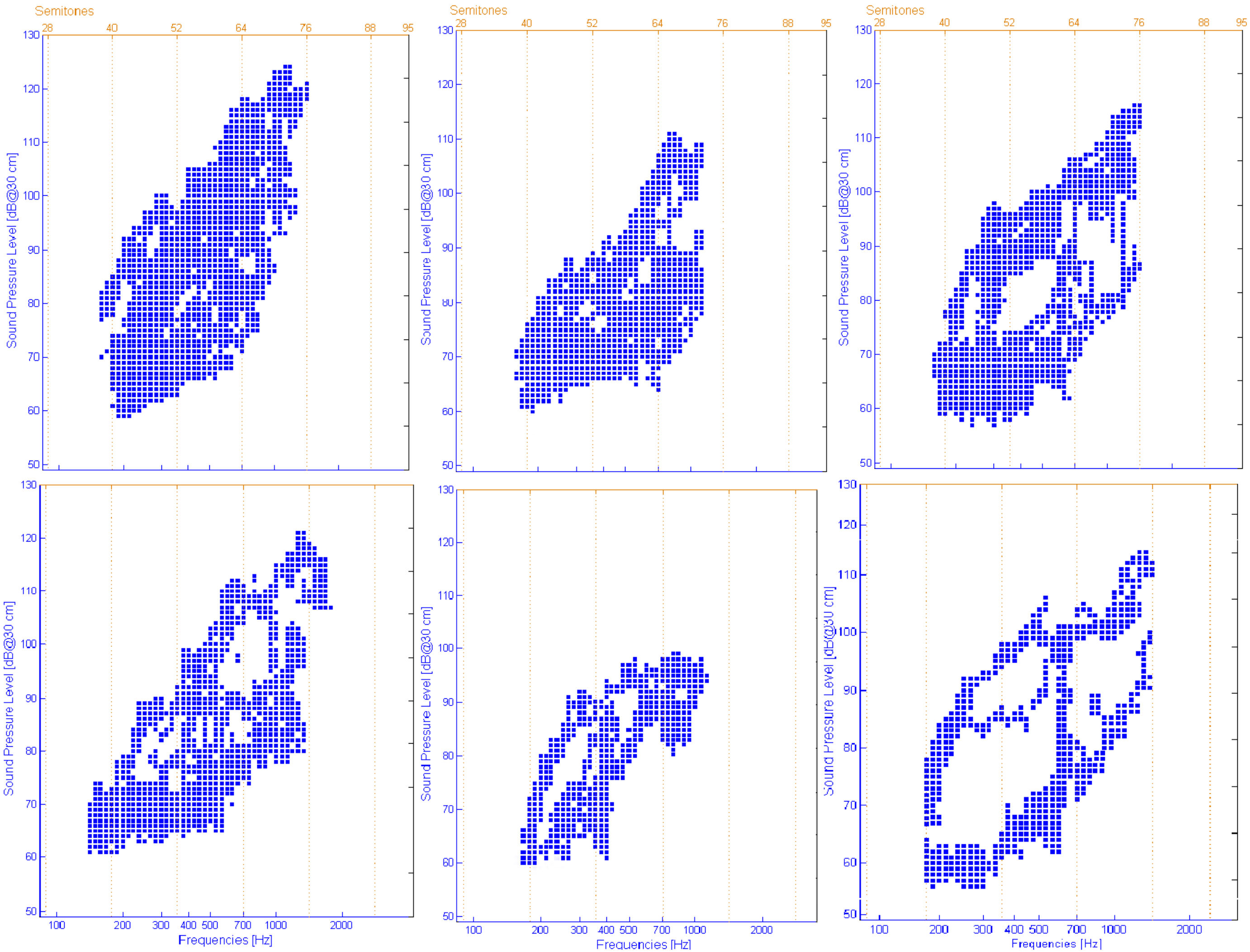


Figure B.1: Physiological VRP (*glissando* task). From the top left, soprano subjects 23 and 25.

