# Registers in Singing. Empirical and Systematic Studies in the Theory of the Singing Voice

Miller, Donald Gray

*Publication date:*
2000

*Citation for published version (APA):*
Miller, D. G. (2000). Registers in Singing. Empirical and Systematic Studies in the Theory of the Singing Voice. Wageningen: Ponsen & Looijen BV, Wageningen.

# REGISTERS IN SINGING

## Empirical and Systematic Studies in the Theory of the Singing Voice

Donald Gray Miller

# Registers in Singing

## Empirical and Systematic Studies in the Theory of the Singing Voice

# STELLINGEN

behorend bij het proefschrift van D.G. Miller

## Registers in Singing

Empirical and Systematic Studies in the Theory of the Singing Voice

1 The task of science, when applied to singing technique, is to describe in objective terms the (physiologic, aerodynamic, and acoustic) mechanisms employed by singers, not to predict their behavior or set normative standards. (Chapter 1)

2 The sound of 'cover,' whose origin is in the need to bridge the gap between the natural registers, is expected by connoisseurs of operatic singing. (Chapter 7)

3 Generalizing on the basis of the singing behavior of untrained or slightly trained persons cannot reveal much about the strategies of the highly skilled. (Chapter 4)

4 Scientific investigation of the singing voice is limited by the fact that most 'knowledge' includes a subjective element of experience that remains hidden from the outside (non-singer) observer. (Chapter 1)

5 The expert ear remains the judge of the singer's sound. (Chapter 10)

6 In the four centuries since virtuosic singing appeared in Italy, the demands composers make on the singing voice have changed, but the vocal organs have remained virtually the same. (Stark, Bel Canto, Toronto, 1999)

7 Sports medicine, which attends to exceptional function, is a good model for the practical application of voice science to the professional singing voice, but singing receives too little attention.

8 Features of musical instruments that might be seen as acoustic defects become their distinguishing characteristics, and technical 'improvements' that have not preserved those features have not survived. (Fletcher and Rossing, The Physics of Musical Instruments, New York, 1998)

9 There is a contradiction between the ideal of the social democratic state, which strives for equality among all members of the body politic, and the ideal of extending 'brotherhood' to all mankind.

10 Reforming the structure of higher education is less effective than improving the effectiveness of individual programs with quality and quantity of staff.

11 The concept of play, which Huizinga develops for art, law, war, etc., can be extended to research, making 'fun' one of its important motives. (Huizinga, Homo Ludens, Groningen, 1985)

12 'Departmental thinking,' which constructs research paradigms that facilitate typical research within a limited field, is a hindrance in dealing with a multi-disciplinary topic, such as voice.

13 The perceived expertise of the investigator has no place in the model scientific experiment, but plays a large role in peer review.

Rijks*universiteit* Groningen

# Registers in Singing

## Empirical and Systematic Studies in the Theory of the Singing Voice

Proefschrift

ter verkrijging van het doctoraat in de
Medische Wetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. D.F.J. Bosscher,
in het openbaar te verdedigen op
woensdag 14 juni 2000
om 16. 00 uur

door

**Donald Gray Miller**

geboren op 21 februari 1933
te Englewood, New Jersey, USA.

Promotor: Prof. dr. H.K. Schutte

Referent: dr. T. de Graaf

Beoordelingscommisie: Prof. dr. J.J. ten Bosch
Prof. dr. L. Boves
Prof. dr. ir. H. Duifhuis

ISBN 90-367-1237-8

| | |
|---|---|
| Paranimfen: | dr. A.M. Sulter<br>drs. D. Skalnik-Poláčková |

**Dedication**

To Dagmar, whose untiring generosity made possible the sometimes improbable completion of this project.

Printed by: Ponsen & Looijen bv Wageningen

Styling: Klaas A. Maat

Miller, Donald Gray
Registers in Singing: Empirical and Systematic Studies in the Theory of the Singing Voice

*Dutch title*
Registers bij het zingen: empirisch en systematisch onderzoek voor de theorie van de zangstem

Author's e-mail address: d.g.miller@med.rug.nl

Thesis University of Groningen, the Netherlands – With a summary in Dutch

ISBN: 90-367-1237-8

# Foreword

A physiologic and acoustic investigation of the registers of the singing voice arose as a serious project one-and-one-half decades ago during my initial assignment in Groningen, a study supervised by the late Prof. Janwillem van den Berg and undertaken during a leave of absence from Syracuse University. Although the project came up relatively late in life, my fascination with the question of registers can be traced back to 1954. This was my second year of singing lessons and a time when the idea that I might someday become a professional singer and singing teacher would have struck me as remote, had I considered it at all. Even to a beginner, the idea of somehow combining the chest and falsetto registers seemed the deepest mystery in an art that was full of mysteries. And this mystery remained at the center of my own struggle with vocal technique as a singer and voice teacher over the next thirty years.

The fulfillment of the project, at least to the extent that the formal study has been concluded, has also been a drawn-out affair. This is partly due to the fact that circumstances allowed me to work deliberately, but also, in no small measure, because of the elusive and varied nature of the subject. While some voice teachers with strong opinions on registration, as the phenomenon is called, have seen it as central in singing technique, voice scientists have tended to shy away from the details of the subject. Because of this relative neglect, together with the rapid development of the technology for the investigation of voice, Harm Schutte and I have been able to happen on some rather fundamental discoveries in the course of this project. These are now assembled in the present volume, which we hope will provide a more comprehensive view of the phenomenon of registration, one in which the whole is more than the sum of the parts.

While the responsibility for the project comes down to a single author, a large number of people have made important contributions, wittingly or unwittingly, to the final result. Of these, I would particularly like to thank the following:

*Individuals who gave an early impulse, ultimately leading to this work:*

My mother, Evelyn A. Miller, who gave not only life and an inquiring attitude toward nature decades before this project was conceived, but also generous financial help when it became a consuming hobby.

William F. May, who guided my early intellectual probings and to this day has continued to provide invaluable support and advice. One does not simply begin a new discipline after reaching the proverbial age of wisdom without having some of the basics already in place. Of all my teachers in 'previous lives,' I especially want to thank Bill.

Jo Estill and Richard Miller, whose personal examples, viewed close at hand, helped to inspire my own efforts to apply voice science to the singing voice.

Martin Rothenberg, whose lively mind did not grow weary of the questions of an autodidact attempting to compensate for a lack of formal training; who provided the best example for his own maxim: Research should be fun.

*My co-authors of papers on the singing voice:*

Arend Sulter, who introduced me to magnetic resonance imaging, as well as to the advantages of the three-author review of text.

James Stark, who in his annual visits brought to Groningen his knowledge of historic treatises on singing, where we did our best to hold them up to the light of fact, on their way to finding a place in his splendid history of vocal pedagogy, *Bel Canto.*

Jan Švec, who arrived from the Czech Republic with some original ideas about voice breaks, and soon established himself and his voice breaks as a permanent fixture of the laboratory, even when his presence was only intermittent.

James Doing, the first professional singer to make practical application of the research on visual feedback for singing instruction, now continuing in the VoceVista project at the University of Wisconsin-Madison.

Professional singers Djoke Winkler-Prins, Charles van Tassel, Michael Sylvester, Linda Kelm, Renate Faltin, Maja Schermerhorn, James Doing, Janice Chapman, Joella Todd, Kevin Smith, Harald Quaaden, Ank Reinders, Ad van Baasbank, Martin Sprenger, Jos van der Lans, Jodi Gilbert, Jeanette van Dijck, and Eugenie Ditewig, who found the time and courage to submit their finely honed voices to invasive procedures for no more reward than the hope of advancing uncertain science.

Sometime pupils Marlies de Waard, Jopie Kuiper, Jimi James, Lukas Kamps, Titia Keiser, Minke Harryvan van Loon, Gonda Jonker, and especially Wies Bouma and Hanny van Lankeren. They not only cooperated in experiments, but provided excellent voices for their teacher to learn from.

The Voice Foundation, for furnishing a framework for the regular exchange and publishing of specialized information on the singing voice. In this context the work of Johan Sundberg, Ingo Titze, and Ron Baken has been most helpful.

My children, who brought me into their own research and writing in a way that helped me to envision this as an academic enterprise.

Louis Journée and Richard Horne, programmers of successive versions of the software program VoceVista, with which the figures for Chapter 2 of this volume are made.

Professor Daniel Kernell, who helped with reading and revision. His respect for and insistence on high historical standards of research contributed substantially to the challenge and the reward of completing this work.

Leo van Eykern, Jan van Dijk, Meindert Goslinga, Ellen van Drooge, and Ina Heidema-Kol, who provided important technical and secretarial support.

The Dutch Organization for Pure Research, and the College of Visual and Performing Arts at Syracuse University, for financial support in the early stages of the project. Also the division Artificial Organs of the Department of Biomedical Engineering at Groningen University for assistance in bringing the project to a successful conclusion.

My greatest debt is to Professor Harm Schutte, founder and head of the Groningen Voice Research Lab, whose experiment directly measuring the pressures across a singing glottis dramatically began this series of inquiries, in which he has always been the essential partner and co-author; who in addition steadfastly provided, even in seasons of scant support and adversity, what must be the best laboratory-cum-library in the world for the empirical study of the singing voice.

Donald Miller
February, 2000

# Table of Contents

Chapter 1 attempts to locate the position of the theory of the singing voice among the academic disciplines, as well as fulfilling the more traditional function of tracing the history of the narrower problem of registers. Medical physiology is seen as the discipline historically most congenial to a scientific theory of the singing voice. Examining the extensive terminology of registers is rejected in favor of a search for the factual basis of prominent discontinuities that pose a challenge to the perceptual integrity of the voice.

Chapter 2 presents a synthesis of the scheme of registers for male and female singing voices, derived from the combination of generally accepted knowledge of the registers, together with new material explored in the published and submitted articles. The fundamental contrast chest/falsetto is described as the basis of the register concept, a concept that is extended to other discontinuities in the desired integral voice, whether these have their origins in the physiology of vocal-fold vibration or the acoustics of the vocal tract. The four most prominent discontinuities for female and for male singing voices, respectively, are identified and described.

**Chapter 3 – Feedback from Spectrum Analysis Applied to the Singing Voice** **63**
(published in *Journal of Voice*, 1990;4:329-334)

Chapter 3 is the first of the articles from peer-reviewed journals that comprise the main body of this dissertation. It introduces a practical method for determining, with sufficient accuracy, the formant (resonance) frequencies of the vocal tract in singing. An explanation is offered as to why this problem, of relatively little importance in speech science, should play such a crucial role in understanding what highly skilled singers do. (The method, using non-periodic vibration of the vocal folds while carefully maintaining a singing posture of the vocal tract, is applied in four of the following chapters and reconsidered in Chapter 10.)

**Chapter 4 – Belting and Pop, Nonclassical Approaches to the Female Middle Voice: Some Preliminary Considerations** **77**
(published in *Journal of Voice*, 1993;7:142-150)

Chapter 4 compares ‘classical’ singing techniques with those of ‘pop’ and ‘belting,’ a primarily female technique tracing its recent origins to the Broadway musical stage. A concise definition of belting is offered in terms of its characteristic physiology (the chest register extended upward to higher frequencies) and acoustics (the first formant matching the second harmonic of the voice source).

Chapter 5 takes a close look at the singing voice under conditions of maximal resonance efficiency: the soprano high voice with a 'falsetto' voice source and the first formant matching the fundamental frequency. The data gathered include the modulation of pressures measured directly above and below the glottis by miniature pressure transducers, revealing a remarkable acoustic back pressure, where the pressure across the glottis actually reverses during the open phase of the glottis.

Chapter 6 explores the highest useful register in singing, identifying as its determinative characteristic the point in the frequency range where the fundamental frequency passes beyond the first formant frequency. The voice remains nonetheless resonant, because the second formant draws close to the first, forming a broader composite formant.

Chapter 7 takes up one of the most conspicuous problems in vocal pedagogy: male *passaggio*, together with ‘cover’ and the extension into the upper end of the singing range. Various alternative strategies of the singer are considered in dealing with this situation, including the forced ‘register violation.’ The solution to the problem is identified as primarily an acoustic adjustment of the vocal tract, rather than an adjustment of the intrinsic muscles of the larynx.

## Chapter 8 – Soft Phonation in the Male Singing Voice 147

(submitted to *Journal of Voice)*

Chapter 8 considers the individually varying discontinuities encountered in diminishing from forte to pianissimo and *mezza voce* (half voice) in male singers. The signals of two highly accomplished tenors are examined, each meeting the challenge in his own way, according to what the structure of his vocal folds will allow. Objective criteria are described for a proposed distinct register, *mezza voce*, including an opening of the typically constricted laryngeal collar, and incomplete closure of the glottis.

## Chapter 9 – Measurement of Characteristic Leap Interval between Chest and Falsetto Registers 165

(submitted to *Journal of Voice*)

Where the previous chapters all approached register discontinuities with the typical singer's task of smoothing and disguising them, Chapter 9 looks at the chest-falsetto transition in its maximally abrupt form. The size of the interval leap between the two registers is measured and shows a tendency to be larger in male than in female voices.

## Chapter 10 – Comparison of Vocal Tract Formants in Singing and Non-periodic Phonation 187

(published in *Journal of Voice*, 1997;1:1-11)

Chapter 10 revisits the material of Chapter 3, exploring sung sounds and their non-periodic ‘imitations,’ frequently used in these studies to ascertain formant frequencies. Magnetic resonance imaging is used to determine the extent to which constant posture of the vocal tract is responsible for the accuracy of the determination of formant frequencies through the use of non-periodic phonation. It is concluded that the accurate reproduction of the first two formant frequencies in non-periodic phonation has an aural, rather than a postural, basis.

# Chapter 1

# Introduction

## The project

The idea of registers in the human voice arises from a wide range of common perceptions, from the involuntary voice breaks of the adolescent boy to the art of the virtuoso yodeler. The abrupt change in the character of the voice is striking, suggesting a change in the character of the speaker or singer, a change exploited theatrically by the ventriloquist. Apparent as the voice break is, it is so common that we accept it without further thought, providing that it does not disrupt some solemn moment. Such changes in register of the voice have undoubtedly been around since long before human history; indeed, listening to the vocal communications of dogs persuades us that similar registers are found in other species of mammals as well.

Pitch, duration and loudness are dimensions of vocal sound that are fully developed in music. Register is not included among these dimensions, but belongs rather to sound quality, which is physically represented in spectral makeup, the balance of the various overtones of a given sound. Pitch and duration can be precisely indicated in musical notation, as can, with a lesser degree of precision, loudness. Quality is quite another matter. Here the composer is limited to specifying the instrument or voice, the vowel, and sometimes a hint regarding the emotion to be conveyed.

Singers were fully aware of registers long before the physical basis of voice in the larynx was understood, and they named the primary registers according to their impression of where the separate 'voices' originated: chest[1] and head. The 'head' voice was also called falsetto, because it contrasted with the 'natural' voice in the male singer. The female singing voice played a secondary role at the time this nomenclature arose (see Stark, p. 59 [Stark, 1999]).

In artistic singing, aesthetic considerations have encouraged the perception of the unified, rather than the segmented, voice. As a consequence of this, the training of skilled solo singers, whose music generally has a more extensive range than the sixth-to-octave compass of the ordinary hymn or folk song, has emphasized the smoothing of the transitions between registers, or the 'even scale.' This minute attention to the evenness of the scale brought attention to

---

[1] The chest voice takes its name from the vibration that can be felt on the chest wall from the reverberation of sound in the subglottal airways, which are fully resonant through ca. *D4* (292 Hz), the point in the ascending scale where the natural voice tends to jump to the falsetto register. It is ordinarily considered the dominant register in the frequency range up to that point.

ter was thus historically extended beyond its original use for chest and falsetto, most prominently in distinguishing between the female middle and upper voice.

With the invention of the laryngoscope in the mid-nineteenth century came empirical knowledge that the distinction between chest and falsetto was located in the pattern of vibration of the vocal folds. The chest and head 'resonances' that singers had associated with the two primary registers thus lost much of their explanatory power among those who sought a scientific explanation for the question of registers. The emphasis shifted to a 'physiological' interpretation of the registers, and 'registration' became a matter of balances among the intrinsic muscles of the larynx, a viewpoint that prevails even today among many thoughtful teachers of singing.

It was not until the second half of the twentieth century that the complex role of the vocal tract in voice production became fully appreciated. The availability of spectrum analysis then made it possible to follow how the resonances of the vocal tract were affecting the individual harmonics of the voice source. The acoustic theory of speech revealed how the strongest resonances of the vocal tract (the first two formants) responded to movements of the tongue, lips, jaw, and larynx to form the several vowels. The subsequent recognition that skilled singers, especially at higher pitches, routinely tune one or the other of these formants to a few available harmonics was slow in coming; nonetheless, it was this insight that made it possible to extend the concept of registers to cover those cases where the discontinuity in sound (or 'placement,' in the language of singers) results from a major shift in the use of resonance. (An example of this can be found in Chapter 7, the transition from chest to full head in the male voice.)

The following pages attempt to give a systematic overview of the classical singing voice with respect to registration, as well as to present individual studies that focus on one or another aspect of registers. The phenomenon itself begins with perception: what the listener hears and the singer feels. What we seek to describe are measurable physical properties at the basis of the perceived registers. These will be found in some combination of the two basic components of the singing voice: the voice source, including the breath pressure and the vibrational patterns of the vocal folds; and the resonances of the vocal tract, particularly the way these are adjusted to select one or another of the harmonics of the voice source to create standing waves in the vocal tract. How this takes place will be explained in detail below.

tract. How this takes place will be explained in detail below.
I have selected eight classes of sung sounds – four male and four female – that singers often speak of as 'registers.' These are presented systematically in Chapter 2 and include both those distinguished primarily by voice source properties (female chest, male falsetto and *mezza voce*); and resonance characteristics (female upper and flageolet, male full head). The remaining two (female middle and male chest) may be considered the 'default' singing registers, from which the others are distinguished. Chapters 3-10 are independent articles describing original research on some aspect of these registers or, in the case of Chapters 3 and 10, on a new and practical method for exploring the role of the vocal tract in singing. Neither the selection of the registers nor the names used are new. The original contribution rests in the description of their physical properties, both physiological and, especially, acoustic.

## Problems regarding scientific study of the singing voice

The ease with which most children repeat a simple tune tends to make us overlook the complex interrelationships among the cognitive and physiological elements involved in such behavior. Even a superficial description of the mechanisms involved in hearing the melody, reproducing the pitches and rhythm, and monitoring and adjusting this reproduction, gives a suggestion of the vast problems involved in constructing a creature with such a capability. If we consider how this design might have arisen in the course of evolution, the complexity is further compounded. A concise speculation on this subject is offered by Klaus Scherer (1991):

> It might well be, then, that the externalization of affect or emotion via vocalization is at the very basis of music and speech. As ethologists have shown, expression and impression are closely linked. In the process of conventionalization and ritualization, expressive signals may be shaped by the constraints of transmission characteristics, limitations of sensory organs, or other factors. The resulting flexibility of the communication code may have fostered the evolution of more abstract, symbolic language and music systems. This development is likely to have occurred in close conjunction with the evolution of the brain. Just as newer neocortical structures with highly cognitive modes of functioning have been superimposed on older 'emotional' structures such as the limbic system, the evolution of human speech

> (and of musical scales and conventions for singing) has made use of the more primitive, analogue vocal affect signaling system as a carrier signal. In making use of vocalization, which continued to serve as a medium for emotion and music, the functions became by necessity strongly intermeshed. Thus, in speech, changes in fundamental frequency (F0) contours, formant structure, or characteristics of the glottal source spectrum, can, depending on the language and the context, serve to communicate phonological contrasts, syntactic choices, pragmatic meaning or emotional expression. Similarly, in music, melody, harmonic structure, or timing may reflect sophisticated constructions of the composer, depending on specific traditions of music, and may simultaneously communicate strong emotional moods. This fusion of two signal systems, which are quite different in function and in structure, into a single underlying production mechanism, vocalization, has proven to be singularly efficient – for the purpose of effective communication and from the point of view of evolutionary survival. It has also proven to be singularly messy and complicated for scientific analysis.'

The present work is an attempt to bring scientific analysis into a limited portion of the 'singularly messy' signal systems described by Scherer. It examines the 'carrier signal,' vocalization, as it is used in the transmission of signals organized as music. The third element in this fusion of signals, speech, is also present in singing, but our assumption shall be that it takes a subordinate position in vocal music, even in the case where the text which is sung may have originally provided the impulse for the musical composition.

### *The claims of the various disciplines*

The question as to which academic discipline has jurisdiction over this investigation is not easily answered. Early scientific inquiries into the mechanism of voice production came under medical physiology. The first half of the 19th century brought a proliferation of instruction books on singing (see Volontieri, 1995). Physiological investigations of the singing voice in Paris, especially those by F. Bennati (1834), an Italian physician with a singer's training, provided background and impetus for the invention of the laryngoscope in 1856 by Manuel Garcia, the most prominent voice teacher and theorist on the singing voice in his day. The combination, in the same person, of expertise in both vocal physiology and theory of the singing voice set a standard that has been difficult to match. Members of the medical profession continue to claim insight regarding the basis and management of the professional singing voice (Seidner and Wendler, 1997), but even when they themselves sing, they seldom have the experiential knowledge of the fully developed singer. In more recent times the rapid advance of technology, working hand in hand with medical research and practice, has greatly increased the possibilities of observing and recording the structures and behavior of the vocal organs, but this has not yet resulted in a corresponding proliferation of measurements of professional (opera) singers.

'Voice science,' a somewhat different current of voice research, has flourished over the last half century, and offers another powerful discipline with a claim to jurisdiction over this material. A key historical development in this discipline was the work done by the Bell Laboratories after World War II in the acoustics of voice. Both the laryngeal voice source and the vocal tract 'filter', which converts the source into audibly distinct phonemes, could be modeled mathematically by the engineers at the Bell Labs, who were interested primarily in the transmission of speech signals. Gunnar Fant's *Acoustic Theory of Speech Production* (1960) followed this work, and acoustic phonetics, which attends to the frequencies of the resonances (the so-called formants) of the vocal tract, offered a rich supplement to the older articulatory phonetics. All of this, of course, took not just mathematically based insight, but also increasingly sophisticated electronic equipment – particularly that for sound spectrography – which was not available when von Helmholtz was discovering the formants 'by ear' in the 19th century (1954).