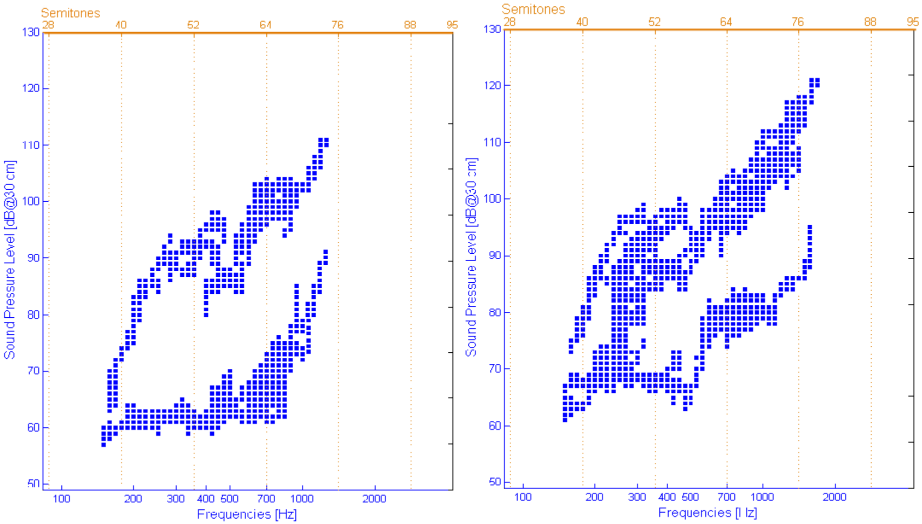


Figure B.2: Performance VRP based on the vocalise task. From the top left, soprano subjects 1, 2 , 3, 6, 8 and 9.

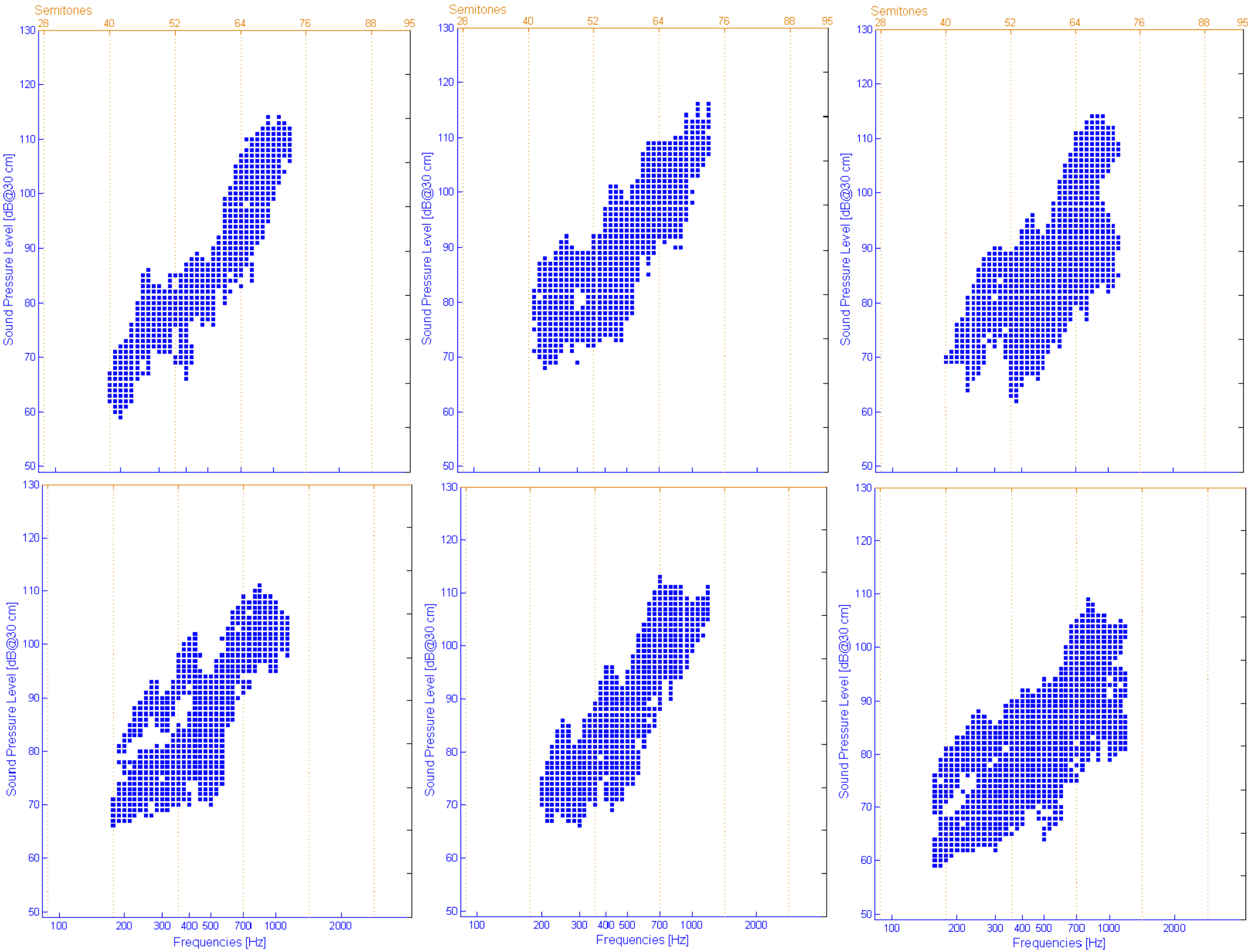


Figure B.3: Performance VRP based on the vocalise task. From the top left, soprano subjects 11,14,15,16,18 and 19.