This line of research has found an important practical application in the production and recognition of speech by computer. It has also made significant contributions to our understanding of possible mechanisms of the singing voice through computer modeling of the action of the vocal folds, (Titze, 1973; Titze, 1974) as well as through electronic synthesis of the singing voice, (Sundberg, 1977a). Particularly illuminating for singing has been the close attention voice science has paid to the acoustics of the vocal tract – an element which had often been neglected by the medical physiologists. However, this line of research has tended to generalize and extrapolate from a reductive model of vocal function that has served well for speech, while paying only limited attention to the types of singing voice and to the concrete strategies singers use to coax highly refined sounds from the vocal organs.

### *Taxonomy*

A theory of the singing voice which can deal adequately with the various

A theory of the singing voice which can deal adequately with the various types of voice production found in Western operatic singing needs more than a single reductive model of the vocal mechanism with variable quantitative parameters. 'Voice science' has furnished what might be regarded as an engineer's model of the singing voice. The biological world of singers and singing needs in addition a kind of natural history, describing and ordering the varied natural phenomena found in the world of singers. A good taxonomy will not only furnish clear labels for the various categories distinguished, but will also show the relationships between them, so that the particular case can be related to the more general principle. The practitioner – say, a voice teacher, whose expertise consists, to a large extent, in the ability to recognize the various 'species' of singers and singing – is served by such a taxonomic theory in gaining a better recognition of the distinguishing characteristics of specific cases, both individual and typological, as well as a deepened understanding of the underlying common mechanisms.

While voice practitioners may agree on the need for attending to the variety of voice types, their own efforts at formulating theories are often reductive as well. The reason for this can be found in the apparent ease with which one identifies with another voice. One 'species' of singer hears another and very quickly arrives at a feeling of what the other is doing at the level of voice production. Such feelings may be largely illusory, but the lack of objective information on what is happening in the hidden organs of another allows the illusions to persist in the form of strong opinions concerning the essentials of singing technique.

A further line of inquiry borrowed from the biological sciences comes from ethology. Human vocalism can be seen as an extended form of the acoustic communication practiced by other species long before the advent of homo sapiens (Tembrock, 1996). As indicated in the quotation from Scherer, vocalization – the 'carrier signal' for singing – is both the result of and a factor in evolution. The signal system we know as music makes its own (aesthetic) demands upon the vocal apparatus, and the individual singer responds with whatever mechanisms he can find to fulfill those demands. This interaction between aesthetic ideals and limited, ad hoc resources results in a variety of specific *strategies* which singers employ, for example, in singing difficult high notes. Part of the taxonomic task of a theory of the singing voice will be to identify these strategies and to relate them to the physical organs that are the results of the (interactive) process of evolution.

### Terminology

A serious stumbling block for any scientific inquiry into the singing voice is the arcane language employed by voice practitioners (the 'adept') for the taxonomy mentioned above. Subtle distinctions, presumably perceptible to the expert ear, are given exotic names, often baffling to the outsider. The use of arcane jargon in itself does not disqualify a field from scientific inquiry (consider, for example, particle physics), but the high degree of skill required for true adeptness, together with the imprecision of language in describing auditory phenomena, has made clear, straightforward definitions of important terms an elusive goal. It is, nonetheless, this language in which the practical, pedagogical theory of the singing voice has been formulated, and its concepts will have to be taken seriously in an empirical examination of the presumed distinctions.

In order to deal with this language, it is expedient to borrow a technique from linguistic anthropology, a discipline in which it is standard practice to have both a native-speaker 'informant' and an expert who can explain, to those outside the language-culture being investigated, the structures of the language, particularly in its relationship to items in the culture of its speakers. Thus, for example, one Australian aboriginal tribe, whose culture attends more minutely to nature than does ours, is reported as having ten different words for 'hole', where we can make do with just a few terms[1].

Although the functions of informant and expert are distinct, these can be combined in a single person possessing both types of knowledge. That will also be in the case in the present study, where the author claims adeptness as a singing practitioner[2].

Returning to the question of the discipline which can claim jurisdiction over such a study as this one, we note that it is characteristic of departmental

---

[1] '*Mutara* is a special hole in a spear, *pulpa* is a rabbit burrow, *makarnpa* is a burrow of a monitor lizard, *katarta* is the hole left by a monitor lizarrd after it has broken the surface after hibernation, and so on. This example also shows that even though a language may not have a one-word equivalent for a word in another language, it is possible to provide an adequate translation by a descriptive phrase (which in the case of katarta may take as many as fifteen English words).' (from (Salzmann, 1993). The special problem of singing, of course, is that even the 'natives' cannot give a clear description of what they mean with 'placement' of the voice.

[2] While not pretending to have a noteworthy singing career, the author has managed to hold his own in the professional singing world, having sung over two dozen leading roles of the standard opera repertory and forty parts in orchestral concerts.

thinking to throw up barriers, allowing only particular (discipline-approved) concepts and methods into the argument. If the messy business of voice is ever to get due consideration, however, interdisciplinary treatment will have to be admitted, even if medical physiology is selected as the 'host discipline' for this interdisciplinary discourse. With its traditional willingness to invade the human body by means of sophisticated diagnostic equipment, medical physiology has historically been the discipline most congenial to the discussion of the mechanisms of singing with regard to the singer's interest. We attempt to follow the examples of such interdisciplinary authors as Garcia and Bennati, each with a foot planted firmly in two disciplines.

### *The Language of Singing*

Admitting the language of singing to a scientific discussion brings up two distinct problems. One of these concerns the degree to which the concepts used represent, in the minds of those using them, the entities they are intended to designate. For example, the words 'red' and 'elephant' refer to entities that are part of common experience, making their use relatively unproblematic (although the fable about the three blind men and the elephant reminds us of potential problems in making this assumption). The use of 'mauve' and 'griffin,' on the other hand, gives us pause, since not everyone has a clear representation of (subtle) mauve in his mind's eye, while the representation of griffin, an imaginary creature, depends on fantasy or memory of versions encountered in art works. Analogously: to anyone with a rudimentary knowledge of music, the pitch *B4* (494 Hz) is a sharable concept. If it is further specified, however, as the penultimate note of the aria 'Nessun dorma', as sung by Luciano Pavarotti, the representations it invokes vary widely, depending upon exposure to recordings and/or performances, talent for hearing, singing experience, etc. Even among opera singers, it would represent something different to a tenor who may (or may not) be able to sing it, than it would to a soprano, who might conceive it as similar to a note an octave higher, or a bass, whose representation might resemble that of a different high note from his own range.

Those are some problems arising in discussing an entity that can be objectively identified. We come up against a second, and more obvious, problem if the entity in question – say, 'forward placement' – has no clear definition, or perhaps no objectively established existence – a sort of griffin inhabiting the singers' world. (It is worth noting, however, that the *representations* of such an entity in the minds of those discussing it are not necessarily more diverse than those of a 'real' sound.)

Further factors hindering the development of a clear language for the theory of the singing voice are these: the physiological mechanisms of voice are largely hidden from visual inspection, preventing realistic visual representation of these mechanisms on the part of the singer; motor control over the 'moving parts' is usually accomplished indirectly, 'by ear' (see Chapter 10); musical notation is precise only for those entities that are clearly quantitative – especially pitch and duration – while language and notation are quite limited in their ability to describe and specify sound *quality*.

In the face of these difficulties, it is perhaps remarkable that singing is taught at all. The practical teaching situation, however, need not depend heavily upon language. It is often sufficient for the pupil to experiment, typically by imitation, and for the teacher merely to recognize a move toward what he considers to be the right direction. In the context of demonstration, whatever the successful maneuver is called – perhaps 'placing the tone in the mask' or 'singing on the breath' – can function as an adequate label for the desired sound. Idiosyncratic use of language, where the nuances of the dialect are shared only by the maestro and his (advanced) pupils, easily becomes the rule in the practice of teaching singing.

In spite of these problems with the language of singing, however, codification of the practices and terminology of vocal pedagogy – in the interests of preserving the 'secrets of singing' for transmission to the following generations – continues to be seen as a desirable goal. The combination of uncertainty as to the meanings of the terms and the high degree of expertise required for familiarity with the world which the terms describe has as a consequence that the terms will often be used without adequate understanding by those who have only partially acquired the expertise. In short, the terminology is necessarily misused much of the time. Both this misuse and the related idiosyncratic use can be seen as normal and inevitable. It is not surprising to find teachers or singers with expert hearing regarding vocal function giving explanations of their practice which are scientifically fallacious or, at the very least, quite incomprehensible without accompanying demonstration. Even where they do not disagree regarding the language, they often 'talk past' one another. (For examples of this, see Jerome Hines [1982]).

### *Singing Expertise*

One of the first motives of a theory of the singing voice is to record and communicate what expert singers and singing teachers 'know.' The portion of that knowledge which receives attention in this monograph is intentionally limited. Large parts of what the practitioner must learn are excluded, including musicianship, musicality, repertory, historical styles, and even expression (insofar as this last can be considered learned). What remains is often referred to as 'vocal technique,' although even that term suggests virtuosic elements of performance which are outside the scope of this inquiry. 'Voice production' comes closer, although it fails to suggest the element – perhaps inborn – of beautiful *sound* associated with an outstanding singing voice. 'Chops,' a slang term borrowed from the jazz world – referring originally to a player's mechanical (mouth) connection to his instrument – best describes that portion of technique-plus-sound which interests us here.

The 'knowledge' residing in 'good chops,' whether learned or resulting from innate disposition, is not found in the form of propositions, but rather in habits embedded in neuromuscular structures. The 'ears' to hear what the 'chops' can produce are an essential part of the whole. Performers need 'ears' to be able to imagine (and monitor) the sounds they produce. Less obvious is that teachers need (at least imaginary) 'chops' to be able to hear what the pupil is doing or, alternatively, might do (Martienssen-Lohmann, 1981). Individuals whose neuromuscular structures are disposed to respond to vocal sound are considered talented. It is not surprising, however, that such individuals usually put more effort into communicating by means of 'chops' and 'ears' than into trying to explain what they are doing in this extraordinarily complex process. Nor should it be surprising that their 'chops' are better than their explanations.

The ability to 'identify' with another person's voice, in the sense of feeling what the other appears to feel, is at the heart of the process of vocal pedagogy, as well as of the emotional effect exerted by the singing voice. At the same time, this ability can present obstacles to discovering the physical mechanisms at the basis of the sound, obscuring, for example, even major differences in production between male and female singing voices. A scientific theory of the singing voice, even if it takes singers' language seriously, will have to search behind the singers' explanations to get at the true expertise embedded in 'chops' and 'ears.'

## Historical aspects of the 'register problem'

The 'problem of registers' is simultaneously one of the most important and most elusive in the theory of the singing voice. The fact that the vast majority of human voices are capable of two plainly distinct modes of production, commonly called 'chest' and 'falsetto,' both of which can be used for singing a tune, has undoubtedly been apparent since prehistoric times, since analogous registers are shared by other mammals (Tembrock, 1996). The codifiers of theory concerning singing, a theory which blossomed in print at the beginning of the 19th century (Volontieri, 1995), traced the notion of the unification of the registers back to Tosi (1723) (Tosi, 1987), who had written a theoretical work which was seen, together with a comparable work of Mancini (1777), as containing the basis of the much-admired Italian school of singing. This virtuosic manner of solistic singing, which had been conspicuous since the late 16th century, typically demanded an extended range, giving rise to the problem of preserving the identity of the voice in different parts of its compass. Tosi (and Mancini, following him) distinguished between 'chest,' or 'natural' voice, and 'head,' or 'falsetto.' Without indicating how the unification of the two was to be accomplished, he left no doubt regarding the effective result: the falsetto must be 'united to the chest voice in such a way that the two cannot be distinguished, since if the union is not perfect, the voice will be of more than one register and as a consequence lose its beauty.' (Volontieri, 1995, p.108).

Before proceeding with an historical consideration of how this requirement of unifying the registers has been met, let us pause for a moment to consider why such a requirement should arise in the first place. The importance of maintaining a constant identity of voice has biologic origins, antedating more strictly aesthetic considerations. In his study of acoustic communication among mammals, Tembrock has distinguished between communication at close hand and that in the far field, where the communication is between individuals not visible to one another (Tembrock, 1996). A large part of the typical 'message' in the latter type is simply the identity and whereabouts of the one calling. Singing can be seen as essentially a form of far-field communication, where the listener wants to hear uncompromised identity. More strictly aesthetic is the motive of diversity in unity, served when vocal integrity is preserved through a wide range of pitch, loudness, and especially, emotional expression. Here the various registers can help widen the expressive range, provided that no disturbing discontinuity is allowed to intrude. Finally, the contrast of 'heavy' and 'light,' as epitomized in chest and falsetto, easily takes on the function of an all-encompassing dualistic principle in the abstraction 'registration,' bearing some resemblance to a masculine/feminine or yin/yang polarization organizing experience. But in all the contrast, it is imperative that the single identity of the voice be preserved.

Most of the wave of books on singing which appeared in the first half of the 19th century had as a model a remarkable volume published in 1803, *Méthode de Chant du Conservatoire du Musique*[1], prepared by a committee of distinguished musicians for use in the Paris Conservatory. The *Méthode* acknowledged 'registers' and Tosi's ideal[2] of unifying the voice, defining the acknowledged 'registers' and Tosi's ideal[2] of unifying the voice, defining the former simply as 'a certain series of sounds of the voice whose character differs from that of another series of sounds which constitute a different register' (Volontieri, 1995, p 90). The authors of the *Méthode* introduced a novelty in recognizing a third register, the 'middle', between 'chest' and 'head,' in the case of the soprano voice, an innovation which began to be applied by subsequent authors to other voice types as well.

The names 'chest' and 'head' were based on where the sound appeared to be formed or resonated, although the mechanisms of the registers were a matter of conjecture. Manuel Garcia, a famous teacher and easily the most influential theoretician of singing in his century, took up the definition from the *Méthode*, refining it considerably:

> By the word register we mean a series of consecutive and homogeneous tones going from low to high, produced by the development of the same mechanical principle, and whose nature differs essentially from another series of tones equally consecutive and homogeneous produced by another mechanical principle. All the tones belonging to the same register are consequently of the same nature, whatever may be the modifications of timbre or of force to which one subjects them (Paschke, 1984; Garcia, 1847).

This definition has withstood the test of time, being quoted as a point of departure in many, perhaps most, serious discussions of register, down to the present day. A key development vis-a-vis the definition given in the *Méthode* is Garcia's attribution of register to 'mechanical principle.' Later, after he had

---

[1] Volontieri sees three extraordinary factors behind this historic event: the end of the era dominated by the castrati, the transplantation of the Italian style of singing to the new center of opera in Paris, and the advent of a more robust style of singing within the Romantic opera. These developments threatened the continuity of orally transmitted traditional Italian practices, prompting the codification in printed form.

[2] Tosi, however, was writing about the castrato voice, and the chest/falsetto discontinuity he referred to was near the top of the soprano staff, in all probability analogous to '*secondo passaggio*' for female voices in this same range, and a far less challenging transition than the move between 'chest' and 'falsetto' in the laryngeal source of the mature male voice which has passed through the structural changes of puberty. Such are the confusions of terminology which arise in the absence of unequivocal definitions and the neglect of taxonomic differences.

were indeed produced by visibly diverse mechanical actions of the vocal folds.

Making visible the laryngeal basis of the distinction between chest and falsetto produced a fundamental change in the theory regarding registers. The presence of registers had usually been explained as the result of the voice's being produced in one or another part of the vocal tract[1], or as depending on the dominance of one or another 'resonance' – explanations that continued to have intuitive appeal for singers. Garcia's advances in exposing the mechanisms of voice production, however, introduced an era in which the theory of the singing voice took on a dominant 'physiological' element. The unification of the registers remained one of the principal goals of practical pedagogy, but after Garcia's discovery of the physiological basis of the chest/falsetto distinction, the requisite smooth joining of the registers tended to be modeled as a gradual and continuous transfer of *muscular tension* between those intrinsic laryngeal muscles identified as responsible for chest and falsetto, respectively. In the latter half of the 20th century, especially after van den Berg's experiments (Berg and Tan, 1959; Berg, 1968; Berg, 1963) with excised larynges, these have been designated as the thyroarytenoids and cricothyroids.

In the early 1970s, when singing teachers such as William Vennard (Vennard, 1967) and John Large were involved in experimental investigations of the singing voice, there was a movement toward an 'integrated' theory of the registers, (Vennard and Hirano, 1973; Vennard et al., 1980; Large, 1972) taking into consideration not only laryngeal adjustments, but also resonance ('acoustic') adjustments as determining features of registers in singing. The basis for this had been laid by the arrival of spectrum analysis and acoustic phonetics, particularly in the work of Gunnar Fant (1960). Johan Sundberg's study (1977b) showing how sopranos tuned the first formant to the fundamental frequency in their upper range was a breakthrough along these lines that found general acceptance, although his intentions apparently did not include the designation of this segment of the female range as a register. The present volume identifies it as such and describes two further segments of the singing voice as registers with an acoustic basis: the male full head register (Chapter 7), and the highest female singing register, the flageolet (Chapter 6).

---

[1] For example, in one of the treatises of the period, De Garaudé described the three registers of the soprano voice as being produced, respectively, in the chest, the upper part of the larynx, and the frontal sinuses (De Garaudé, 1837).

The movement toward the integrated theory ended before it could bear much fruit, and since that time singing teachers have been less prominent on the frontiers of voice science. The concept of register adopted in the present work, however, is clearly one that points to such an integrated theory.

The historical literature is not heavily emphasized in these pages. This is not because it would be irrelevant to know what Garcia meant with 'falsetto' (the use of which he approved for operatic tenors), or what his contemporary Gilbert-Louis Duprez, the first tenor reported to make public use of the '*ut de poitrine*' (high C in chest voice), meant with '*voix mixte fortissimo*'; on the contrary, the history of the theory of registers could make a fascinating book, not only from a musicological standpoint, but also for its value to practical pedagogy (for an historical account with greater detail see Stark's excellent study [Stark, 1999]). Teachers of the present have much to learn from the insights of masters like Garcia, since the physical basis of the human voice is presumably essentially the same as it was in his day. Operatic singing may have become more robust in the course of the nineteenth century, evoking some novel pedagogical approaches (e.g., Staumethode [Armin, 1931]), but the raw material of voice has not changed.

## The aim of this work

The present work, however, aims not at translating the terminology of the methods of *bel canto* into a more contemporary singers' jargon, but at providing a map of the physical reality which the historical theories of registers point to, in their varying and inexact terminology. It presents, in terms which are empirically verifiable, a schema of what are offered as the determinative mechanisms of the registers most commonly recognized by voice teachers, whether they be acoustic or laryngeal.

The chief problem with the theory of registers has not been one of uncontrolled proliferation of terminology, but of the failure to identify and specify the mechanisms distinguished by the terms. The position taken in these pages is that when the *phenomena* are understood by singing practitioners, the terms used to designate them will be a matter of secondary importance. And so long as the phenomena are not understood, the theories describing them will remain in a state of flux and be debated with inappropriate fervor. There is a children's game, 'rock, scissors, paper', where each of the items designates a form of the hand – respectively, a closed fist, two extended fingers, and a flat palm. These are presented simultaneously by the participants, and the object is to triumph over your opponent's offering: rock breaks scissors, paper covers rock, and scissors cut paper. Similarly, in the world of voice teaching, it is always possible (and tempting) to move beyond the received interpretation: a two-register system is refined by adding further and more subtle distinctions; the many registers yield to the ideal system of the 'single register'; this cannot remain ascendant for long before it founders on the rock of a discontinuous voice which resists unification. Thus the spiral never ceases, as new and slightly different voices enter the game and move it one phase further, altering or reinventing the terminology.

What is attempted here is a description and classification of the various registers in terms of objective measurements, as well an account, likewise in objective measurements, of the 'registration events' which characterize the move from one register to another. All such empirical work is of course subject to refinement and correction, either in part or on the whole, by anyone else with comparable equipment.

Terminology is not the primary issue, but value is nonetheless seen in terms which connect with the historic tradition, e.g., 'head' (Chapter 6), or have an intuitive connection with the mechanism identified, e.g., 'flageolet' (Chapter 10). It appears inadvisable to create new terms (such as 'loft,' which seemed to be in favor within recent memory, but is now in decline [Miller, 1986]) for the purpose of having a term free of unscientific historic associations[1].

## Measurement techniques

Prominent in any empirical investigations of this kind are the instruments and methods of measurement. Where the subject is vocal function, there are various possible points of access along the 'speech chain', from the intention of the speaker to the perception of the listener (Rothenberg et al., 1988). Our measurements have been largely concentrated in the physiological, aerodynamic, and acoustic realms, monitoring, respectively, the movements of the speech organs, air pressures and flows, and sound pressures and spectra of the acoustic signals. The measurements reported in these pages include the following well-established techniques: spectrum analysis; electroglottography (measuring vocal fold contact area – see glossary) (Childers and Krishnamurthy, 1985; Baken, 1992); esophageal pressure by esophageal balloon (from which mean subglottal pressure can be estimated) (Schutte, 1980);

---

[1] This use of uncontaminated terms also has a long tradition: F. Bennati, the physician who cared for the singers at the Paris Opera, was careful to avoid unscientific and misleading terms in his *Etudes physiologiques et pathologiques sur les organes de la voix humaine*, identifing the registers as 'one' and 'two' (Bennati, 1834). All such efforts to the contrary, the misuse of terminology is inevitable.

mean airflow rate (Isshiki et al., 1967); and stroboscopic visualization of the vocal folds (Kitzing, 1985). In addition, we have done pioneering work in applying the following measurement techniques to the singing voice: magnetic resonance imaging (Sulter et al., 1992) (see also Chapter 10); direct measurements of sub- and supraglottal pressures with wideband pressure transducers (Miller and Schutte, 1985); estimating formant frequencies in singing with the help of a non-periodic voice source (Chapters 3 and 10). Other methods used, but not reported here, include inverse filtering, phonetography, X-ray photography, and videokymography (a new technique developed in the Groningen laboratory). Significant techniques reported in the literature, but not attempted in our series of measurements, include electromyography, high-speed cinematography, (spectrum) analysis by synthesis, and computer modeling of the vocal folds.