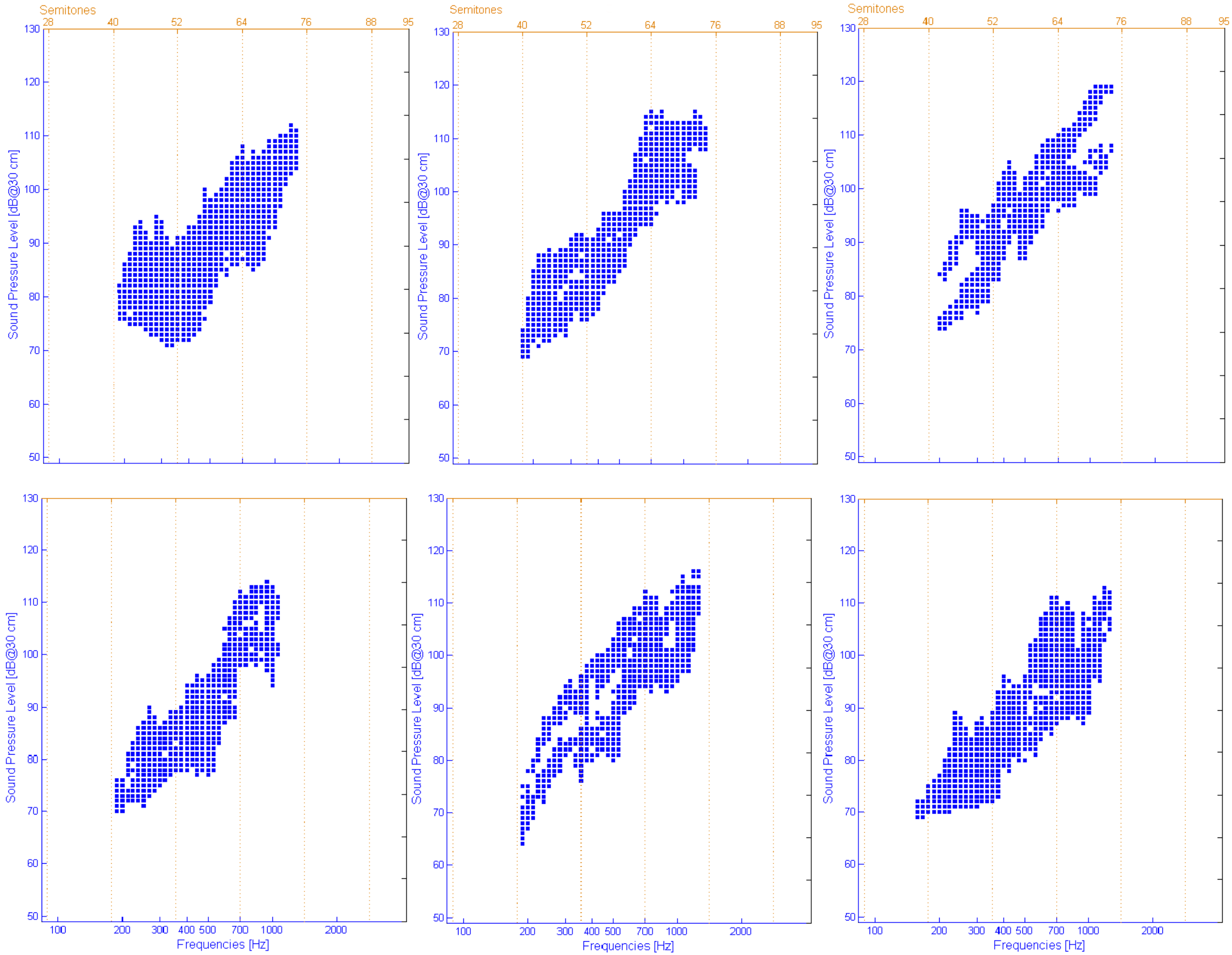


Figure B.4: Performance VRP based on the vocalise task. From the top left, soprano subjects 20,22, 23 and 25.

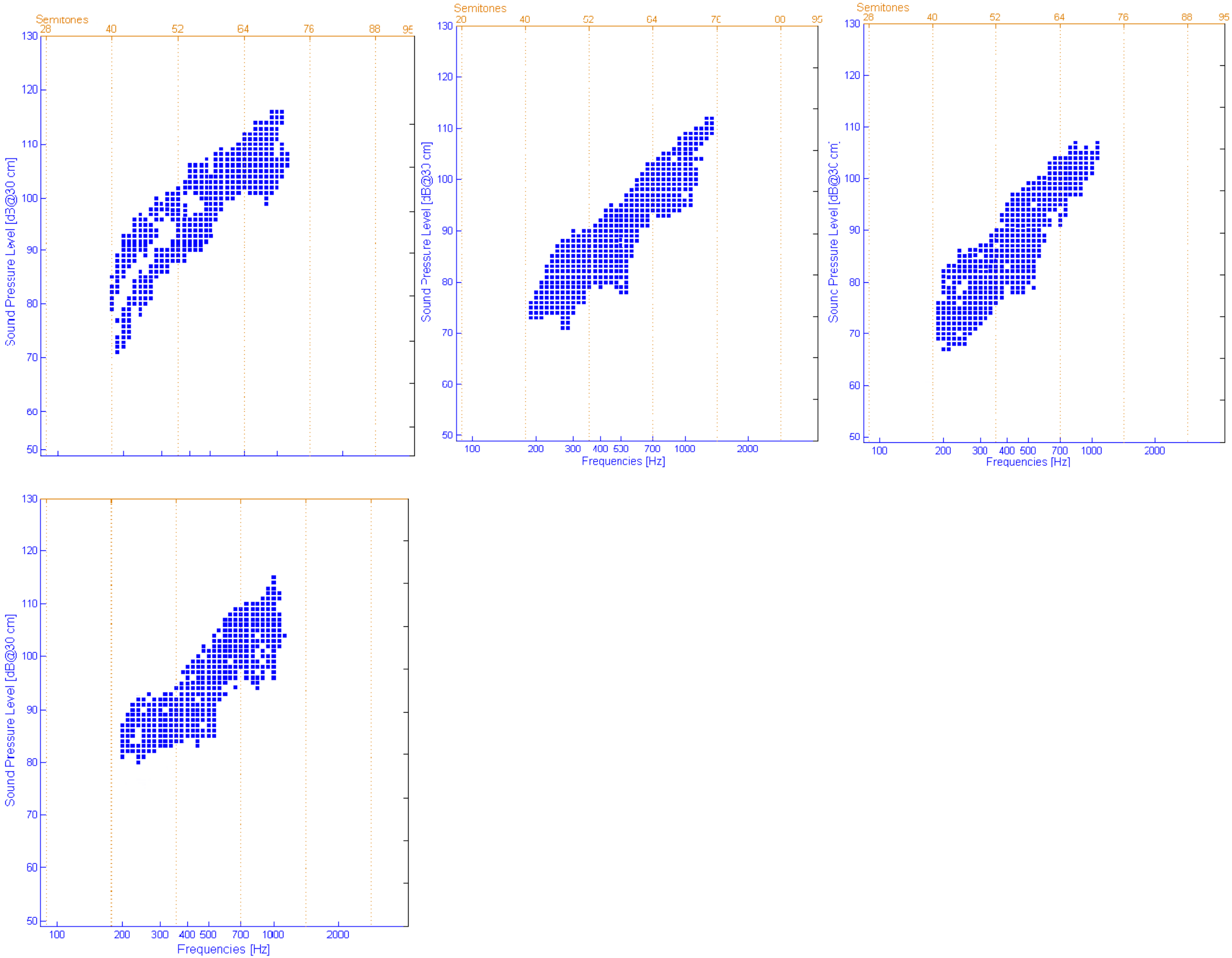


Figure B.5: Physiological VRP (*glissando* task). From the top left, mezzo-soprano subjects 21 and 30.

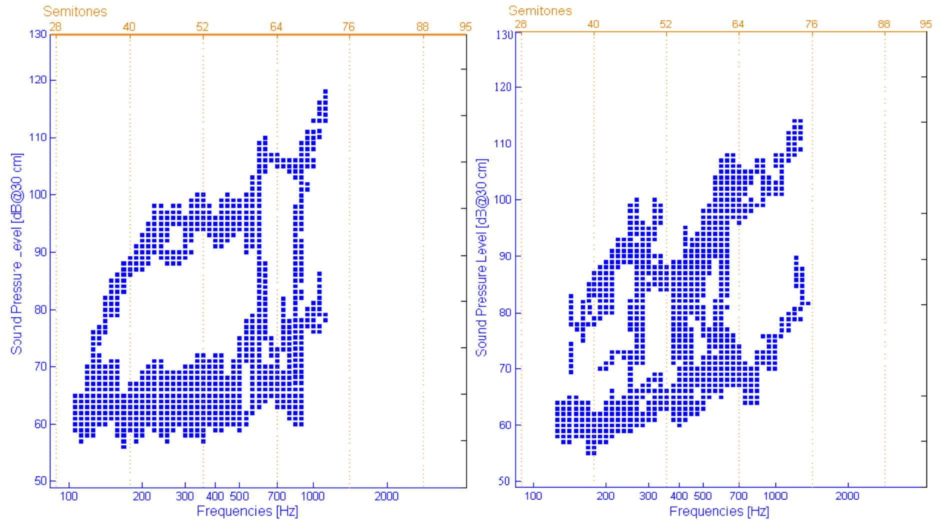


Figure B.6: Performance VRP based on the vocalise task. From the top left, mezzo-soprano subjects 4, 5, 7, 10, 12 and 13.