The measurements reported were all made in the Voice Research Lab in Groningen, under Professor H.K. Schutte, in the years 1984-2000. They began with the pressure measurements employing a catheter passed through the glottis (see Chapter 5), the most invasive of these methods. Some previously uncharted signals resulted from these measurements, which were made on a number of subjects, including other professional singers. Since that time, it has been our experience that once the less accessible parameters (for example, the time-varying acoustic pressures below the glottis) have been investigated, they can often be estimated on the basis of more accessible parameters. Thus it has become possible to locate the determinative mechanism of, say, a particular register, with methods that are increasingly less invasive. An example of this process can be found in Chapter 10, where MRI is used to gather direct information, which can thereafter be inferred from less intrusive spectrum analysis.

It should not be overlooked that all the measurements were made under the monitoring ear of the author, who, as already noted, claims an appropriate singing practitioner's expertise. While this characteristic of the author should have no effect on the (objective) measuring process itself, it plays an important role in the selection of material for presentation, where the intention is always to find voices and tokens which are *representative* of the phenomena in question. Related to this is the fact that the author underwent all the measurement techniques as a subject, and his data were often included among those reported.

## Technology and the Practice of Vocal Pedagogy

The author finds a close correspondence between the acoustico-physiological mechanisms described in these pages and the theory of registers as it is found in oral tradition and at least some historical treatises on singing. The correspondence, of course, cannot be exact. Both the lack of precision in singers' language and the disagreements – not merely semantic – among the various authorities preclude any exact correspondence. One way to attempt a verification of this correspondence would be a perception test with expert listeners. Beyond the problem of authenticating expertise (a problem that also applies in the case of the author's claim to this expertise), there is the still greater problem of isolating, in the sound examples to be presented to the judges, the factors considered determinative. Such isolation is only possible with a synthesized signal, and it is doubtful whether perceptual judgments of synthesized signals can yield definitive information concerning subtleties in actual singing (Carlsson and Sundberg, 1992). The real test of validity for the descriptions and definitions of registers offered here will have to come from the practice of singers and voice teachers, once they have access to the sorts of signals shown in these pages. If the patterns of the signals agree with what they hear as distinctions of register, then the signals will become part of their definitions, supplementing and clarifying their hearing. If they find that the signals are irrelevant, or that their hearing contradicts the explanations offered here, they will reject these explanations.

The last three years have brought advances in technology to the point where signals such as spectrum analysis and electroglottography can be routinely available in voice studios (chapters 12 and 13 on VoceVista [Miller and Schutte, 1999a; Miller and Schutte, 1999b]). Such technology non-invasively gives information concerning the glottal source and the vocal tract that for centuries had been hidden. It may be that this will bring about a minor revolution in singers' language, which will then relate what the ear hears to concrete, measurable patterns of glottal closure and resonance adjustments, instead of to quasi-imaginary spaces and 'resonances' in the vocal apparatus. This could mean a certain emancipation of the singing practitioners from the 'priests' of the medical profession, who lose some of their monopoly of knowledge of hidden bodily processes. But it will also emancipate the pupils of singing from the private and sometimes absurd claims of the maestri, since both master and pupil will have access to the same objective information.

## References

Armin, G. (1931). *Die Technik der Breitspannung*. Berlin-Wilmersdorf: Verlag der Gesellschaft für Stimmkultur.

Baken, R. J. (1992). Electroglottography. *Journal of Voice*, 6 , 98-110.

Bennati, F. (1834). *Etudes physiologies et pathologiques sur les organes de voix humaine*. Parigi. (Cited after Volontieri 1995) trad it. Milano 1834.

Berg, J. v. d. (1963). Versuche mit menschlichen Kehlkopfpräparaten. *Monatschrift für Ohrenheilkunde und Laryngo-Rhinologie*, 97, 522-527.

Berg, J. v. d. (1968). Sound Production in Isolated Human Larynges. In M.Krauss (Ed.), *Sound Production in Man* (155 ed., pp. 18-27). New York: The New York Academy of Sciences. Notes: Conference Sound Production in Man, held by The New York Academy of Sciences on November 7, 8, and 9, 1966

Berg, J. v. d. & Tan, T. S. (1959). Results of Experiments with Human Larynxes. *Practica Oto-Rhino-Laryngologica*, 21, 425-450.

Carlsson, G. & Sundberg, J. (1992). Formant frequency tuning in singing. *Journal of Voice*, 6, 256-260.

Childers, D. G. & Krishnamurthy, A. K. (1985). A critical review of electroglottography. *Crit Rev Biomed Eng*, 12, 131-161.

De Garaudé, A. (1837). *Méthode complète de chant ou Théorie pratique de cet art mise à la portée de tous les professeurs*. (opus 40 ed.) Parigi: presso l'Autore. (Cited after Volontieri)

Fant, C. G. M. (1960). *Acoustic Theory of Speech Production*. The Hague: Mouton & Co.

Garcia, M. (1847). *Traité complet de l'art du chant*. Paris: Chez l'Auteur.

Helmholtz von, H. (1954). *On the Sensations of Tone*. New York: Dover Publications, Inc.

Hines, J. (1982). *Great Singers on Great Singing*. New York: Doubleday & Company Inc.

Isshiki, N., Okamura, H., & Morimoto, M. (1967). Maximum phonation time and air flow rate during phonation: Simple clinical tests for vocal function. *Annals of Otology, Rhinology and Laryngology (St Louis)*, 76, 998-1007.

Kitzing, P. (1985). Stroboscopy : a pertinent laryngological examination. *J Otolaryngol*, 14, 151-157.

Large, J. W. (1972). Towards an integrated physiologic-acoustic theory of vocal registers. *The NATS Bulletin*, 29, 18-40.

Mancini, G. B. (1777). *Pensieri e Riflessioni sopra il Canto figurato*. (3 ed.) Milano: Giuseppe Galeazzi. (Cited after Volontieri)

Martienssen-Lohmann, F. (1981). *Der wissende Sänger: Gesangslexikon in Skizzen*. (3 ed.) Zürich: Atlantis Musikbuch-Verlag.

Miller, D. G. & Schutte, H. K. (1985). Characteristic Patterns of Sub- and Supraglottal Pressure Variations within the Glottal Cycle. In V.L.Lawrence (Ed.), *Transcr XIIIth Symp Care Prof Voice* (pp. 70-75). New York: The Voice Foundation.

Miller, D. G. & Schutte, H. K. (1999a). The Use of Spectrum Analysis in the Voice Studio. In G.Nair (Ed.), *Voice Tradition and Technology. A-State-of-the-Art-Studio* (pp. 189-210). San Diego: Singular Publishing Group.

Miller, D. G. & Schutte, H. K. (1999b). The Use of the Electroglottograph (EGG) in the Voice Studio. In G.Nair (Ed.), *Voice Tradition and Technology. A-State-of-the-Art-Studio* (pp. 211-225). San Diego: Singular Publishing Group.

Miller, R. (1986). *The Structure of Singing: System and Art in Vocal Technique*. New York: Schirmer Books.

Paschke, D. V. A. (1984). *Complete Treatise on the Art of Singing: Part One*. New York: Da Capo Press. (Translation of Garcia, 1847)

Rothenberg, M., Miller, D. G., & Molitor, R. (1988). Aerodynamic Investigation of Sources of Vibrato. *Folia Phoniatrica*, 40, 244-260.

Salzmann, Z. (1993). *Language, Culture, and Society*. Boulder, (CO): Westview Press.

Scherer, K. R. (1991). Emotion expression in speech and music. In J.Sundberg, L. Nord, & R. Carlson (Eds.), *Music, Language, Speech, and Brain* (pp. 147). London: MacMillan.

Schutte, H. K. (1980). *The Efficiency of Voice Production*. Thesis University of Groningen.

Seidner, W. W. & Wendler, J. (1997). *Die Sängerstimme: Phoniatrische Grundlagen für die Gesangsausbildung*. (3., erw. Aufl. ed.) Berlin: Henschel Verlag.

Stark, J. (1999). *Bel canto: A history of Vocal Pedagogy*. Toronto: University of Toronto Press.

Sulter, A. M., Miller, D. G., Wolf, R. F., Schutte, H. K., Wit, H. P., & Mooyaart, E. L. (1992). On the relation between the dimensions and resonance characteristics of the vocal tract: A study with MRI. *Magnetic Resonance Imaging*, 10, 365-373.

Sundberg, J. (1977a). Synthesis of Singing. In V.L.Lawrence (Ed.), *Transcr VIth Symp Care Prof Voice* (pp. 14-16). New York: The Voice Foundation.

Sundberg, J. (1977b). The acoustics of the singing voice. *Sci Am*, 236, 82-92.

Tembrock, G. (1996). *Akustische Kommunikation bei Säugetieren: die Stimmen der Säugetiere und ihre Bedeutung*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Titze, I. R. (1973). The human vocal cords: a mathematical model. Part I. *Phonetica*, 28, 129-170.

Titze, I. R. (1974). The human vocal cords: a mathematical model: Part 2. *Phonetica*, 29, 1-21.

Tosi, P. F. (1987). *Observations on the Florid Song*. London: Stainer & Bell. Notes: Translation from original: Opinioni de' cantori antichi e moderni (1723), translated by J.E.Galliard

Vennard, W. D. (1967). *Singing: the mechanism and the technic*. New York, N. Y.: Fischer.

Vennard, W. D. & Hirano, M. (1973). The physiological basis for vocal registers. In J.W.Large (Ed.), *Vocal Registers in Singing* (pp. 45-58). The Hague / Paris: Mouton & Co.

Vennard, W. D., Hirano, M., & Ohala, J. J. (1980). Chest, head, and falsetto. In J.W.Large (Ed.), *Contributions of Voice Research to Singing* (pp. 165-184). Houston, Texas: College-Hill Press.

Volontieri, E. F. (1995). *Le Regole della 'Canora Repubblica'*. Firenze: Firenze Libri.

# Chapter 2

# Synthetic Overview of the Registers

*Combining generally accepted knowledge with new material taken from the published and submitted articles (Chapters 3-10), the author presents a synthesis of the major registers employed by male and female voices in the western classical tradition.*

## Registers and registration

### Registration phenomena: voice source or vocal tract?

Much confusion and disagreement concerning registers in the singing voice (the number of registers, differences between male and female registers, etc.) arises from the fact that it is possible to approach the question from opposite directions: either as elaboration on the primary empirical fact of the natural registers – 'chest' and 'falsetto'[1] – or, alternatively, as a fine classification of discontinuities in sound that must be smoothed in order to produce the 'seamless scale' and the *messa di voce* (crescendo-diminuendo).

From the first approach has come the definition of registration as an exclusive feature of the voice source, a vibratory pattern of the vocal folds (Hollien, 1974). Empirical support for this view has come, since the middle of the nineteenth century, from laryngoscopic evidence of the difference between 'chest' and 'falsetto,' terms that have a considerably longer history.

An alternative to this is suggested by the experience, as well as the language, of singers and voice teachers, who historically have made 'third registers' an important feature of the map of registration, insisting that there is more to registration than merely 'chest' and 'falsetto.' This approach supports an 'integrated' approach to registration, including adjustments of the vocal tract, as well as those of the source (Large, 1972).

The approach that will be followed here is a pragmatic one. While the aim in describing registration phenomena is always to discover their physical properties, verifiable by instrumental measurement, the phenomena to be selected as register categories are chosen according to their importance in the experience and perception of singers, as manifested in the pedagogical literature.

---

[1] The use of the terms 'chest' and 'falsetto' (in single quotation marks) in this chapter is restricted to designating two categorically different vibrational patterns of the vocal folds. The terms are not identical to the chest and falsetto registers of the singing voice, which refer not only to the voice source, but as well to attendant uses of the vocal tract.

Thus we select what appear to us to be the four most significant discontinuities in the male and female singing voices, respectively, from the perspective of the 'classical' western tradition. Since in some cases the essential features of the discontinuities are identified with resonances of the vocal tract, we have obviously chosen the 'integrated' approach.

The names applied to the phenomena will not please everyone, even among those who agree that the particular phenomena in question are accurately described. In choosing terms, we have kept two considerations in mind. The first of these is that, while greater clarity is achieved by introducing a new or neutral term, much is lost in giving up terms that are widely used or make a direct connection to the (historical) pedagogical literature.

The second is that one cannot prevent terms that are only dimly understood from being used and taught. This implies that images immediately suggested by the terms need to be taken seriously, whether they are misleading, or, more happily, show evidence of the intuition of singers' language. In any case, it is hoped that a clearer objective understanding of the phenomena of registration will reduce mere allegiance to terms that sometimes characterizes pedagogical disputes, particularly those over registers.

Before we turn to the phenomena themselves, a further general observation is in order. As we have already mentioned in the introductory chapter, one of the principal aims of most voice training is the elimination, or at least minimization, of perceptual discontinuities between apparent segments of the voice, whether in frequency or intensity. Since the classical vocal literature demands extensive ranges, professional singers are expected to be able to pass the borders smoothly between some of the major registers described below, making the transitions, in an ideal case, imperceptible to the listener. Thus the pedagogical literature, where most of the theory of the singing voice can be found, puts a strong emphasis on eliminating the distinctions between the registers.

In contrast to this pedagogically motivated attention to smooth transition, here we shall first pay attention to sung sounds that belong firmly within one registration category or another, in order to establish the salient and determinative features of the categories. Only then can we identify correctly the problems associated with moving from register to register, as well as the more complex question of 'mixing' the registers.

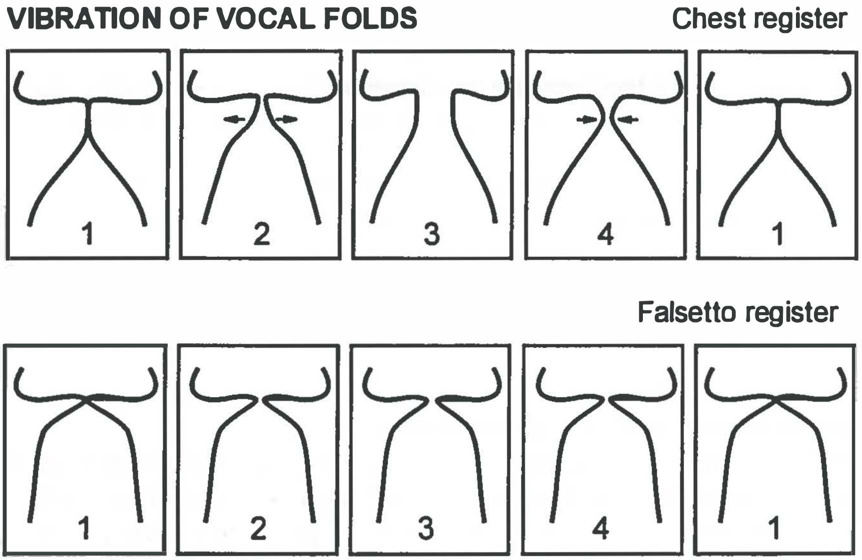


**Figure 1**
Schematic view of vocal-fold vibration. Above: 'chest' pattern shows vertical phase difference, where the closing of the lower margins precedes that of the upper. Below: in the 'falsetto' pattern the lower portion of the vocal folds is essentially stationary, and vibration is limited to the upper margins.

### *Registration phenomena: the natural registers*

We use the term registration phenomenon to refer to any readily perceptible discontinuity that may occur along an (ideally smooth) continuum of pitch or loudness. The most striking example, simultaneously the model, of this is the break from 'chest' to 'falsetto' in an ascending scale, a change that occurs principally in the glottal source.

The great majority of voices, both male and female, exhibit a primary register transition (PRT) somewhere within the total F0 range. In spite of more and less subtle differences between male and female voices with respect to this 'break' (explored in Chapter 9), there is general agreement regarding some distinctions between these natural registers, which shall be designated 'chest' and 'falsetto.' Among these distinctions are the following: 'chest' is greater

1. with respect to closed quotient (usually >50% in moderately loud singing, compared with <40% for 'falsetto');
2. with respect to depth of contact between the vocal folds ('chest' is characterized by vertical phase difference – the upper margins of the vocal folds lag behind the lower margins in the glottal cycle – while in 'falsetto' vocal-fold oscillation is limited to the thin upper margins (Fig. 1);

3 with respect to high-frequency components of the voice source (the spectrum of the voice source has a more rapid fall-off in 'falsetto');
4 with respect to the ratio between the forces applied by the muscles which contract (thyroarytenoid) and elongate (cricothyroid) the vocal folds.

In 'chest' the vocal folds are relatively short and thick and have a larger lateral excursion. (The longer vocal folds of 'falsetto' may seem counterintuitive, since long strings have low F0s; the greater stiffness and much reduced mass of 'falsetto' vocal folds, however, more than compensate for the increased length.) Finally, 'chest,' of course, also produces a lower F0 at a given level of effort.

### *Register transitions and the register violation*

Besides this primary register transition (PRT), there are other discontinuities associated with changes in the source, as well a group characterized chiefly by changes in resonance. We shall consider both types in detail below. Before proceeding to specific registers, however, let us look for a moment at the general question of register transition.

If a tone is sustained at a comfortable level of loudness and F0 (say, a perfect fourth above habitual speaking pitch), one can expect a relatively easy, habitual adjustment of the vocal folds. If this same adjustment is maintained while the subject sings progressively higher pitches, the normal result, certainly for the untrained singer, is an increase in subglottal pressure, glottal resistance, and intensity of sound. If this is continued into the upper F0 range, the singer either 1) maintains the initial adjustment, resulting in somewhat forced upper notes, 2) 'breaks' to a 'lighter' register, sacrificing continuity, or 3) achieves the smooth transition prescribed by classical technique. In the third case, the continued rise in subglottal pressure and sound intensity is accepted as the natural result of singing higher pitches.

There is a natural reluctance to give up the initial adjustment as the pitch rises, and this leads to the first alternative given above, one in which the lower, 'heavier' register is maintained past the point where, ideally, it should have yielded to a 'lighter' register. This is the typical case of what we shall call a register violation. If it occurs in loud voice, as is usually the case, the excessive effort associated with it makes subsequent transition to a lighter register still more difficult. Thus it is generally tolerated only for the highest notes in a phrase, where the potential abrupt discontinuity in the scale is not exposed.

Although not all voices are equally disposed to produce this rather natural fault, it is commonly encountered and typically gets considerable attention in the training of the singing voice. The usual strategy in dealing with it consists of strengthening the lowest tones of the upper adjacent register and learning to avoid getting ‘locked’ into a particular adjustment for the top tones of a heavier register.

In both male and female voices there is a natural tendency to give up the more effortful 'chest' vibrational pattern at around 300 Hz, where it becomes difficult to produce soft tones without resorting to 'falsetto.' Because this natural point for the PRT occurs in a relatively low part of the female F0 range, where subglottal pressures are no more than moderate, and because the source adjustment from 'chest' to 'falsetto' is less radical in female vocal folds (see Chapter 9), the discontinuity between the female chest and middle registers is considered tolerable in the classical tradition, particularly after it has been disguised by training.

In another singing style, the register violation resulting from the maintenance of the 'chest' voice source beyond its easy range is familiar in the sound of so-called belting, a technique of singing which cultivates the loud female voice up to about 600 Hz (*D5*), and sometimes even beyond this point (see Chapter 4). In this case virtue is made from register violation, exploiting the louder tones and more robust emotions appropriate to them.

The male voice encounters the primary register transition high in its frequency range, at a point where the effort, and usually the loudness, of singing is already considerable. This increases the obviousness of the change in the character of the voice, and the abrupt transition to falsetto production is not generally acceptable in the classical tradition. The solution to this problem is found in allowing that fraction of the voice range that lies above the PRT to continue with a ‘chest’ voice source, but a different use of resonance (see Chapter 7).

This change of registers as a shift from one dominant resonance of the vocal tract to another constitutes the other basic type of register transition. In these cases the boundary of the lower, heavier register is reached when the vocal tract encounters a ceiling on the upward adjustment of the dominant resonance. In the case of the male chest register, this occurs when the (dominant) F1 reaches the practicable upper limit on an open vowel. If F1 is forced higher by raising the larynx, the result is a register violation, recognizable by a tone that is ‘open’ in a pejorative sense (see Chapter 7). The soprano faces the same sort of transition around *C6*, at the end of her full upper voice, where the temptation is also to 'open' the vocal tract and continue following the rising F0 with the first formant (see Chapter 6).

One can also speak of a register violation of the opposite sort, where the lighter register is carried down into the proper range of a heavier register. While such a result may be undesirable, engaging less than the full potential of the voice, its use poses less of a threat to vocal hygiene than does the push-from-below violation.

### *Messa di voce and the blending of registers*

A final point in this general consideration of registration concerns registers on the continuum of intensity. This has been of some historical interest, since softer tones have often been identified with lighter registration. Many teachers have regarded the *messa di voce*, a maneuver which swells a single tone from pianissimo to fortissimo and back again, as the touchstone of not only registration, but of singing technique altogether. Some of these teachers, however, have tended to see the ideal voice-source adjustment between 'chest' and 'falsetto' as a gradual one, with intermediate degrees characterized by some combination of these functions (Van den Berg, 1968). In the view presented here, even in the case of a voice where it is possible to achieve close similarity between them, 'chest' and 'falsetto' remain discrete vibratory patterns and cannot be combined, like black and white, to produce a shade of gray. If a tenor reduces the vibrating mass of the vocal folds within the 'chest' function, the resultant sound may have a greater resemblance to the female-imitative sound we recognize as falsetto. We would not call this, however, a combination of 'chest' and 'falsetto' vocal-fold functions. Like the father of laryngoscopy, Manuel Garcia II, we consider this an incorrect description (García, 1847).