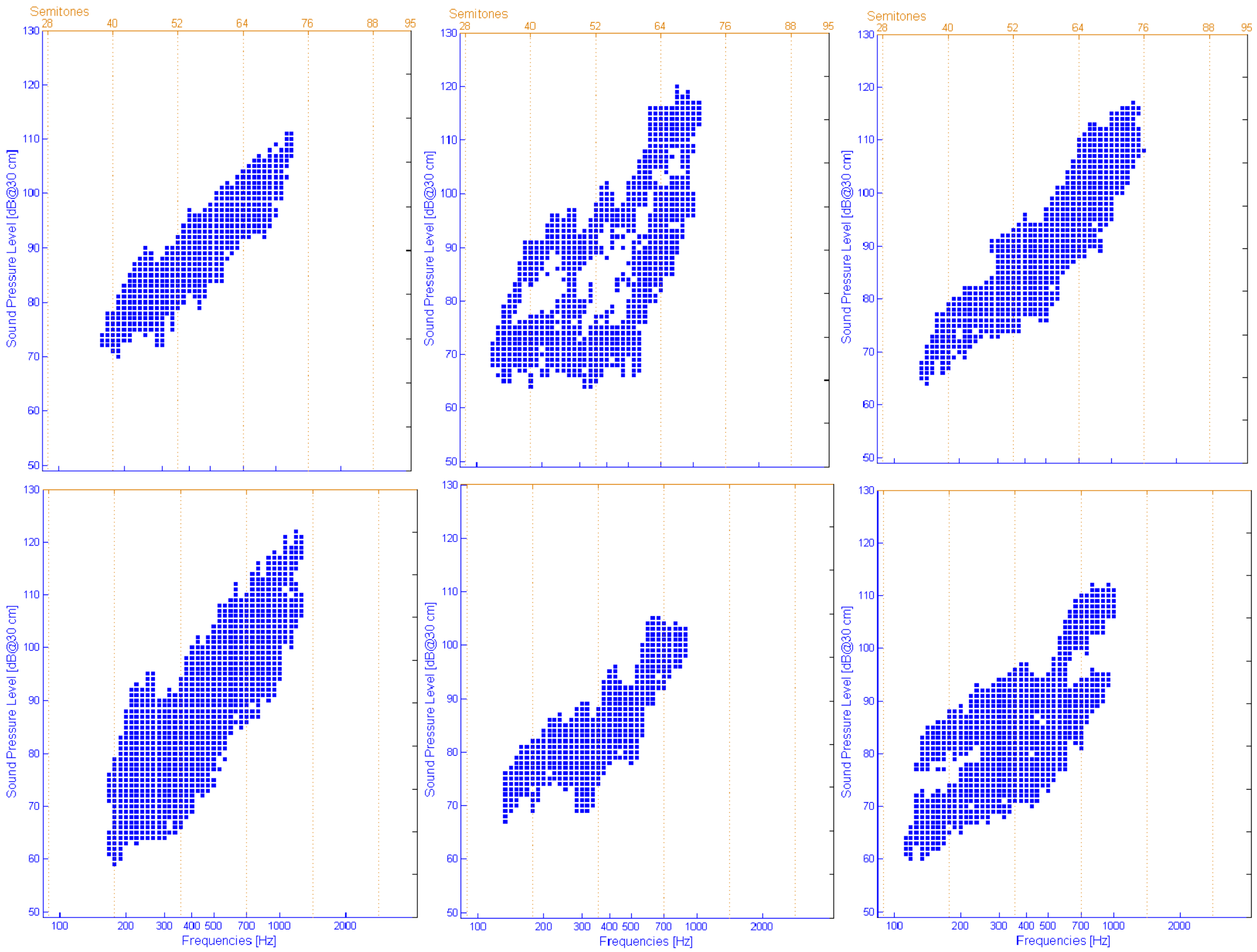


Figure B.7: Performance VRP based on the vocalise task. From the top left, mezzo-soprano subjects 21 and 30.

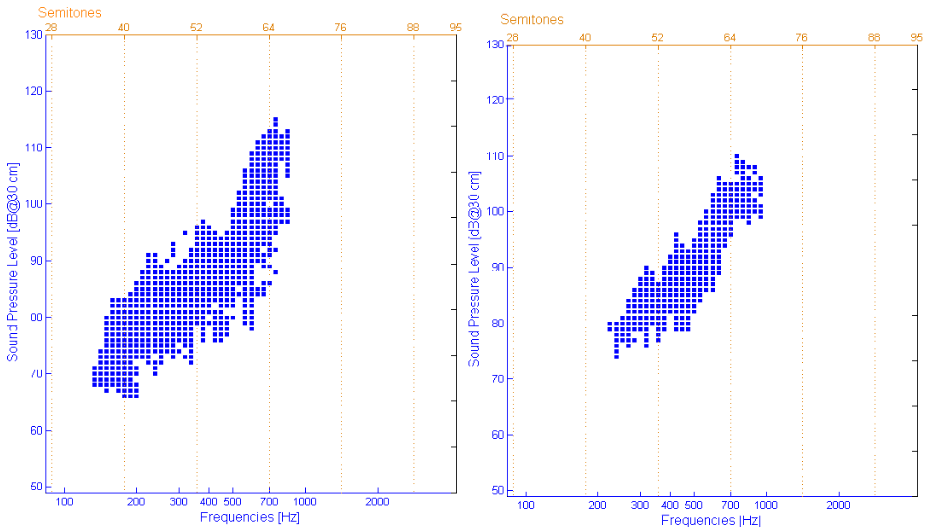


Figure B.8: Physiological VRP (the glissando task). From the top left, contralto subjects 17,24, 26, 27, 28 and 29.

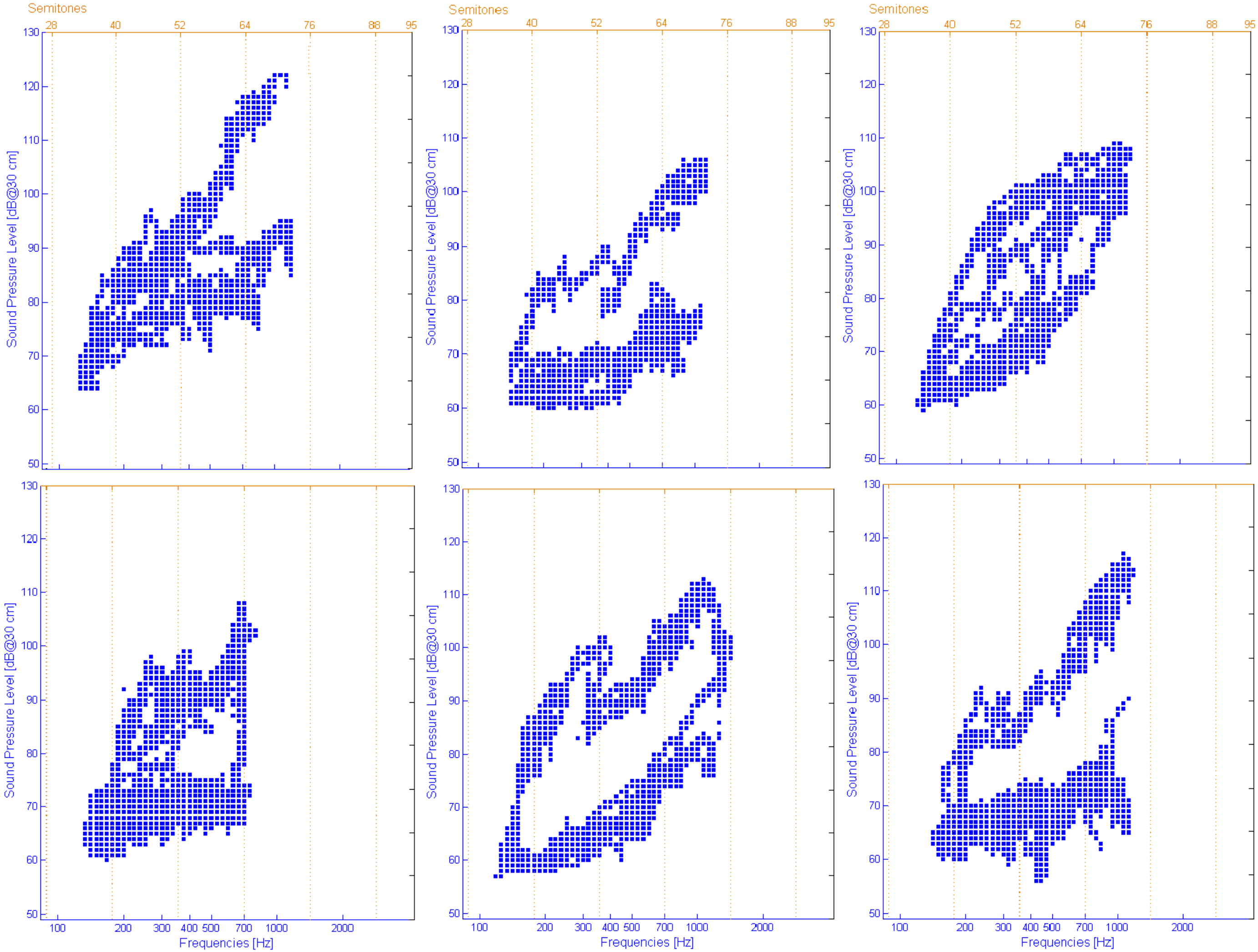


Figure B.9: Performance VRP based on the vocalise task. From the top left, contralto subjects 17,24, 26, 27, 28 and 29.