Abruptness of transition between registers can be smoothed in two basic ways. One of these entails making two distinct functions more closely resemble one another: for example, the vibrating mass of the vocal fold can be reduced in the 'chest' function, as mentioned above, or, conversely, mass can be added in the 'falsetto' function, resulting in higher closed quotients (sometimes exceeding 50% in female voices). The other derives from the fact that the voice source and the vocal tract are to some extent independent of one another. The 'chest' voice source can be combined with the vocal tract of the female middle register, for example, as we shall see below. In order to describe the combination accurately, however, one must give a correct specification of the elements combined.

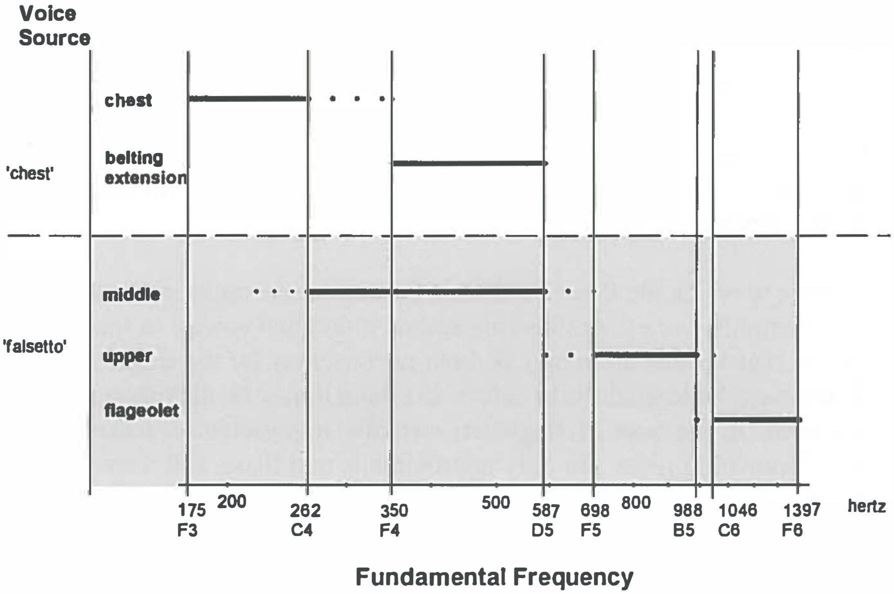


**Figure 2**
Overview with approximate ranges of the registers of the female singing voice. In the registers above the dashed line, the voice source has the 'chest' vibratory pattern; below the dashed line the vibratory pattern is 'falsetto.' The pitches where transitions between registers occur are given only approximately, and can vary by voice type (soprano, mezzo, etc.), as well as individually.

The voice segmentation reflected in registers is primarily one of the F0 range, with discontinuities in intensity taking a secondary place. In the following section, what we consider the eight most prominent male and female registers are selected for attention: only one of these, the male *mezza voce*, appears primarily as a discontinuity in the intensity dimension.

## **Registers in the female singing voice** (see Fig. 2)

Dividing the female F0 range into segments according to the most prominent registration events yields the following four registers:

1 Chest, the register used for most speech, from the lowest pitches up to the primary register transition at about 300 Hz. (*D4*), with a possible extension of about one octave (belting), to *D5*;

2 Middle, regarded here as the 'default' register of singing (insofar as this is not done in the chest register), comprising approximately the octave *D4*-

*D5* plus a transition to the next register, which is completed by *F5* (700 Hz), at the top of the staff;
3. Upper, the segment *F5-B5-flat* (with extension to *C6* [1046 Hz]), characterized by greatly reduced vowel definition and potentially high SPL;
4. Flageolet, the highest useful segment of the singing voice, a less effortful appendix to the upper register. In the flageolet register the dimension 'vowel' disappears entirely.

The degree to which the three transitions between these registers present barriers to continuity varies considerably among individual voices: in some voices a given register transition may be both unconscious for the singer and difficult for even the trained ear to detect; in others it may be disturbingly obvious or even, in the case of flageolet, virtually impossible to achieve. The points of transition given are only approximate, and these will vary according to the structure of the voice, a complex property including the build of both vocal folds and vocal tract, as well as the ability to vary the dimensions of these[1]. However, provided that the essential features of the registers and their transitions are understood, and that there is opportunity to confirm these with objective measurement, these individual variations can be readily described and characterized.

***Chest and middle*** (see Chapters 4 and 9)

The chest and middle registers can be best understood from a consideration of the changes that take place at the transition between them. This transition is identical to the primary register transition; thus the determinative change takes place in the voice source in the move from 'chest' to 'falsetto' function. Even if the contrast between 'chest' and 'falsetto' is generally less sharp in female than in male voices, the singer can usually identify the difference in the (proprioceptive) 'feel,' particularly if the tone is sung loudly.

This difference in the source is reflected in the sound as well. In accordance with the differences between the natural registers, a 'chest' source offers greater glottal resistance than a 'falsetto' source in the same F0 range, making it possible to increase subglottal pressure, and thereby loudness. Secondly, the 'chest' source creates a flatter source spectrum, and thus a greater high-frequency component in the sound, enhancing its audibility in the low F0 range.

---

[1] A more detailed listing of transition points, according to voice type, can be found, for example, in Miller, 1986; pp. 117, 134 and 135.

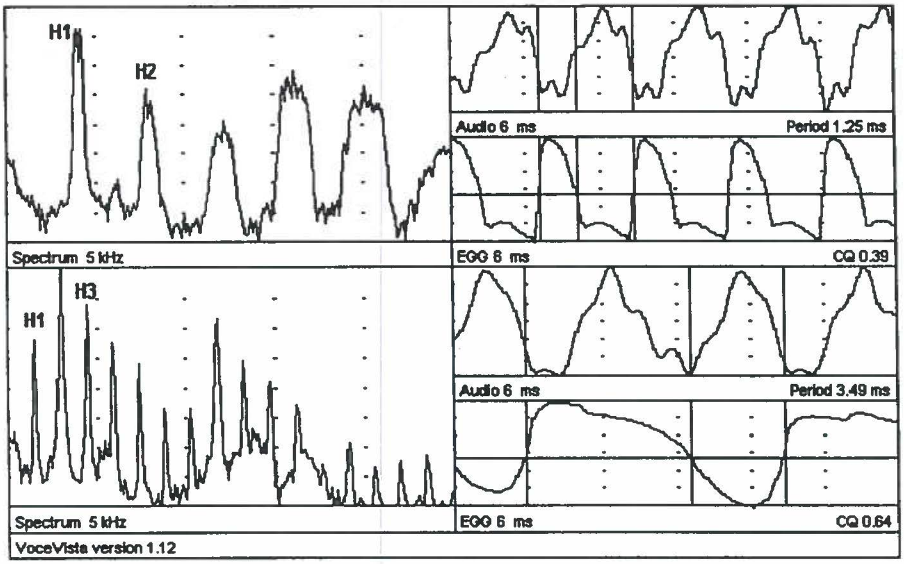


**Figure 3**
Push-pull effect in the open phase of the glottis. Power spectrum, audio signal and electroglottograph (EGG) signal. Above: soprano singing forte *A5-flat* (ca 800 Hz) on vowel /ɑ/. The relatively large open phase of the glottal cycle in 'falsetto' function (closed quotient is 39%), in combination with the dominant wave at the fundamental frequency (Fl = H1), is favorable for producing rising pressure (push) on glottal opening, as well as falling pressure (pull) on glottal closing. Below: Bass-baritone singing forte *D4* on vowel /ɑ/. The relatively small open phase in 'chest' function (CQ = 64%), in combination with the dominant wave at twice F0 (F1 = H2), produces a favorable push-pull on every second cycle of the dominant wave, while each first cycle resonates in the closed phase. N.B.: For a detailed explanation of the information displayed in this and other figures from VoceVista, see Appendix 1.

Along with these differences attributable to source, there are others deriving from the vocal tract. The >50% closed quotient of loud 'chest' makes this register especially efficient in resonating H2, providing F1 is tuned to it (see Fig. 3). Thus it is not surprising that we find this particular species of formant tuning on open (high-F1) vowels near the upper boundary of the chest register, both in female and male voices. In what we have described as a register violation, F1 is raised still further – typically by allowing the larynx to rise – in order to follow H2. The spectrum of such a tone shows the characteristic dominant H2, effected by an elevated F1, a sound that is described in the pedagogical literature as 'open' (Chapter 4, Figures 2 and 5). In the female voice it is characterized by a certain masculinity (Miller, 1986, p 136).

The transition from chest to middle offers several options for combining the registers. Because of the association of chest register (in ‘open chest’) with register violation, the vocal tract articulation that has a higher larynx (and F1) is perceived as characteristic of that register, while the lower larynx (and F1) is characteristic of middle. This makes it possible to combine the characteristic source of one register with the vocal tract of another.

The practical application of this is seen especially in the use of a low-larynx articulation of the vocal tract in combination with a 'chest' source. Approaching the PRT from chest register in a low-larynx articulation encourages the move to the middle register to come earlier and to appear less abrupt, in that the approaching change in the vocal tract has been anticipated. For similar reasons, but with the opposite effect, the abruptness of the change in sound is emphasized if the source and vocal tract of the chest register are maintained above the point of easy transition, as in the so-called ‘open chest’ production.

The other general type of mixing concerns adjustments in the source itself. While, as noted above, it is not accurate to speak of combining 'chest' and 'falsetto' functions in singing, it does seem likely that the widely held pedagogical view of a gradual adjustment between the two has some objective basis. The mechanism could be an adjustment, within the 'chest' function, in the mass of the vibrating vocal folds, with greater mass characterizing heavier registration. This adjustment presupposes a fine control over the vocal folds (tension is maintained while mass is reduced) that may or may not be learnable. We shall return to this in the discussion of male *mezza voce* (Chapter 8).

The female middle register is similar in several respects to the male full head register, whose F0 range (ca. 300-500 Hz) it overlaps. In general, the major differences between male and female voice are found in the voice source: the lengthening of the vocal folds of the male in puberty lowers his F0 range roughly an octave, and the ‘chest’ vibratory pattern becomes more robust. There are also secondary differences in the vocal tract, chiefly with respect to size, giving the male somewhat lower formant frequencies. The relative similarity of vocal tract in men and women, however, often makes it possible for both sexes to use the same resonance strategies for identical F0’s. Thus the strategy of second formant tuning (see Chapter 7) is often used by women in the middle register, enhancing ‘brightness’ in a range with natural disadvantages, vis-à-vis chest and upper registers, in regard to audibility.

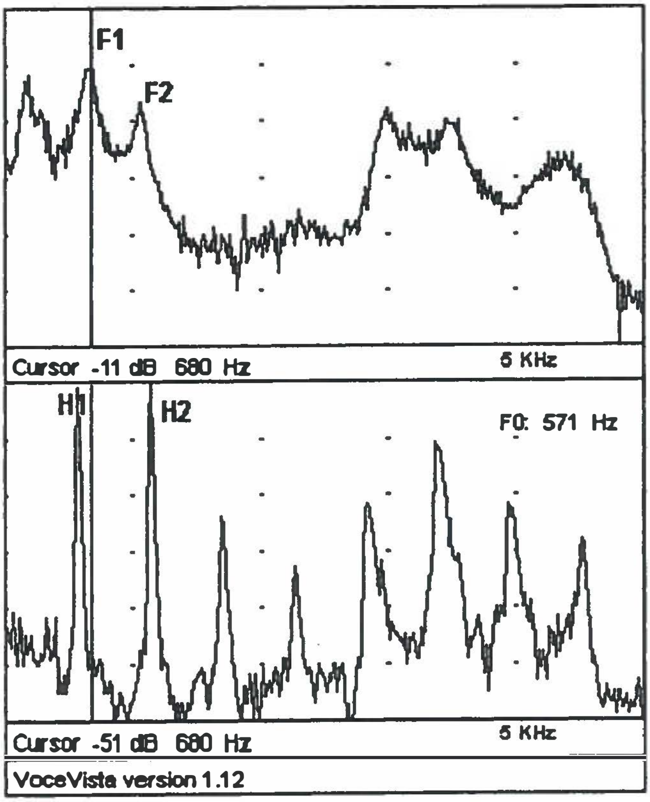


**Figure 4**
Soprano singing *D5* (ca. 600 Hz, near upper boundary of middle register) on vowel /ɑ/; power spectrum, audio signal, and EGG signal. The upper panel shows a continuous spectrum of vocal fry, produced with the vocal tract in the same posture as the sung *D5*, revealing the formant frequencies (see Chapters 3 and 10). The cursor is placed at F1, which has the frequency 680 Hz (the first peak in the upper spectrum is idiosyncratic, and below the range of the first vowel formant), revealing that Hl (the fundamental frequency) is still about 100 Hz below F1 and thus not maximally resonated. N.B.: For a detailed explanation of the information displayed in this and other figures from VoceVista, see Appendix 1.

At least two considerations limit the use of this male strategy by women in the middle register, however. The first is that the female voices are using a 'falsetto' source. The larger closed quotient of the 'chest' source used by male voices in this range gives a distinct advantage in resonating harmonics higher than H1 (see Fig 3). The other is the need for the female voice to prepare the transition to the upper register, where a dominant first formant will be tuned to H1. This transition is favored by a relatively low first formant, thus a 'darker' sound.

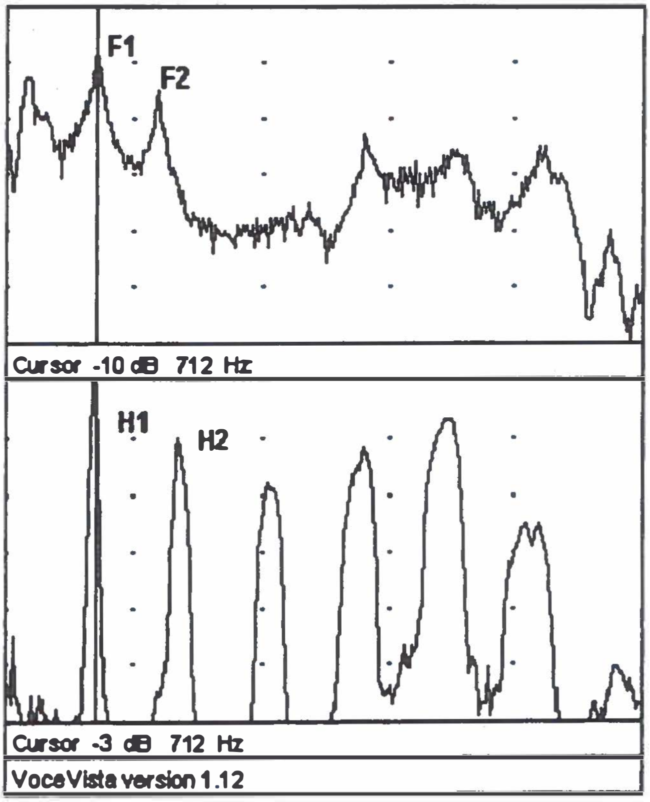


**Figure 5**

Soprano singing *F5* (ca. 700 Hz, at lower boundary of upper register) on vowel /ɑ/; power spectrum, audio signal, and EGG signal. The upper panel shows a continuous spectrum of vocal fry, produced with the vocal tract in the same posture as the sung *F5*, revealing the formant frequencies. The cursor is placed at F1 (712 Hz – the first peak in the upper spectrum is not a formant), revealing that H1 (the fundamental frequency) matches F1 and is thus fully resonated, rising at least 10 dB above neighboring harmonics. F2 now falls below the frequency of H2, reducing the level of that harmonic. N.B.: For a detailed explanation of the information displayed in this and other figures from VoceVista, see Appendix 1.

***Transition to upper*** (see also Chapter 5)

The determinative feature of the upper register is the natural tendency to tune F1 to H1 on the open vowels (see Figures 5 and 6).

F1-H1 tuning can also be effectively employed throughout the whole middle register, but only when the vowel has a first formant low enough to match F0. This can apply to the close vowels /i/ and /u/ all the way down to the normal point of PRT. The higher F1's of 'half close' vowels /e/ and /o/ cannot follow

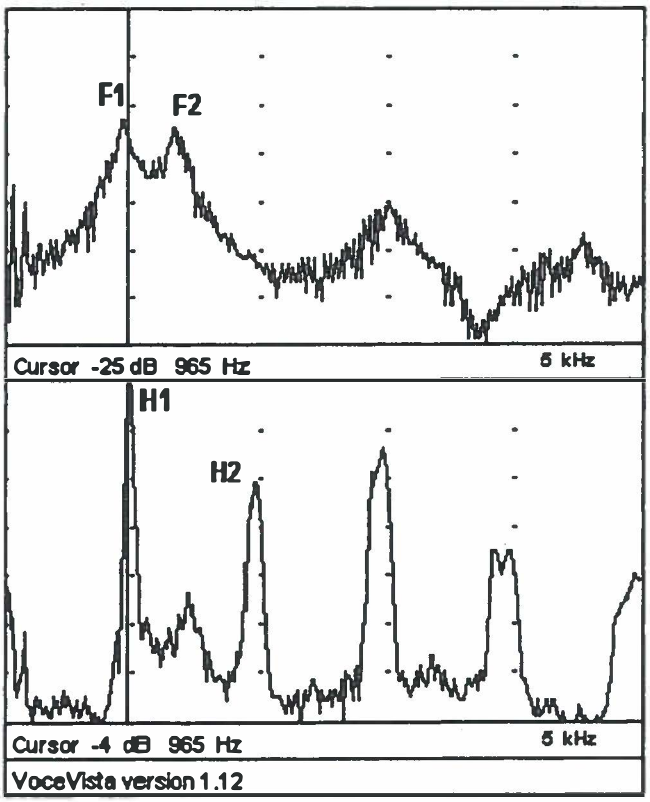


**Figure 6**
Soprano singing *B5-flat*, near the upper boundary of the upper register. The upper panel shows a continuous spectrum of vocal fry, revealing the formant frequencies. F1 aligns with H1 of the sung spectrum, which is 20 decibels above the level of H2. The slight discrepancy between F1 in the vocal fry 'imitation' and the sung H1 is presumably due to a small error in execution of the maneuver. N.B.: For a detailed explanation of the information displayed in this and other figures from VoceVista, see Appendix 1.

so far down the scale. Conversely, the close vowels must be 'opened' to keep pace with rising F0 (Sundberg, 1977), until all vowels have virtually the same F1 in the upper register.

At *D5*, nearing the end of the middle register, the first two harmonics are at approximately 600 and 1200 Hz. If the vowel is /ɑ/, the second formant (F2) is typically very close to the second harmonic (H2), while F1 is higher than H1. The result of this configuration is that H2 has an amplitude as strong or stronger than H1 (Fig 4). If there is a tendency to 'lock' on this use of resonance, moving F2 higher to follow H2 as F0 ascends, the singer will experience difficulty with the transition to the F1-H1 resonance strategy of the upper register. On front vowels, particularly /i/, the reluctance to give up an F2-H3 resonance can produce a similar register violation.

In the majority of voices the smoothing of the transition to the upper register can be accomplished with relatively little difficulty, making this the least problematical of the three female transitions. Nonetheless, it often gets (and deserves) the most attention in voice training, since all female voices in the classical tradition need to be able to negotiate it expertly, while simultaneously varying the color of the voice for expressive purposes.

### *Upper register*

The determinative attribute of the upper register, as we have already noted, is the tuning of the first formant to the first harmonic, without regard to where F1 might fall in speech. Not surprisingly, this has marked consequences for the intelligibility of the vowels: these, instead of being specified by two variable formant frequencies, all share a common F1, artificially determined by the sung pitch. This leaves only F2 – also somewhat constrained by the unnatural postures of the vocal tract on high notes – to differentiate between vowels.

The vowel is not lost entirely, however. According to the position of F2, a 'basic front vowel' can still be distinguished from a 'back' counterpart in which F2 is brought quite close to F1, helping to reinforce H1. Furthermore, skillful singers manage to convey considerable linguistic information in the quick transients between the steady-state vowel segments.

Although the contact area between the vocal folds diminishes as these are stretched and thinned with increasing F0, the ('falsetto') voice source in the upper register is not categorically different from the middle register. The closed quotient of this source combines well with F1-H1 resonance strategy, making it easy to produce high levels of sound pressure (see Figures 3, 5 and 6).

The tension in the vocal folds required to produce the high pitches helps provide the glottal resistance to build up sufficient subglottal pressure for loud sounds. There is, however, another, less evident, source of glottal resistance in this register, meriting our attention. It derives from the phase relationships (the ‘timing’) between the glottal cycle and the standing wave in the pharynx.

The vocal tract as resonator is in its most effective state when F1, reinforced by a proximate F2, resonates the fundamental frequency, produced by a ‘falsetto’ voice source. The powerful standing wave falls into a phase pattern such that the pressure difference across the glottis is maximal during the closed phase of the glottis and minimal in the most open phase, coordinating the ‘floodgates’ and the ‘tide’ in such a way that the flow is greatly reduced (see Chapter 5). Thus F1-H1 formant tuning creates an acoustic backpressure which limits flow, making it possible to sustain a normal phrase length, even with the weaker (muscular) glottal resistance of the 'falsetto' source.

A further consequence of being able to get powerful acoustic results without high glottal adduction is the facilitation of modulating the amount of contact between the vocal folds. From a maximal closed quotient of about 40-50%, skilled sopranos can reduce both the CQ and the vibrating mass of the vocal folds with relative ease, achieving seemingly effortless *pianissimi* without a radical change in ‘identity’ of the voice (see Chapter 8).

When we consider the special characteristics of this register – the highest and most efficiently produced SPL, the farthest removal from the acoustics of speech, the ‘magic’ of ‘floating’ soft tones – it is not surprising that it is often regarded as the classical singing technique *par excellence*.

***Flageolet register*** (see Chapter 6)
At a given pitch, when ascending through the upper register, F0 will reach F1 of the most open (in the sense of high-F1) vowel. This is not the end of the register, but rather the point after which the further raising of F1 (in tracking F0) must be implemented by an effective reduction in the size of the back cavity, usually by allowing the larynx to rise. Soon this extension, a sort of sanctioned register violation, is also exhausted. If the singer is to continue her ascent, she will have to allow F1 to fall below F0.

The resonance adjustment that allows her to continue is the primary characteristic of the flageolet register. In it the close approximation of F1 and F2, typical of the highest notes of the upper register, is maintained, providing a formant cluster within which H1 is resonated (Fig. 7). The basic two-formant vowel structure, compromised in the upper register, is now lost altogether, and the resultant tone has no more vowel than does a whistle.

In the typical flageolet range (*B5-F6*, or 1,000-1,400 Hz – the frequencies are those of the first two formants of the vocal tract in this register), the higher harmonics do not ordinarily fall near formants. This gives further relative prominence to H1, reducing the complexity of the sound. This may help to explain why loud flageolet sounds are seldom described as beautiful. Rather than for occasional loud high notes, this register is employed most effectively in the demonstration of control beyond the usual human limitations. Like the violin tone from which it borrows its name, its beauty is bound to an illusion of weightlessness. In this connection it has a characteristic voice source, one with greatly reduced modulation of contact between the vocal folds (Chapter 6, Figures 3, 4, 5), and apparently no complete closure of the glottis.

Since the flageolet has distinctive characteristics of both source and vocal tract, it is possible to combine it with the upper, much as we have seen chest and middle can be combined. One way this happens is the use of the upper-register source for loud tones in the flageolet range. While this may carry some danger of forcing, it is to be distinguished from the more aggressive forcing which pushes F1 beyond its limit in order to extend the upper register.

The other possibility for mixing involves a ‘reverse’ register violation. Here the light flageolet source is used in the range of the upper register, a tendency that is considered by some teachers to be destructive for the voice (Martienssen-Lohmann, 1981, p 285). In order to pass judgment on this view, however, one would have to distinguish clearly between the flageolet source and a pianissimo middle/upper source, a task that still remains to be accomplished.

## **Registers in the male singing voice** (see Fig 8)

The registers created by the most prominent discontinuities in the male singing voice have a different aspect from the female categories. This can be attributed to three basic considerations:

1 Because larger laryngeal structures of the voice source give the male a natural F0 level nearly an octave below that of the female, the primary register transition occurs in the upper part of the male, and in the lower part of the female F0 range;

2 The smaller difference between male and female vocal tracts (in contrast to voice sources) means that formant adjustments ordinarily take place in

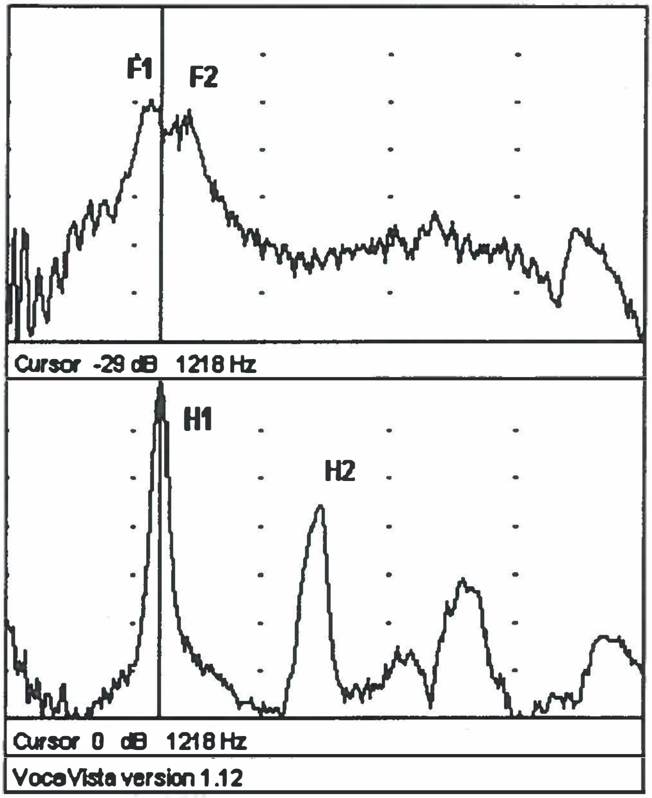


**Figure 7**
Soprano singing *E6-flat* in the flageolet register. The upper panel shows a continuous spectrum of vocal fry, revealing the formant frequencies. The cursor is placed at the F0 of the sung tone, 1218 Hz. F1 is ca. 100 Hz below F0, and F2 is ca. 200 Hz above it. The two formants create a cluster that strongly resonates H1, even where this has risen above the first formant. N.B.: For a detailed explanation of the information displayed in this and other figures from VoceVista, see Appendix 1.

the same octave (as, for example, in female middle and male full head);

3. The mechanical adjustment in moving from 'chest' to 'falsetto' voice source is notably greater in the male voice (see Chapter 9).

The major registers are these:

1. Chest, from the lowest pitches up to the PRT (roughly 300 Hz);
2. Full head (*voce piena in testa*), the F0 extension beyond the PRT, with a 'chest' source;

3 Falsetto, the upper portion of the total F0 range, produced with a 'falsetto' source, and showing substantial overlap with the range of chest register;
4 *Mezza voce*, a special adjustment of the 'chest' source, used to produce soft tones.

It will be noted that, of possible transitions between registers, only that between chest and full head is fixed within a narrow range of pitches, whereas all three of the prominent female transitions are so fixed. (Variability in the F0 location of these transitions due to differences in structure among individual voices applies equally to male and female.) *Mezza voce* is, at least theoretically, a distinction in the intensity dimension only, and independent of F0.

***Chest and full head*** (see Chapter 6)
In the male classical singing voice, the question of registers arises most problematically around the primary register transition. Some more or less subtle adjustments may occur within the chest register – roughly defined as the large F0 range, sung with a 'chest' source, below the PRT – but these adjustments present only minor obstacles to realizing the ideal of the continuous scale, when compared with the challenge of *passaggio*. This Italian term is commonly applied to the point of transition from chest to full head, as well as to the pitches just below it (*zona di passaggio*), where the transition is 'prepared,' often with minute pedagogical attention.

The upper limit of the chest register is presented by F1 of the open (high F1) vowels (thus 600-750 Hz), as this resonance is engaged by H2 in the 'call' of the voice (Fig. 3). The strong standing wave generated by the tuning of F1 to H2 produces the highest sound pressure levels in the chest register. In shouting there is a strong tendency of F1 to follow H2 still higher (by raising the larynx), resulting in the detrimentally open sound of the register violation. The trained singer may avoid this tendency by 'covering' (keeping F1 low), usually by some combination of lip-rounding and larynx-lowering (or at least not allowing the larynx to rise).

The full head register begins, for open vowels, at the point where F1 falls below H2, and it continues to the end of the F0 range of 'full voice' (not including falsetto). The source remains the 'chest' function, with the closed quotient, in some robust voices, even marginally higher than that of the chest register.

The perceptible distinction between chest and full head usually appears in the

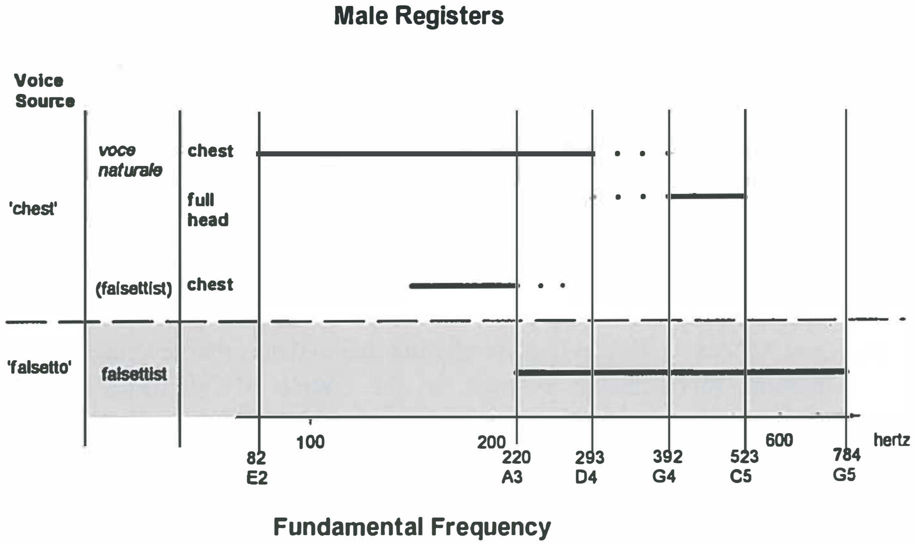


**Figure 8**
Overview with approximate ranges of the registers of the male singing voice. In the registers above the dashed line, the voice source has the 'chest' vibratory pattern; below the dashed line the vibratory pattern is 'falsetto.' The pitches where transitions between chest and full head registers occur vary by voice type (tenor, baritone, etc.), as well as individually.

use of resonance. In open vowels, up to the *passaggio* point the strongest resonance is most often F1. At this point F1 begins to fall away from H2, losing much of its effectiveness as a resonance, since there is no longer a harmonic very close by its center frequency. If the voice is to remain fully 'resonant,' a higher formant must find a harmonic to enhance[1]. The two most effective resonance strategies in full head are F2-tuning, where the second formant amplifies H4 or H3, or a powerful singer's formant, which can also carry the major part of the SPL (Fig. 9, see also Chapter 7, Figures 5, 7, and 9). A sound dominated by the singer's formant enjoys a considerable acoustic advantage by virtue of occupying a frequency region where there is little competition from most orchestral instruments (Sundberg, 1977). The high-frequency 'brightness' that is a desired element in the singer's formant, however, can also be supplied by F2-tuning.

A good command of the full head register is of course most important for the tenor voice, but other male singers who aspire to careers in opera must have

---

[1] except in the case of a close (low-F1) vowel, where F1 can resonate H1

it as well. Baritones need it frequently, but even basses, who accomplish *passaggio* at *D* or *E4-flat*, must master it in order to avoid the labored sound of register violation.

***Lighter registers in the male voice*** (See Chapter 8)
As the typical musical pattern of the yodel suggests, the interval between chest and falsetto registers, using similar levels of vocal effort, is a major sixth or an octave (see Chapter 9). If the two registers produce a tone with the same F0, the falsetto tone is, as a rule, so much softer and 'lighter' in registration that the perceived integrity of the voice is threatened (see Chapter 1, 'Historical Aspects'). Thus it is generally not deemed acceptable to use falsetto production, recognizable as such, in the course of full-voice singing (Miller, 1986). However, the soft tone that is relatively high in pitch can be an important, and even a necessary, part of the singer's technique. How are such tones to be produced?

One possible answer to this question is in the register which we designate *mezza voce*. As a species of the 'chest' source function, this register offers a contrast to full voice with less threat to the integrity of voice. It differs from the full voice, whether this is chest or full head, by reducing to a minimum both the oscillating mass of the vocal folds and the modulation of contact between them. In keeping with the softness of the sound desired, the subglottal pressure is very low, on the order of values occurring in conversation (see Chapter 8). In order to maintain oscillation at this low pressure, the glottal resistance, and thus the adduction of the vocal folds, must also be low, resulting in a glottis that does not fully close. As we have observed, this means a source spectrum with a steep fall-off, and a sharply reduced high-frequency component (see Fig. 3c in Ch. 8).

Producing this register in a reliable manner, especially at relatively high pitches, requires considerable skill, or at least a disposition of the instrument that may not be learnable. Less common than this is the ability smoothly to diminish the sound from forte to the pianissimo of *mezza voce*. The change from a completely-closing to a leaky glottis will be characterized by some abruptness, and for this reason we consider it one of the major discontinuities. (The approach to pianissimo in the 'falsetto' function, with its reduced oscillating mass, is considerably easier. For this reason we find no intensity adjustments among the most prominent discontinuities in the female voice.)

This register is sometimes called '*voix mixte*,' suggesting the commonly

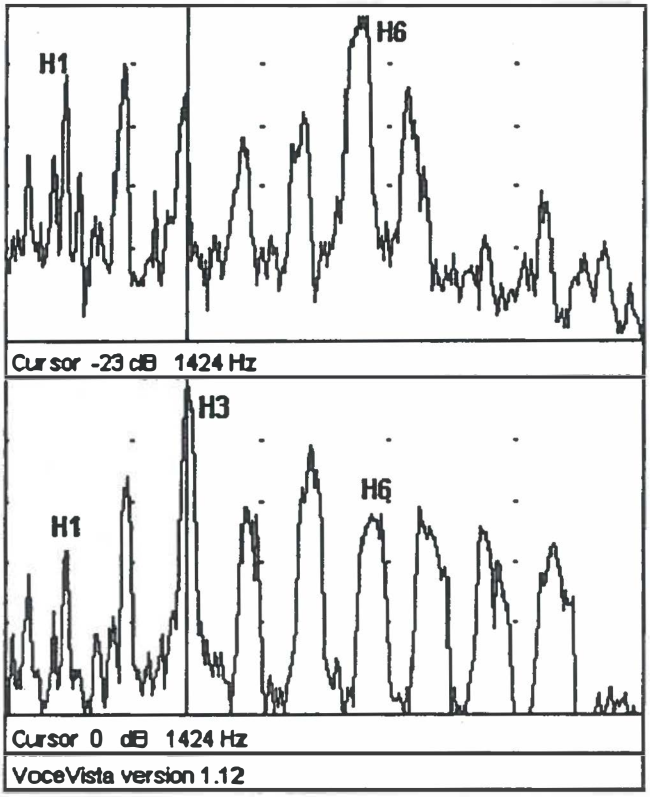


**Figure 9**
Two different resonance strategies for the final *B4-flat* of the aria 'Celeste Aida' on the vowel /o/, taken from commercial recordings. Below: the tenor Luciano Pavarotti, who relies on a dominant F2 – H3 resonance. Above: the tenor Placido Domingo, who favors the singer's formant strategy. N.B.: For a detailed explanation of the information displayed in this and other figures from VoceVista, see Appendix 1.

found view that it is a mixture of 'chest' and 'falsetto' functions. As we have already observed, such a combination of functions is not applicable in singing. Soft falsetto can be distinguished from *mezza voce*, both in its EGG pattern and (possible) complete closure of the glottis. Besides its low subglottal pressure, however, there is a further feature of *mezza voce* which is similar to falsetto: in many instances the vocal tract tends to adopt the F1-H1 resonance strategy appropriate to the small closed quotient.

The artistic use of falsetto, as in the practice of singers who call themselves variously countertenors, falsettists, and male altos, remains to be considered. Although such singers rely primarily on a 'falsetto' source, they must, like low-voiced women, at least occasionally cross the primary register transition and use a 'chest' source. What we have observed about the female voice in this regard applies here as well: the less subglottal pressure on the voice, the less apparent and disruptive the transition will be; thus one generally tries to keep the transition low in F0.

The resonance strategies appropriate to falsetto are similar to those used in the female middle and upper registers. Two differences in the structure of the (average) male and female voice should be kept in mind, however. The first is that the larger dimensions of the male vocal tract, especially the pharynx, give him lower natural F1 values, thus shortening the range of (the equivalent of) the upper register. The second is that the greater difference between the 'chest' and 'falsetto' source of the male (see Chapter 9) tend to handicap him, even more than the female, in the pitch range just above the transition. Here his intensity range with open vowels, where he cannot employ the F1-H1 resonance strategy, is usually limited.

## References

García M. *Traité complet de l'art du chant*. Paris: Chez l'Auteur, 1847.

Hollien H. On vocal registers. *Journal of Phonetics* 1974;2:125-143.

Large JW. Towards an integrated physiologic-acoustic theory of vocal registers. *The NATS Bulletin* 1972;29:18-40.

Martienssen-Lohmann F. *Der wissende Sänger: Gesangslexikon in Skizzen.* 3 ed. Zürich: Atlantis Musikbuch-Verlag, 1981.

Miller R. *The Structure of Singing: System and Art in Vocal Technique.* New York: Schirmer Books, 1986.

Sundberg J. The acoustics of the singing voice. *Scientific American* 1977;236(3):82-92.

Van den Berg Jw. Register Problems. In: Krauss M, editor. *Sound Production in Man*. New York: The New York Academy of Sciences, 1968;129-134.

# Chapter 3

# Feedback from Spectrum Analysis Applied to the Singing Voice

# Feedback from Spectrum Analysis Applied to the Singing Voice

Donald G. Miller and Harm K. Schutte

## Abstract

Despite the general availability in the last two decades of both the spectrum analyzer and prominent pedagogical theories concerning the conscious tuning of vowel formants to enhance the singing voice, there has been little reported use of spectrum analysis to track formant frequencies in singing. An important exception is Sundberg's (1973) work on the soprano voice. The reasons for this neglect are considered: in the singing range where information on formant tuning would be most helpful, the wide spacing of the harmonics renders the formants difficult to locate by spectrum analysis. Methods are described for obtaining continuous spectrograms with the vocal tract in the varied articulations of singing by using sweep tones and non-harmonic voice sources, and thus locating quickly and accurately the frequencies of the first five formants.

## Introduction

It is a curious fact that the spectrum analyzer has not had a greater impact on the theory and pedagogy of singing. The role of formants in the phonetics of speech and the quality of the singing voice has been understood for a number of decades (Bartholomew, 1934). William Vennard's widely read treatise, *Singing, the Mechanism and Technique*, with its prominent and effective use of sonagrams, is more than twenty years old (Vennard, 1967). Berton Coffin, another well-known singing teacher, developed a whole singing method based on consciously matching formants with the harmonics of the voice source (Coffin, 1976), although he neglected the further step of confirming the method in practice by using spectrum analysis. While they are still expensive, spectrum analyzers have been generally available since the mid-1970's, and many people have access to them; nonetheless little information based on experiment, about the tuning of the lower formants to achieve an optimal sound has appeared in the literature.

One significant exception has been Johan Sundberg's study of a professional soprano (Sundberg, 1973), from which he made the important observation that the soprano resonance is greatly enhanced by systematical adjustment of the lowest formant to match the fundamental frequency in her upper range. The investigative method used in this study, however, gives us an indication of why the spectrum analyzer has not been used more widely in locating formants in singing.

In order for the formants (resonances of the vocal tract) to make any significant contribution to the acoustical power of a tone, one of the harmonics (F0 and its multiples) must fall close to the center frequency of the formant. If the F0 is 100 Hz (about two *A-flats* below middle C) as in average male speech, then the distance between a formant and the nearest partial can be no more than 50 Hz, and this is close enough to insure that the formant will not be 'empty', or too far from any partial to produce actual resonance. As we go higher in F0, however, the location of formants becomes more critical. Three octaves above this male-speech pitch, at *A-flat* above the treble staff, the 800 Hz interval between harmonics means that a formant can easily fall where it is too far from a harmonic to produce effective resonance. From the spectrum analyzer we get a *line spectrum* of a sung tone, i.e., a picture of the frequencies and relative strengths of its component harmonics, whereas the formants are only revealed accurately in a *continuous spectrum*, where there are no gaps between harmonics. Locating the formants on the basis of the displayed harmonics becomes increasingly a matter of guesswork as the interval between harmonics gets larger with higher pitch, in other words, exactly as tuning the formants becomes more critical for the resultant sound in singing.

Sundberg (1973) dealt with the problem by substituting a non-harmonic sound source for the harmonic voice source. The soprano carefully held the vocal tract articulation used for a given tone but stopped singing, and a vibrator, applied to the thin outside wall of the pharynx on the neck, produced a non-harmonic sound inside the vocal tract, thus filling in the gaps in the spectrogram and allowing the determination of formant frequencies. Sundberg then confirmed these findings by the method of ‘analysis by synthesis’ or arranging the (artificial) formants of his voice synthesizer in such a way that the spectrum of the synthesized tone approximated the spectrum of the tone actually sung by the soprano.

Analysis by synthesis is hardly the answer to the problem of the use of the spectrum analyzer in singing, since it is a laborious process even for the very skilled. The approach with the vibrator on a silently held articulation, on the other hand, is relatively quick and straightforward. Its biggest drawback is that it produces a weak sound that is poor in high-frequency components, when compared with the actual voice source. Indeed, in some throats the sound that gets through the neck tissue is so damped that the configuration of formants is vague.

## Method and results

The purpose of this paper is to describe some relatively simple and effective methods of supplementing the information displayed in spectrograms of sung tones, especially with respect to formant frequencies. The trick is always that of filling in the gaps between the harmonics, thus producing a continuous spectrum in the frequency region of interest. For this we use two basic strategies: non-harmonic voice sources and sweep tones.

A very effective non-harmonic voice source can be produced by ingressive airflow through rather tightly approximated vocal folds. The glottal flow pattern of this source has sharply defined closed intervals of irregular length separating the puffs of inward-flowing air. It can produce a sound as loud as normal speech and with a spectral balance much like that of singing in chest voice, at least to 5 kHz, but some skill is required (just as in the vibrator method) to hold the articulation of the sung tone whose formant configuration is being investigated. The method is thus not foolproof, but it does give clear pictures of the transfer function of the vocal tract.

Other techniques we employ, described in more detail below, are vocal fry (as alternate non-harmonic source) and two sweep tones: a chromatic scale and a wide 'trill'.
For the experiment a FFT Real Time Spectral Analyzer, Model 4512 (Princeton Applied Research), incorporating a built-in Hanning weighting network, was used. Each of the spectra shows the average of at least two seconds of sound, and the scale passages take more than twice that long. The samples of spectra in the four figures were produced by one of the authors, a bass-baritone with considerable professional singing experience.

Figure 1 gives an example of a spectrogram produced by the ingressive source, along with some spectrograms of sung tones, chosen to illustrate the problem we address here. The vowel /ɑ/ is sung at three pitches: 100, 200, and 400 Hz (just above *G2*, *G3*, and *G4*). The closely spaced harmonics at the lowest octave permit us to find the approximate locations of the formants without much difficulty, at least in the cases of the first and second formants (F1 and F2). An octave higher the density of harmonics is only half as great, and for accurately locating the formants we need the information from the ingressive spectrogram, which was produced with the same vocal-tract articulation as the 200 Hz sung tone. When we consider the tone another octave higher at 400 Hz (but produced with the articulation of 200 Hz), it is apparent that the sung spectrogram alone cannot tell us even how many formants are present, let alone their frequencies and bandwidths.

Figure 2 gives another example of the vowel /ɑ/ at 400 Hz, but this time the vocal tract is allowed to adjust to a more resonant articulation by the 'natural' (for trained singers) process of vowel modification. An ingressive spectrogram of that articulation is then compared with the articulation for 200 Hz. We see that the improved resonance is achieved by dropping F1 close to the fundamental and concurrently lowering F4 to form a stronger singer's formant with F3.

Since it is not easy to hold precisely the same articulation for the ingressive phonation as for the sung tone, we can expect some error in our determination of formant frequencies. In order to get a clearer picture of the magnitude of this error, we introduce an alternative method of producing a non-har-

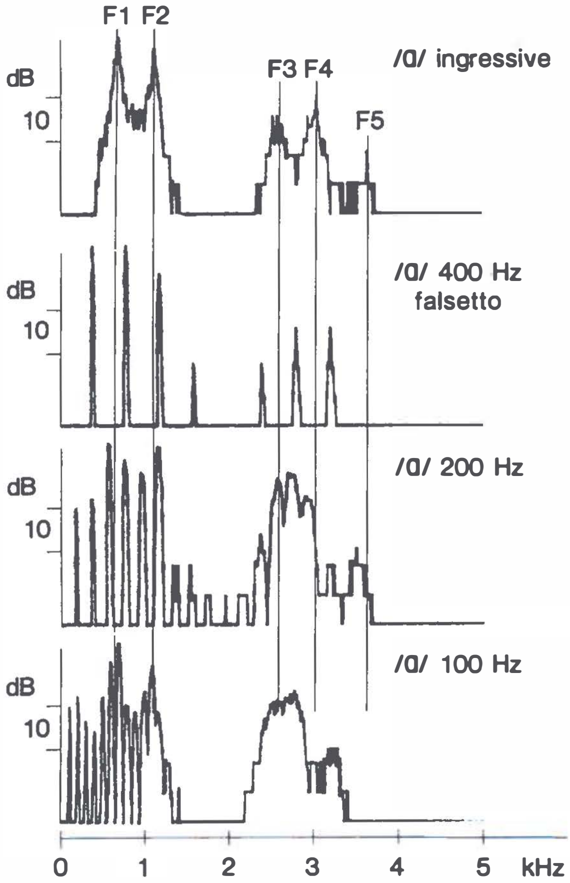


**Figure 1**
Spectrograms of the vowel /ɑ/ produced by a non-harmonic (ingressive) voice source and singing of three octaves of a pitch just above G. Both the ingressive and 400 Hz examples maintain, as far as possible, the articulation of the 200 Hz example. The 100 Hz example is produced with the singer's normal articulation for that pitch, with the result that formants 3-5 have lower frequencies.

monic source, that of vocal fry, or 'creaky voice'. Since it is egressive, we might expect a tendency to raise the larynx from its level in the sung phonation, in contrast with an expected tendency to lower the larynx in ingressive phonation. The two are shown together in Figure 3 in the articulation of a sung *D4* (292 Hz) on the vowel /e/. For the first two formants the two non-harmonic phonations are in substantial agreement, which is confirmed by visual inspection of the sung spectrogram, indicating that the magnitude of

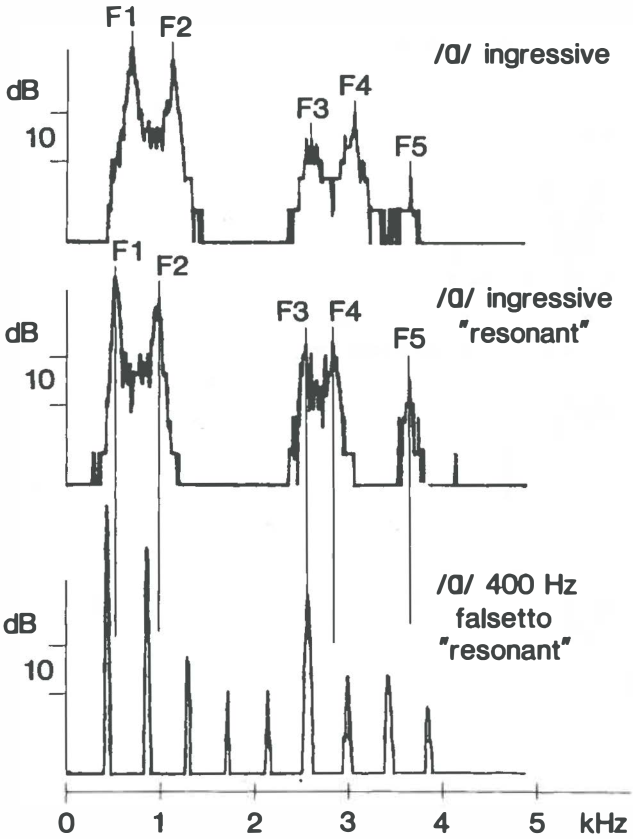


**Figure 2**
Another 400 Hz example, where the vowel has been modified to produce a 'more resonant' sound. Above it are ingressive spectrograms of this 'resonant' articulation and of the articulation used to produce the 400 Hz tone in Figure 1. The first and sixth partials are particularly enhanced by the altered configuration of formants, all of which are lower than those of Figure 1.

any error is small. The higher formants are less accurately displayed, particularly in the vocal fry phonation, which consistently overestimates their frequencies. The weaker sound of vocal fry also gives a less accurate reflection of the formant levels than does the ingressive phonation, which is fairly close to the sung phonation in this respect. Vocal fry, however, is easily produced, and in singer-subjects without previous experience in this technique we have had good results in locating the first two formant frequencies, i.e., those which come into question for conscious formant tuning in singing.

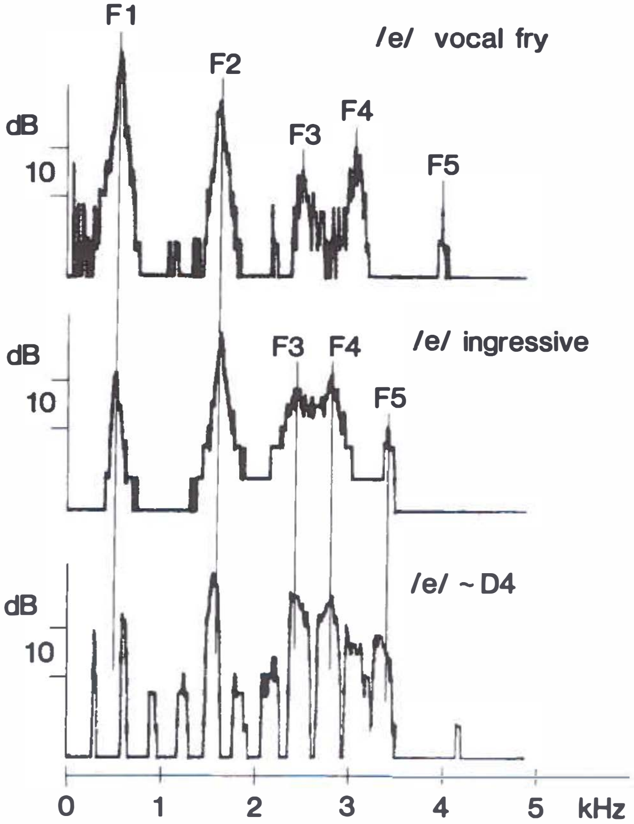


**Figure 3**
Spectrogram of the vowel /e/ sung at *D4* (292 Hz), along with those of two differing non-harmonic voice sources in the 'same' articulation. Correlation with the sung spectrogram shows that formants 3-5 are estimated too high by the vocal fry method. Both non-harmonic sources confirm the presence of a tuned F2 and a detuned F1.

Figure 3 contains further evidence of the practical value of feedback from spectrum analysis, in that it presents a good example of a sung tone in which the second formant is the one which is tuned (to the 5th harmonic), while the first formant is ineffectively positioned between the first two partials. Such dominant F2 tuning, which is not mentioned in any of the pedagogical treatises that recognise tuning of the vocal tract, was first described in a paper by the present authors given in 1987 at the 16th Voice Symposium (Miller & Schutte, 1990). Second formant tuning appears to have an important role in vowel modification near and above the transition from 'chest' to 'head' reg-

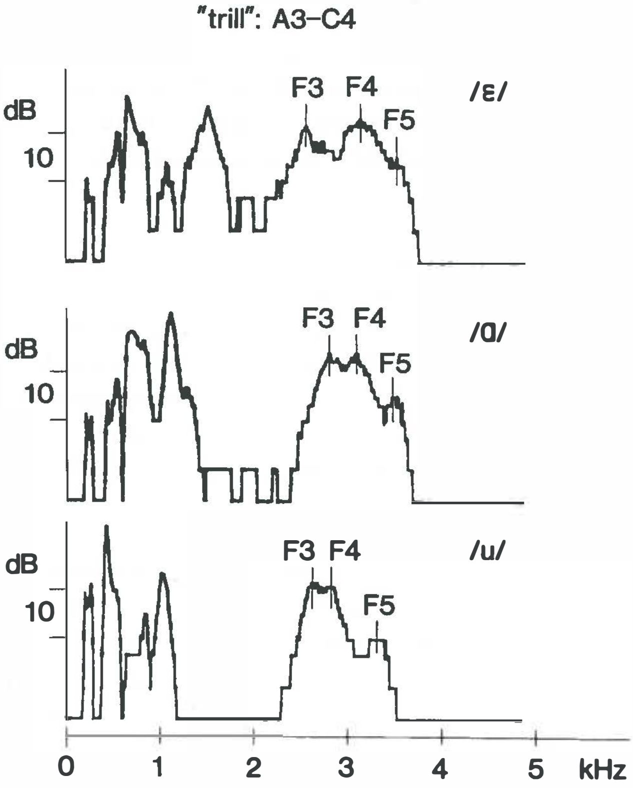


**Figure 4**
A few seconds of a wide 'trill' covering the frequencies 200-250 Hz, affording a clear picture of various configurations of the 'singer's formant', reflected in formants 3-5 in the vowels /ɛ/, /ɑ/ and /u/.

isters in the male voice, the details of which are beyond the scope of this paper.

While ingressive and vocal fry phonations give us a good picture of formants one and two, the accurate location of formants three to five (in whose range the so-called singer's formant will be found) presents somewhat different problems. On the one hand, the discrepancy between sung and non-harmonic articulations appears to be greater in the higher formants (in Figure 3, even in the relatively accurate ingressive phonation, F5 is noticeably higher than in the sung phonation). Fortunately, there is a convenient method (illustrated in Figure 4) for producing a kind of sweep tone, while actually singing, which allows us to locate the higher formants accurately. We know that the frequency extent of vibrato is multiplied in the higher harmonics (a vibrato frequency extent of 20 Hz at a fundamental frequency of 200 Hz, for example, produces an extent of 200 Hz at the 10th partial, thus 'filling in' all the spaces between partials above 2,000 Hz), making it possible to locate the higher formants directly from the sung phonation.

In Figure 4 the extent of vibrato is intentionally enlarged to make a 'trill' of a major 3rd, or about 25% of the fundamental frequency, which fills in the spectrum from the fourth partial upwards, or two octaves above the F0. The spectrograms of different vowels show each the average of a few seconds of such a 'trill' covering the frequencies 200-250 Hz. In practice this means that our spectrogram can give a complete picture of the higher formants for all sung tones but those of the female upper range (from the pitch F5 [700 Hz] and higher).

Figure 4 shows spectrograms of the vowels /ɛ/, /ɑ/ and /u/. The differences in the configuration of formants three to five are revealing: the /ɛ/ has the largest frequency spread of these formants, producing a 'wide' singer's formant; the /ɑ/ is more 'compact', and in the /u/, which is also 'compact', F3 and F4 are so close together, producing a strong synergetic effect of formants in close approximation, that the contribution of F5 to the singer's formant appears to be less important. The whole configuration of formants of /ɛ/ differs in the absence of a wide and deep trough in the frequency range between F1 and F3. Such spectral patterns are indications of complex perceptual differences in the 'quality' of these vowels at this frequency range in this voice, but more detailed analysis goes beyond the limits of this paper. The fact that the higher formants appear to have, individually, less sharp resonances than the lower formants is probably a characteristic of these formants under the conditions of singing, where their center frequencies may vary with small movements of the vocal tract in vibrato.

Figure 5 introduces still another method of filling in the gaps between partials to produce a continuous spectrum: the chromatic scale passage. In this case the scale includes all the semitones in the interval *E4-flat* to *B4-flat* (ca. 300 - 450 Hz.), assuring a continuous spectrum above 600 Hz. It is presented together with an example of the wide 'trill' covering approximately the same F0 range, as well as a single tone that falls within this range, all sung by a soprano on the vowel /ɑ/. (The single tone is *F4-sharp*, close enough to the 400 Hz falsetto tones of Figures 1 and 2 to allow the curious reader to

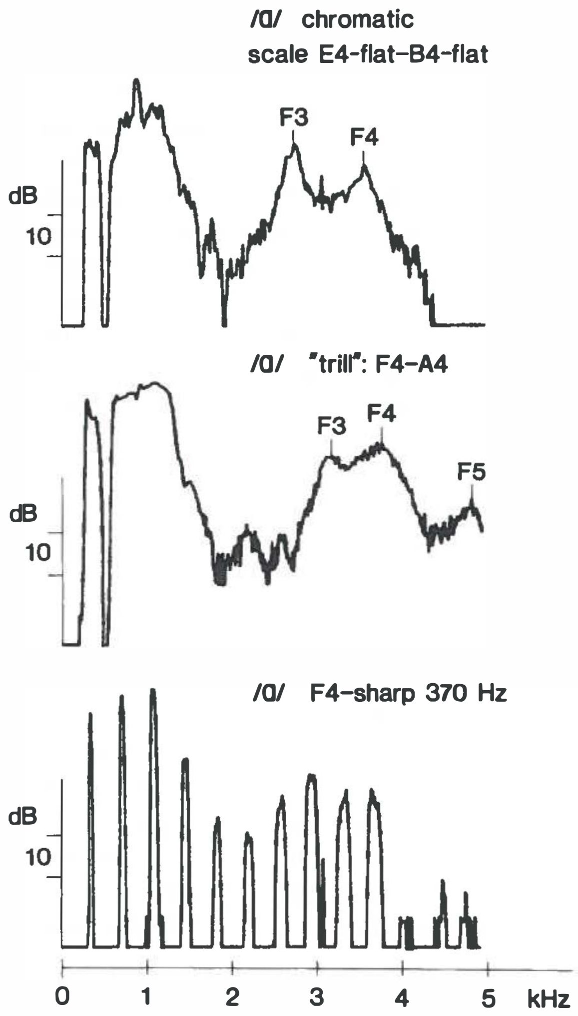


**Figure 5**
Spectrograms of the vowel /ɑ/ sung by a soprano voice: a chromatic scale passage, containing all semitones from *E4-flat* to *B4-flat*; a wide 'trill', with target pitches *F4-A4*; a sustained *F4-sharp*, which lies within the F0 range of the upper two phonations. The formants of interest here, F3-F5, have less consistent locations than in the spectra of the male singer, but they still allow some general observations (see text).

make a comparison.) Since over this range we can expect some adjustment of the lower formants in order to 'tune' the vocal tract to the changing harmonics, we confine our attention here to the configuration of the higher formants in the region of the 'singer's formant'. While there is some discrepancy between the formant frequency values indicated by the two methods (the 'trill' gives somewhat higher values), the spectra, when considered together, give us important information on the configuration of formants that amplify harmonics 7-10 of the single tone. From visual inspection we can reasonably hypothesize that there are two formants in this region (F3 and F4), separated by at least 600 Hz, and that F5 lies above the effective 'singer's formant' region. These observations contrast with the 'singer's formant' of the male voice displayed earlier. His singer's formant also shows typically higher levels, at least on those tones that are sung.

## Discussion and conclusions

We have presented some techniques that we have found helpful in obtaining more precise information about the position of the formants in the singing voice. This information is potentially useful in singing pedagogy for both the articulation of the singer's formant and the careful tuning of the first two formants of the vocal tract for optimal enhancement of the laryngeal source. It should be apparent, however, that the usefulness of any of these, or similar techniques is limited in two ways. First, a certain skill, not identical with the skills of a trained singer, is required for their execution. Second, experienced judgement is required for their interpretation in order to recognize and correct for inevitable error and artifacts. Regarding the first of these limitations, it can be observed that the new skills required (ingressive voice, vocal fry, holding an articulation) are relatively simple when compared with other skills acquired in a singer's training, and the importance of the information gained easily compensates, in our opinion, for the extra effort. Regarding the second limitation, the availability, as control, of alternative methods, and especially the comparison with the sung spectrogram, should not only keep the error within acceptable limits, but also function as feedback, allowing expertise in handling the method to accrue with experience. The spectrum analyzer gives complex information, appropriate to the complexity of the acoustic phenomenon it analyzes, in this case, the singing voice. The inexperienced person should not expect to be able to interpret its pictures exhaustively, anymore than he can interpret X-ray film.

Although the examples we present here touch on some important points in the theory of the singing voice and singing pedagogy (vowel modification for enhanced resonance, second formant tuning, the configuration of the 'singer's formant'), they are merely illustrations of the sort of questions that can be addressed with the spectrum analyzer, rather than the usual 'results' of a scientific investigation. With informal procedures and only two subjects, however highly trained and experienced, we do not seek to prove any of the points our examples indicate. From our background of experience with a range of other singer-subjects, however, we have selected them as characteristic.

A final disclaimer is in order. While in principle it is possible to use the spectrum analyzer for feedback in the training of the singing voice, the cost of the equipment and the required expertise to use it effectively make it unlikely that it will soon have widespread use in vocal pedagogy. The usefulness of such investigation is primarily in what it reveals about the voice production of singers who are taken as models for emulation. The ear of the listener still remains the arbiter of the singing voice.

## Acknowledgements

This work was supported by a grant from the Netherlands Organisation for Scientific Research (NWO) and the assistance of the College of Visual and Performing Arts, Syracuse University, Syracuse New York, USA. The authors appreciate the assistance of Frouke Wildeboer, Franka Steenhuis, and, last but not least, Meindert Goslinga in preparation the manuscript and figures.

## References

Bartholomew, W. A physical definition of 'good voice-quality' in the male voice. *Journal of Acoustic Society America*, 1934;6:25-33.

Coffin, B. *The Sounds of Singing*, Boulder, Colo: Pruett Publishing Co., 1976.

Miller, D.G. and Schutte, H.K.: Formant tuning in a professional baritone. *Journal of Voice*, 1990;4:231-237.

Sundberg, J. Observations on a professional soprano singer. STL-QPSR, 1973;1:14-24.

Vennard, W. *Singing, The Mechanism and the Technic*, New York: Carl Fischer. 1967.

# Chapter 4

# Belting and Pop, Non-Classical Approaches to the Female Middle Voice: Some Preliminary Considerations

# Belting and Pop, Non-Classical Approaches to the Female Middle Voice: Some Preliminary Considerations

Harm K. Schutte and Donald G. Miller

## Abstract

There is a commonly perceived difference in the sound produced in the approximate range *D4 – D5* by female singers in the western opera and concert tradition, on the one hand, and certain other styles, including rock, pop, folk, and some Broadway musicals, on the other. The term 'belting' is sometimes used to refer to at least one approach to such 'non-classical' singing. In this study, based on spectrographic, electroglottographic, and sub- and supraglottal pressure measurements on representative voices of the 'operatic' and 'non-classical' tradition, acoustic and laryngeal differences between the two traditions are described, and an objective, specific definition of 'belting' is offered.

## Introduction

Because singing, one the one hand, is so easily recognized and defined, while on the other hand, its varieties are infinite, the classification of types of singing has always been a problematical subject. One of the most useful designations is that of historical tradition, such as the 'western operatic and concert' tradition which is the subject of much of the scientific and pedagogical literature on the singing voice. There is of course a huge range of variation within this tradition, but there is enough agreement on the type of sound desired and its basic manner of production that we can usually recognize it in a general way, commonly referring to it as 'classical' singing.

The negative formulation of the subject of this investigation is a reflection of the practical questions that confront voice teachers, most of whom have their formal training in this 'classical' tradition. Often enough it happens that aspiring or active singers who do not want to 'sound like opera singers,' but rather to emulate the singing stars of the broader world of popular entertainment consult such teachers. The teacher may accept such pupils, feeling that his basic knowledge of singing technique will make it possible to help develop or improve singing skills of any sort, but remaining aware that a somewhat different aesthetic will rule the selective ear of his pupil.

This familiar aesthetic, or complex of aesthetics, the historical development of which has been sketched by Boardman (1989), we shall refer to as 'non-classical,' reflecting the vantage point of the 'classically' trained voice teacher. For such a teacher, two questions of high priority are: What are the distinguishing characteristics of these sounds and their production? and, Are they compatible with vocal hygiene?

Some of the general features of this style of singing are widely recognized. To begin with, the texts of the songs have a more dominant position in the total effect than do the texts of 'classical' counterparts. The texts are often witty or carry important nuances of emotion, and it is essential that they be understandable, even on first hearing. This means that the singer is permitted less of the vowel modification used by 'classical' singers to enhance the instrumental beauty of the vocal line.

Second, a high value is put on naturalness of sound, often even at the expense of the beauty cultivated in the 'classical' tradition. The increased size of the voice and its extended range often count against it if they make it sound 'too trained.' The dark sound of the 'covered' voice seems particularly offensive to this aesthetic of naturalness. On the other hand, unevenness and certain features peculiar to individual voices that would be considered faults in the 'classical' world are readily tolerated.

This brings us to a third observation, namely, that the performer generally has a higher place vis-à-vis the composition than in the world of 'classical' music: in the 'classical' world it is understood that singers should not attempt to sing a given piece in public unless they can meet the (high) demands it makes on their technical skill; in the 'non-classical' world it is not unusual to adapt the song to the strengths and weaknesses of the individual voice and temperament.

Some of the differences between the 'classical' and 'non-classical' approaches to singing are connected with the fact that amplification of the singer is presupposed in most of the 'non-classical' world, now even in works for the theatre, where 'belting' once had the function of allowing the female singer to be heard in low-middle range over a brassy orchestra (Boardman, 1989).

These general guidelines could be developed in detail, but our purpose here is rather to describe in objective, measurable terms some of the acoustic and physiological features characteristic of this style of singing. Here again, comparison with the more unified and codified 'classical' tradition seems to offer a promising point of departure. The literature in both camps, however, is rather sparse. Using the refined ear and knowledge of the singing teacher, as well as a gift for writing, Osborne (1979) analyzed recordings of the leading singers of the Broadway stage, but he offered no acoustic and physiological measurements. Large (1973) measured paired tones in 'chest' and 'middle' registers in what we call here 'classically' trained singers, but he overlooked the key role of formant tuning, which we describe below.

What follows is not a catalogue of the practices of 'non-classical' singers, let alone a revelation of the 'secrets' that distinguish the outstanding ones, but rather a physiological and acoustic sketch of some of the basic choices these singers make, along with the consequences of such choices. If our observations prove to correspond to the perceptions of voice teachers and singers, perhaps they can serve as a basis for further investigation. In any case, we consider that the success of the study will be greater, the closer we come to identifying the control mechanisms singers use in selecting or rejecting a sound according to its appropriateness for a given style.

## Specialized terms

In the following section of this study, various vocal-technical strategies for singing in the female middle range are described with the aid of physiological and acoustic measurements. Since the specialized terms and instruments used for these measurements may not be familiar to the reader, they are explained briefly below. The curious reader will find it useful to consult a textbook on speech acoustics for more extensive further information.

A simplified model of the voice divides it into source and filter. The source is formed by the effect of the oscillating vocal folds on pressurized air coming from the lungs. In each glottal cycle the vocal folds open and (nearly) close, producing a fundamental frequency (F0) at the repetition rate of the vocal fold oscillations, for example, at 400 Hz (cycles per second). The source also produces the rest of the harmonic series (H2, H3, etc.), which are whole number multiples of the fundamental or first harmonic (H1). A 'vocal fry' source (the low, cracking sound of air escaping in tiny puffs through closed vocal folds, also called 'creaky voice'), by contrast, is non-periodic (has no regular repetition rate).

Much information about the source is provided by the electroglottograph (EGG), an instrument, which measures changes in vocal fold contact area. With it we can not only determine F0 by 'counting' cycles per second, but it is also possible to get an estimate of the portion of the glottal cycle belonging to the closed and open phase. This is important in distinguishing between the basic vocal fold functions of 'chest', with its long (> 50%) closed phase, and 'falsetto', with its short (< 40%) closed phase. (The terms 'chest' and 'falsetto' are shunned by some who wish to avoid historically imprecise associations of the terms; in this study they are used only for this specific distinction in vocal-fold function. The equivalent terms, 'modal' and 'loft' registers have found some currency in the voice science literature.)

The 'filter' function in voice refers to the effect of the resonator on the 'source'. This is comprised (in non-nasalized vowels) of the air space between the vocal folds and the mouth opening. This complex, highly variable space, called the vocal tract, has a series of resonances or formants (designated F1, F2 etc., in ascending order, according to center frequency). The configuration of the vocal tract, and thus of the formants, is varied by changing the positions (articulations) of the lips, jaw, tongue, velum, and larynx, resulting in the various vowels and some of the differences in 'quality' of sung sounds. While a larger (smaller) vocal tract is generally characterised by lower (higher) formant values, typical values for the various vowels in speech have been measured, and these can be compared with the formant values of sung vowels, which in turn imply information about the configuration of the vocal tract.

The instrument used in this study for gathering quantitative information about the vocal tract is the spectrum analyzer. The products of the spectrum analyzer are spectrograms, which give an analysis of the frequency components of a given 'time slice' of sound. (Examples of these are the figures accompanying this study.) All sung sounds are more or less complex, containing not only a fundamental (first harmonic), but several higher harmonics, arranged in a spectrogram along the x-axis according to frequency in thousands of cycles per second (kHz). The y-axis represents the level in decibels (dB) of the several harmonics. The varying decibel levels of the harmonics, which presumably emerge from the glottal 'source' with uniformly diminishing levels as frequency increases, are the result of the effect of vocal tract resonances (formants). The formants of particular interest in this study are F1 and F2, which usually resonate (increase the level of) one of the first three or four harmonics in this female middle range. Generally speaking, the closer (in frequency) a formant is to a harmonic, the higher the decibel level of the harmonic.

In sung tones, sound is present only at the frequencies of the harmonics, resulting in a line spectrum, with gaps between the harmonics. In the range of interest here, the gaps are from 300 to 700 Hz wide, making it difficult to infer the location of the formants. Thus we also use spectra from a non-periodic, vocal-fry source. These contain all frequencies, resulting in a continuous spectrum with sharp peaks, at least at the frequencies of the first two formants. By locating the formants more precisely, we are able to estimate when a singer is adjusting ('tuning') a formant frequency to follow (or 'track') a given harmonic, in order to maintain 'resonance' in the voice, as the sung pitch changes.

The pitch-numbering system employed is one in which the pitches of each octave (*C*-*B*) have the number of the several *C*s of the piano keyboard, starting at the bottom. Thus '*middle C*' (262 Hz) is *C4*.

## Measurements and discussion

The emphasis, in the 'non-classical' music world, on naturalness of sound and the toleration, even encouragement, of personal idiosyncrasy make the choice of what to measure a difficult one. The selection of subjects is to a large extent arbitrary: The well-known stars, at one end of the scale, are out of reach: at the other end, the line between professional and dilettante is indistinct. Furthermore, although the pedagogy for such singers has increasingly drawn the attention of 'serious' institutions for professional training, a comparable literature to that concerning 'classical' pedagogy has not yet appeared. Having taken preliminary acoustic and physiological data on a number of subjects, we chose to make measurements of what appeared to be characteristic points of difference between the basic styles. Several of these are illustrated here in sung phonations of a mezzo-soprano with a firm foot in both 'classical' and 'non-classical' styles. Later we make comparisons with some well-known recorded examples.

### *'Middle' and 'chest' registers*

As a point of departure, let us consider some of the differences between two tones sung in the lower middle range.

Figures 1 and 2 show spectrograms of the vowel /ɑ/ on the pitch F4-sharp, sung by a 'classically' trained soprano with a well-developed mid-range, first in her customary 'classical' manner ('middle' register) and then in 'chest' register, making no attempt to match the other tone. The most striking difference between the tones, with respect to what the singer feels in producing them, is that of vocal fold function. The 'chest' register function has a distinctly longer closed phase of the glottal cycle (>50%), requiring greater vocal effort, than that of the 'middle register' (<40%). This difference, which was confirmed by EGG measurements, is widely recognized in voice science literature as the distinction between 'chest' and 'falsetto' registers (Lecluse, 1977; Schutte and Seidner, 1988). (The 'middle' register does not enjoy equally wide recognition in the scientific literature.)

The differences in vocal tract articulation, which accompany the difference in glottal source, are also striking. In the habitual 'middle' register phonation, the first two formant frequencies are lower, with F1 falling below the second harmonic (H2) and F2 strongly resonating H3 at 1,080 Hz. The pronounced tendency of F1 to follow H2 near the upper limit of 'chest' register, which the authors have described in the male voice (Miller and Schutte, 1993), is also

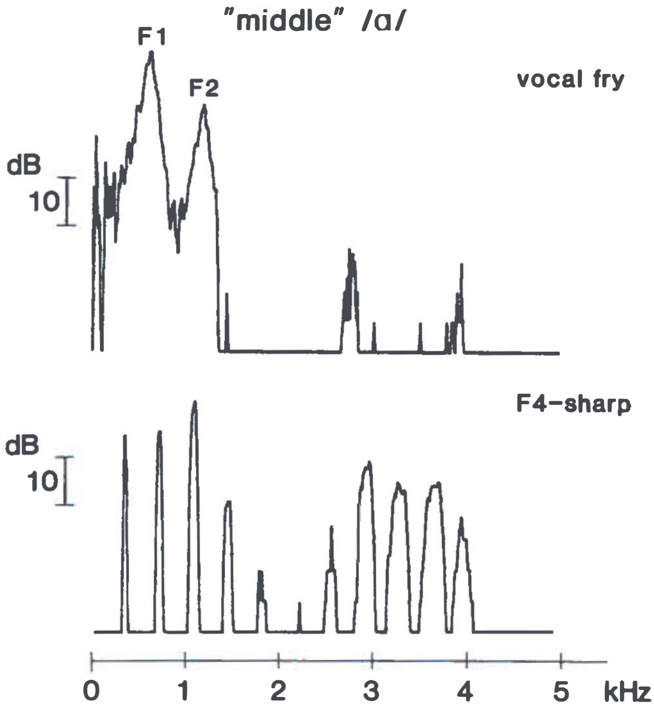


**Figure 1**
Spectrogram of the pitch *F4-sharp*, sung by a 'classically' trained soprano in 'middle' register. The vocal-fry spectrogram, in the same articulation, supports the interpretation that F2 = H3 and F1 falls below H2.

apparent here. F1 rises ca. 100 Hz to keep pace with H2 (note the 21-dB difference between H1 and H2 levels), and the frequency of F2 is evident from its impact on H4.

Our remarks about the frequencies of the first and second formants in these spectrograms, as well as in the following set, are based on visual inspection of the sung spectrogram together with a non-periodic spectrogram, made immediately afterwards, of a vocal-fry source while the vocal tract articulation of the sung phonation is carefully maintained. Although the *precision* of the formant frequencies measured in the non-periodic spectra is subject to error, especially (as was the case here) with subjects inexperienced in this method, these spectra clearly support our description of the varying positions of the first two formants. (For a more complete description of this use of spectrum analysis, see Miller and Schutte, 1990).

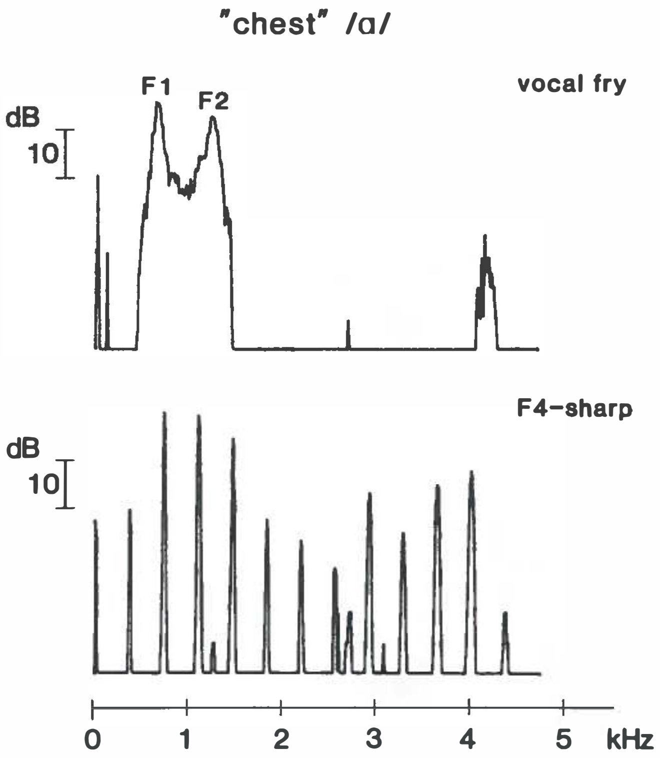


**Figure 2**
Spectrogram of the same pitch, vowel and singer as in Figure 1, here in 'chest' register. The first two formant frequencies are higher, with F1 moving up to resonate H2. The 'chest' register production also shows a reduction in vibrato (the 'thickness' of the harmonics) and an increase in higher-frequency components, features which are perceptually apparent.

The higher formant frequencies of the 'chest' register articulation, characteristic of more 'open' singing, are closer to average speech values (Peterson and Barney, 1952) than are those of the more 'covered' sound of the 'classical' articulation. To what extent this is a direct result of F1-H2 tracking or, alternatively, just a change in 'color' cannot be determined from this data. The two go hand in hand and take us away from the aesthetic of the 'classical' sound toward the 'non-classical.'

### *'Non-Classical' modes*

The next series of phonations (Figures 3-5) were produced by our versatile mezzo-soprano. The first, which we shall designate 'pop,' is from a sustained /ɔ/ (actually the phoneme /ou/, as in 'boat', leaving room for discretion as to

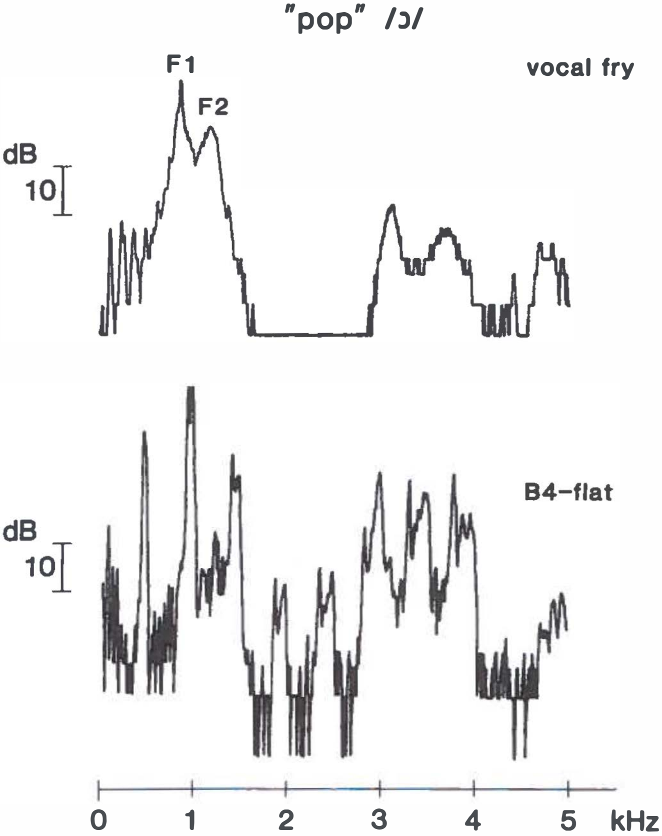


**Figure 3**
Spectrogram of a sustained tone from a popular song, sung in habitual manner by a mezzo-soprano, who has also had 'classical' training. The paired vocal-fry spectrogram shows that the second harmonic is resonated by the approximated first two formants, one above and one below H2.

how open or close to sing the vowel) on the pitch *B4-flat*. This is sung in a phrase from a popular song, which the subject has sung frequently over many years; thus it represents a habitual 'non-classical' mode.

The phonation is repeated in a 'classical' mode as well as in 'chest' register, in the sense of maintaining the long (>50%) closed glottal phase. From visual inspection of the levels of the first three harmonics in these, confirmed by information from the non-periodic spectra, we infer that the dominant H2 in the 'pop' example results from the impact of F1, located not far below the

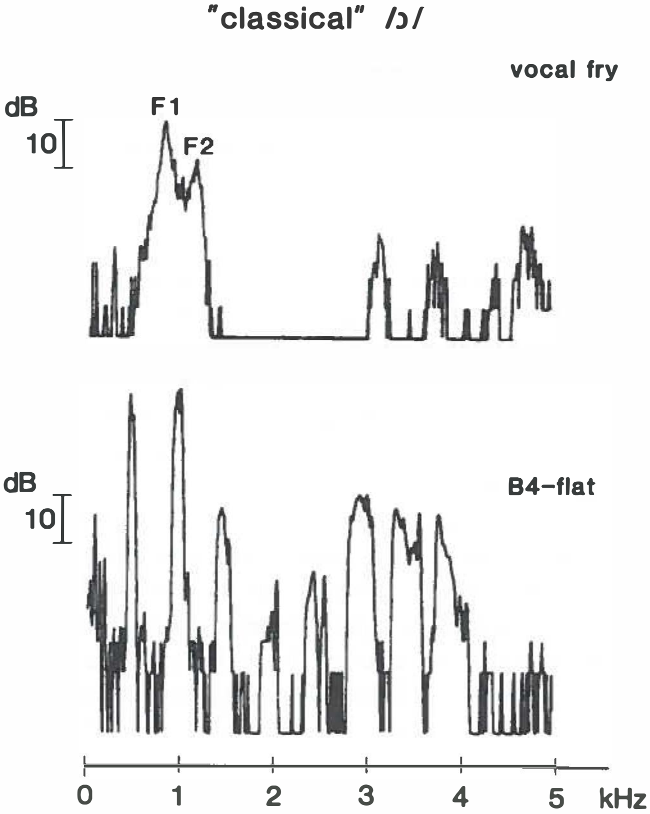


**Figure 4**
Spectrogram of the same tone as Figure 3, this time sung in a 'classical' mode. Both F1 and F2 have moved lower, and now the first harmonic is also prominent. The perceptual result is 'rounder' and 'darker'.

harmonic. In the 'classical' example both F1 and F2 have moved lower, making the first two harmonics nearly equal in level and giving the sound a 'darker' and 'rounder' perceptual quality.

For the 'chest' register example the formants are moved in the opposite direction. The 27-dB difference in level between H1 and H2 is a strong sign that the charateristic tracking of H2 by the first formant in the high part of the 'chest' register is taking place, putting F1 >950 Hz. Further differences between this phonation and the other two are the relative lack of vibrato (especially visible in the excursion of the upper harmonics) and the 'spread' of the high frequency components, which in 'pop' and 'classical' are concentrated in harmonics six through eight, forming a sort of 'singer's formant.' Both of these differences can be attributed to the high larynx position needed to get the first formant above the second har-monic, a point we shall return to later. The perceptual result is a loud sound with a bright, somewhat harsh quality that conveys the excitement of high tension.

***'Legit' and 'belt' voices***

Generalizing from observations on these three different approaches to the same tone, we arrive at a series of statements, as detailed below.

Whereas the 'classical' approach to the middle range is characterized by the relatively low first formants that result from the widely endorsed 'comfortably low' larynx position, the 'non-classical' approach prefers the higher (first) formants of speech. On the open vowels the first formants rise still higher than average speech values in the middle range to keep F1 in the vicinity of the second harmonic. If the vocal-fold function is allowed to relax into a 'falsetto' adjustment, F1 can stay below, but close to, H2, permitting a high but non-extreme larynx position. This is the basis of the so-called 'legit' Broadway voice: a pretty, but nonetheless 'open' sound in the middle range with text articulation seemingly not far removed from that of speech. The vocal-fold function of 'chest' register in the middle range, however, presents a different picture: there is greater effort both at the glottal source, where the long glottal closed phase demands and contains increased lung pressure, and in the vocal tract, where the first formant must be raised all the way to the frequency of the second harmonic. This is the basis of the so-called 'belt' voice, which we define as follows:

> Belting is a manner of loud singing which is characterised by consistent use of 'chest' register (>50% closed phase of glottis) in a range where larynx elevation is necessary to match the first formant with the second harmonic on open (high F1) vowels, that is, approximately *G4-D5* in female voices.

Although these measurements do not include any formal monitoring of larynx position, both informal observation and theoretical considerations lead us to assume that the higher values of F1 on open vowels in 'non-classical' singing are the result primarily of higher larynx positions. The assumption applies *a fortiori* to the extreme F1 values seen in high 'belting.' The theoretical argument supporting this assumption is found in Fant (1970): For open back vowels, F1 is affiliated with the back cavity (between the glottis and the point of constriction of the tongue); a change in the volume or length of this

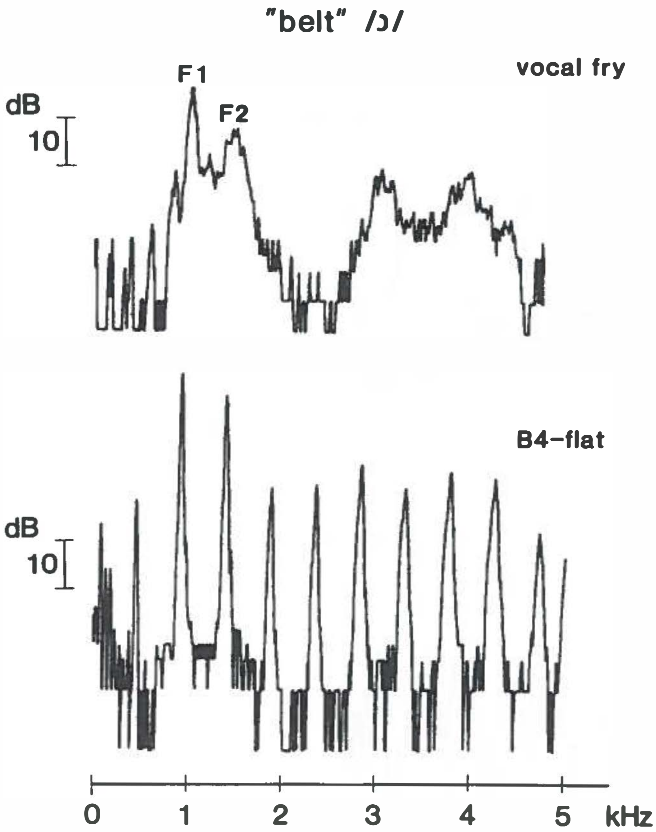


**Figure 5**
Spectrogram of the same tone as Figures 3 and 4, this time in 'chest' register (> 50% closed phase), also called 'belting.' F1 and F2 move higher, with F1 following the second harmonic. Vibrato is diminished and higher-frequency component increased, as in Figure 2. Perceptually this is a loud, bright, 'edgy' sound.

cavity will raise or lower all the formants, but especially F1; therefore, we conclude, a maximal F1 will intuitively be achieved by reducing the back cavity to minimal size.

### *Upper limits*

The upper limit of the female middle voice presents a challenge to all three approaches. By the time she reaches *D5* (592 Hz), a 'classical' singer who keeps her F1 frequencies below, say, 700 Hz, will already have arrived at the point where F1 can resonate the first harmonic. Above this point F1 will follow H1 for another sixth or seventh. The adjustment we have labelled 'pop'

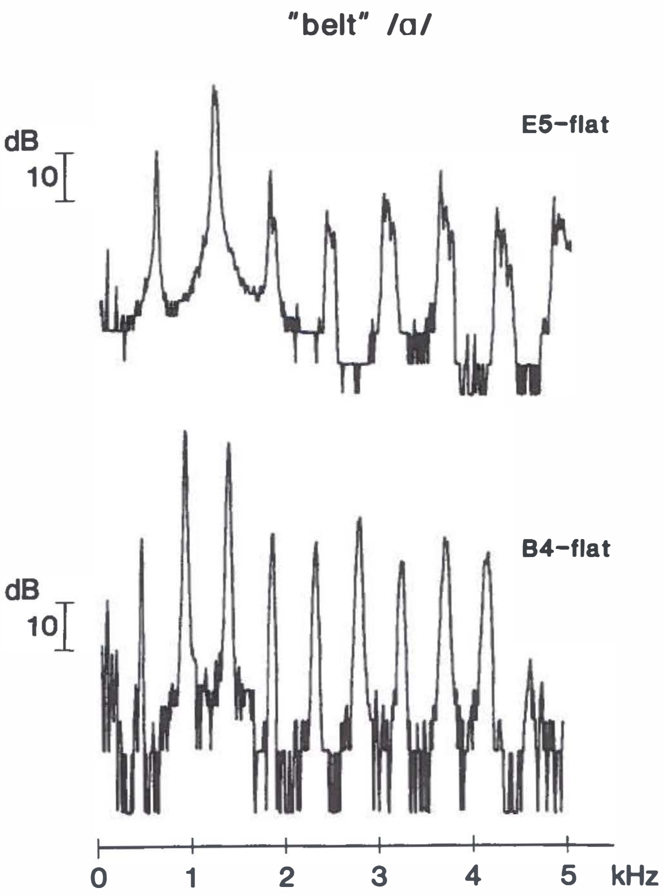


**Figure 6**
Spectra taken from an *E4-flat* to *E5-flat* scale sung in 'chest' register by the same subject as Figures 3-5 (note the similarity of the *B4-flat* with that in Figure 5). The spectrogram of the *B4-flat* has a sharper profile and is less 'spread', a feature especially noticeable on the difference in levels between the first two harmonics. Perceptually the *B4-flat* sounds fuller and less constrained.

(akin to 'legit') will arrive at *D5* with F1 adjusted upward to stay close to H2, thus with an elevated larynx. From this position the move to bring F1 close enough to resonate H1 is a 'long' one, since the distance between the two harmonics is considerable (600 Hz, one octave). The perceptual effect is too abrupt and generally rejected as 'operatic,' even if the change in larynx position can be accomplished skilfully. If a 'classical' pop voice (where the first formant tracks the fundamental) is called for, the practiced 'legit' singer will avoid getting caught with an over-high larynx in the middle range.

The most dramatic encounter with the upper end of the middle range is found in the 'belt' voice. Part of the excitement of extending this technique upwards is the clear sense that one is on a collision course with a finite termination of the range. The listener is aware of the risk the singer is taking. We had our mezzo-soprano subject sing the octave scale *E4-flat* to *E5-flat* in 'chest' register. Spectrograms of the *B4-flat* and *E5-flat* are shown in Figure 6.

Even with the high level of vocal effort in such voice production there is perceptually a certain freedom in the first fifth, through the *B4-flat*, which seems to diminish as she sings higher. If, as we assume, F1 follows H2 all the way to 1240 Hz on the *E5-flat*, then there appears to be some loss in the sharpness (i.e. greater bandwidth) of the resonance on the high note where the differences in levels of H1 and H2 is 15, instead of 25 dB. These last remarks are speculative, but they are consistent with the observation from practice that 'belters' (for example, those in the musical *One Mo' Time*) are able to use *B4-flat* as a sort of 'reciting tone,' apparently without great fatigue.

Carefully distinguishing between the contribution of the glottal source and the vocal tract (the configuration of formants) to the final sound gives the 'classically' trained teacher an approach to 'non-classical' singing in terms of more familiar elements. For the vocal techniques we have considered here, these elements are indicated in Table 1. The 'belt' voice takes the 'chest' register (source) to higher, sometimes much higher pitches than is generally advised in 'classical' singing (Garcia, 1982), but because of the high-to-extremely-high larynx position (vocal tract) the sound is quite different from the 'classical chest' voice. The voice production we have labelled 'pop' here, which uses not a 'chest' register source, but a laryngeal adjustment with a shorter closed phase, much like the 'classical middle' voice, still tends to use the higher larynx positions of speech, bringing the advantages and problems we have touched on here. The ability to move smoothly between the 'belt' and 'pop' voices in the upper middle range demands a high degree of skill, not to say wizardry. (See Osborne's description of what he calls 'belt-mix' voice, 1979). A continuous adjustment of the glottal source between 'belt' and 'pop' is theoretically possible, but hard to find in practice.

This concludes our brief survey of some techniques used by female singers, called forth by the original question about the distinguishing characteristics of the non-classical sound and its production. With the help of objective measurements we have outlined two strategies, 'pop' and 'belting', for dealing

**Table 1**
Physiological and acoustical parameters characterize four types of female voice production.

| | vocal fold adjustment | larynx position* | subglottal pressure | frequency range |
|---|---|---|---|---|
| pop | 'falsetto' | intermediate to high | moderate | low to middle |
| legit | 'falsetto' | low to intermediate | moderate | middle to high |
| belt | 'chest' | high to very high | high | middle |
| classical chest | 'chest' | low to intermediate | moderate | low |

Note *. This parameter could be more directly designated 'first formant frequency,' of which larynx position is merely the most determinative factor on open vowels. Other factors (mouth opening, pharynx width) are also operative in determining F1.

ing with the female middle range, indicating concrete differences from the 'classical' approach. Although the measurements presented support our explanation of the strategies, any assertion about the *extent* to which these strategies are followed by non-classical singers would require a study with a larger number of subjects. In such a study, the interpretation of spectrograms introduced here could reveal much pertinent information. It can, for example, be applied to recorded singers, as we have done in what follows.

Barbra Streisand is a singer who is well known for her versatility of technique in producing a range of expressive effects, largely in the non-classical tradition. Figure 7 consists of spectrograms taken from the end of her recording of 'Being Alive': the held final note (*D5* /ɑ-/), as well as the short higher note (*E5* /ɔ/ 'force'). These notes fit our definition of 'belting', exhibiting the typical H2 dominance (in both cases more than 20 dB above H1), and implying that she raises her first formants to about 1175 Hz (sustained) and 1,300 Hz (short). The pattern of the spectrograms, together with the strong perceptual continuity of the highest notes with typically 'belted' notes of the middle range, are persuasive evidence, even in the absence of EGG confirmation, that she uses the laryngeal adjustments of 'chest' register on these notes.

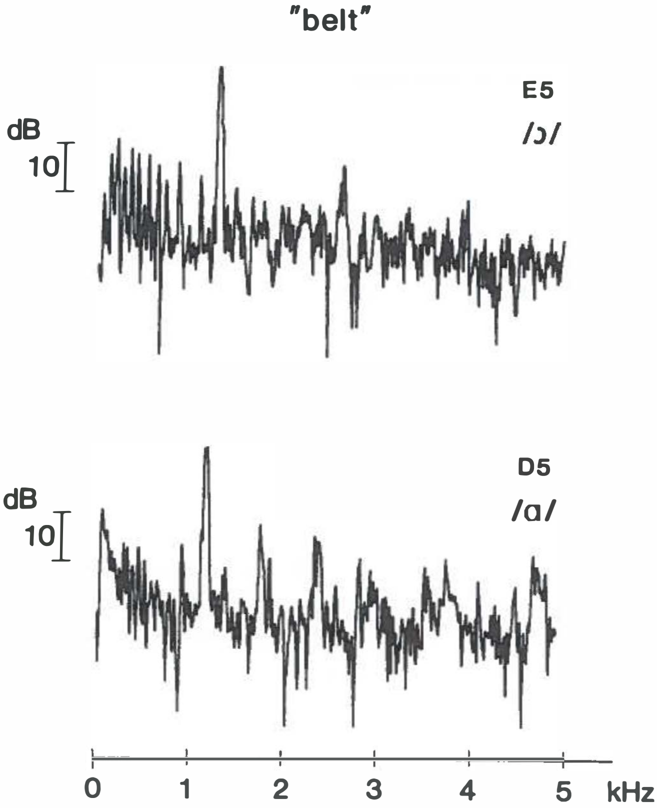


**Figure 7**
Spectrograms made from 'belted' high notes on a recording of 'Being Alive,' sung by Barbra Streisand: the short E5, 'force,' and the long last *D5*, 'alive'. In both cases the second harmonics dominate other components of the voice and the orchestral.

## Vocal health

We now turn to the question of the effect on vocal health of these 'non-classical' modes of singing in general and 'belting' in particular. Is there a medical-physiological basis for the unqualified rejection of continued use of the 'chest' register in the range *G4-D5*, such as voice teachers from Garcia (1982) to Osborne (1979) have stated? We identify three potential sources of vocal abuse in the techniques described here: high (as opposed to 'comfortably low') larynx, 'chest' (closed phase >50%, as opposed to 'falsetto' closed phase <40%) register, and high breath pressure. That none of these factors need in itself lead to vocal abuse is evident from the fact that skilled operatic tenors sing with high pressure and 'chest' register (long closed phase) glottal source, while Chinese singers of 'Peking opera' evidently sing with high larynges and high pressure. 'Belters' who use the larynx positions required for extreme high notes, however, seem to expose their voices to a unique degree of risk by challenging all three risk factors simultaneously. That there are robust voices that stand up under the rigors of 'belting' without immediately developing acute problems is an undeniable fact. Even with these voices, however, well-coordinated, firmly closing vocal folds and avoidance of such extreme larynx positions that all freedom of movement is lost would seem in order, along with respect for the limits imposed by fatigue.

'Non-belters' who use a variety of the 'pop' approach we have described here are better off with respect to all three of these potential sources of abuse: they use less pressure, a 'falsetto' glottal source, and need less extreme larynx positions than 'belters' do to sing the same pitch. It should be apparent from our descriptions of 'pop' and 'belt,' however, that the higher larynx position is often part of the aesthetic, reducing the voice teacher's discretion in altering it. On the other hand, 'old style' musical roles which require unforced notes as high as *F5* (700 Hz) and higher, will allow the sound of the lower first formant (thus lower larynx) in the middle range as well.

Even if their hands are tied with respect to larynx positioning and related techniques for 'darkening' the voice, there is often much teachers can do to improve the coordination of breath with vocal fold action in these singers, who sometimes have little awareness of how they are using their voices. In a study of voice ailments among 'belters' (presumably our 'non-classical' singers), Lawrence found that voice training correlated positively with vocal health (1979). Whether training leads to less abusive techniques of singing or merely to heightened consciousness of the vulnerability of the vocal instrument, it appears that the trained singer is less likely to succumb to the potential abuses of 'nonclassical' singing.

## Acknowledgements

This work was supported (partially) by a grant from the Netherlands Organization for Scientific Research (NWO). The authors appreciate the assistance of Ella Blokzijl, and especially Meindert Goslinga for his meticulous preparation of the figures.

## References

Boardman S.D. Singing Styles on Broadway. *NATS Journal* 1989;4(1):4-10,14,24.

Fant C.G.M. *Acoustic Theory of Speech Production.* The Hague/Paris: Mouton & Co., 1970:63-90.

Garcia M. *Hints on Singing.* Translated by Beate Garcia. First Printing:1894, New York: Joseph Patelson Music House, Ltd., 1982:15.

Large J.W. Acoustic Study of Register Equalization in Singing. *Folia Phoniatrica* 1973;25:39-61.

Lawrence V.L. Laryngological Observations on Belting. *Journal of Research in Singing* 1979;2(1):26-8.

Lecluse F.L.E. *Electroglottografie.* Thesis Rotterdam. Elinkwijk, 1977.

Miller D.G. and Schutte H.K. Feedback from Spectrum Analysis Applied to the Singing Voice. *Journal of Voice* 1990;4(4):329-34.

Miller D.G. and Schutte H.K. Towards a Definition of Male 'Head' Register, *Passaggio*, and 'Cover' in Western Operatic Singing. *Folia Phoniatrica Logopaedica* 1994;46:157-170.

Osborne C.L. The Broadway Voice, Part 1: Just Singin' in the Pain. *High Fidelity Magazine* 1979;29(1):57-65. Part 2: Just Singin' in the Pain. *High Fidelity Magazine* 1979;29(2):53-6.

Peterson G.E. and Barney H.L. Control Methods Used in a Study of the Vowels. *Journal of the Acoustical Society of America* 19-52;24:175-84.

Schutte H.K. and Seidner W.W. Registerabhängige Differenzierung von Elektroglottogrammen. *Sprache-Stimme-Gehör* 1988;12(2):59-62.

# Chapter 5

# The effect of F0/F1 coincidence in soprano high notes on pressure at the glottis

# The effect of F0/F1 coincidence in soprano high notes on pressure at the glottis

H.K. Schutte and D.G. Miller

## Abstract

Using high-frequency miniature pressure transducers, direct measurements of sub- and supra-glottal pressure are taken on two sopranos singing high notes, in a range where F1 is close to F0. When the peak in supraglottal pressure occurs well into the open phase of the glottal cycle, the instantaneous transglottal pressure approaches zero. When, in one of the subjects, this peak occurs in a later phase as a feature of the vibrato modulation, the transglottal pressure even drops appreciably below zero. The implications of these findings for the glottal volume velocity waveform, as well as for the singer's experience of the second 'passaggio' point, are discussed.

## Introduction

Although it has been known since the time the laryngoscope made the voice source visible that glottal closure is weaker in falsetto than in modal voice, it remains difficult to explain how the high female singing voice, with its falsetto-like vibration pattern, can sustain relatively long phrases, instead of running out of breath prematurely, as one might predict from its relatively short closed quotient. The shortening of the glottis, or 'damping' of the vocal folds by holding the cartilagenous portion closed (Vennard, 1967), as well as the relatively small lateral excursion of the vocal folds, when compared with modal voice, may reduce the flow of air through the glottis in the falsetto pattern. Nonetheless, even a glottis of reduced area would hardly seem capable of retaining subglottal air under considerable pressure when the closed quotient is small or indeed, when complete closure is not established at any point in the cycle.

From their collective experience in singing, practicing phonatrics, teaching voice and investigating the physiology of singing, the authors speculate that average subglottal pressure, as well as maximal phrase duration, are, on the average, smaller in female than in male concert and opera singers, an hypothesis that might be worth testing (Schutte [1980] has, however, found a lower mean airflow rate in women's voices than in men's, when investigating sustained phonations by normal non-singer subjects). Nonetheless, sopranos and even countertenors sing longer phrases than one would expect from considering a comparable breathy male falsetto. Furthermore it is the subjective experience of many female singers of high notes that a well-produced tone offers an increased resistance to breath pressure, seeming to reduce the airflow. Taking as our point of departure Sundberg's observation that the soprano greatly increases the intensity of her voice by tuning the first formant close to the fundamental frequency on high notes (Sundberg, 1973), we decided to measure the effect of this tuning on the instantaneous transglottal pressure in the expectation that it might be factor in reducing airflow.

## Procedure

Our procedure, described in detail in Miller & Schutte (1985), consists briefly of recording subglottal and supraglottal pressures with miniature high frequency pressure transducers on a small catheter that passes through the glottis. Also recorded are the radiated sound and the EGG waveform. The

SPL is derived automatically from the audio signal, and the transglottal pressure wave, not reported in the previous paper, is derived by a simple electronic analogue subtracting network as the difference between the supraglottal and subglottal pressures.

Our previous article was limited to a single subject, one of the authors, and to singing in modal voice, while dealing only in passing with the patterns of high notes in that voice. A more detailed description of falsetto and of transition to the upper range will appear later.

In the present study we shall focus on a kind of acoustic back pressure, which takes the form of a reduction in the pressure differential across the glottis (transglottal pressure) during the phase of high positive supraglottal pressure. We have selected phonations from two dissimilar soprano voices, both singing high notes, in order to illustrate this effect. Subject A is a trained amateur singer who produces a sound relatively weak in high partials. Subject B is a professional opera singer with a bright, strong high range, particularly from *G5* upwards. Both are in their early thirties.

## Results

### *Ratio of modulations of supra- and sub-glottal signal*

Our figures (Figs. 1-3) show the details of interest as well as the whole phonations in which they occur. On comparing the curves in Figs. 1 and 3 with the figures of the male modal voice (Miller and Schutte, 1985) one is struck by ratio of $p_{supra}$ AC to $p_{sub}$ AC, attaining values of as high as 7:1, considerably higher than anything we measured in modal voice. This results not only from the enhanced amplitude of the $p_{supra}$ signal by near coincidence of F0 with F1, but also from a reduced subglottal AC, since the F0 exceeds the fixed subglottal resonance (about 640 Hz in these subjects), diminishing the acoustic response that we find in the subglottal space at lower frequencies. Instead of playing an active, sometimes dominant, role in regulating vocal fold motion, as they do in parts of the modal voice, subglottal modulations are restricted, revealing a weakened acoustic subglottal response even during the well-defined closed phase in subject B. The asymmetrical shape of the $p_{supra}$ waveform, revealing a mean positive supraglottal pressure in the order of 5% of $p_{supra}$ AC during vowel phonation, is also apparent, particularly in the case of subject B.

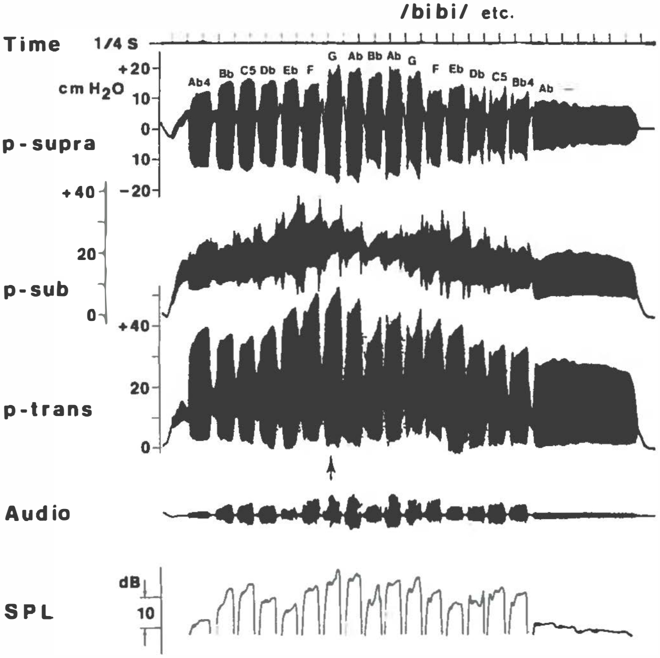


**Figure 1**
Nine-note scale on the repeated syllable /bi/, sung upwards from and returning to *A4-flat*, by subject A, a trained amateur singer. Note the sudden increase in $p_{supra}$ AC, concurrent with a reduction in $p_{sub}$ AC, in the step *F5 - G5*, a correct, if somewhat crudely executed '*secondo passaggio*'. The arrow indicates the note from which the detail in Figure 2 is taken.

### *Small peak in $p_{sub}$ during the open phase*

The feature of the waveforms depicted here that interests us most in this study, however, is the small maximum in $p_{sub}$ that appears during the open phase in the glottal cycle (arrows in Figs. 2 and 4). We find no explanation for this in the acoustics of the subglottal tract, nor is there any reason to suppose that it is caused by a secondary closing of the vocal folds. It corresponds rather to the point where the positive supraglottal pressure, coupled with the subglottal space in the open phase, approaches the value of the subglottal pressure, preventing its further decline and actually raising it a small amount. (The 6 cm distance between the transducers, representing 0.2 ms at the speed of sound, accounts for the apparent lag of this effect in the subglottal wave after the maximum in the supraglottal wave.)

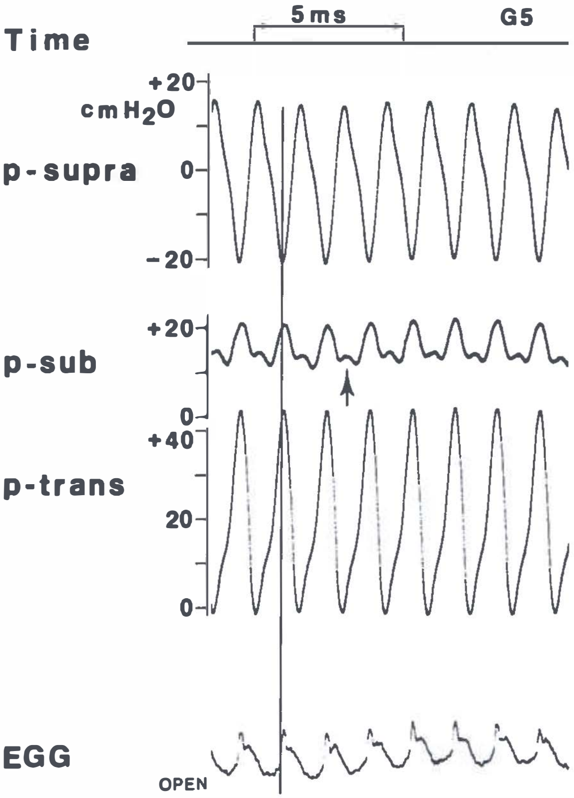


**Figure 2**
Portion of a sustained /i/ at 790 Hz. The vertical lines here and in Figure 4 are drawn through representative closing points in the glottal cycle. The values corresponding to the point marked with the arrow are 15, 14 and -1 cm water pressure for supra-, sub-, and trans-glottal pressures, respectively. This small maximum point in $p_{sub}$ lags 0.2 ms behind the $p_{supra}$ maximum it reflects because of the 6 cm that separate the transducers. Note the weak EGG signal, probably without full closure.

## *Coupling of sub- and supraglottal spaces*

A closer look at the curves reveals several other points of interest. In subject A, the amateur singer with a voice lacking in brilliance, the EGG waveform indicates a large open quotient; in fact it is doubtful whether the narrow 'finger' at the closing moment results from a complete closure (Fig. 2). In the middle of the open phase, where the supraglottal pressure maximum is reflected in the sub-glottal pressure wave, indicating considerable coupling between the supra- and sub-glottal systems, there is a segment of about one-

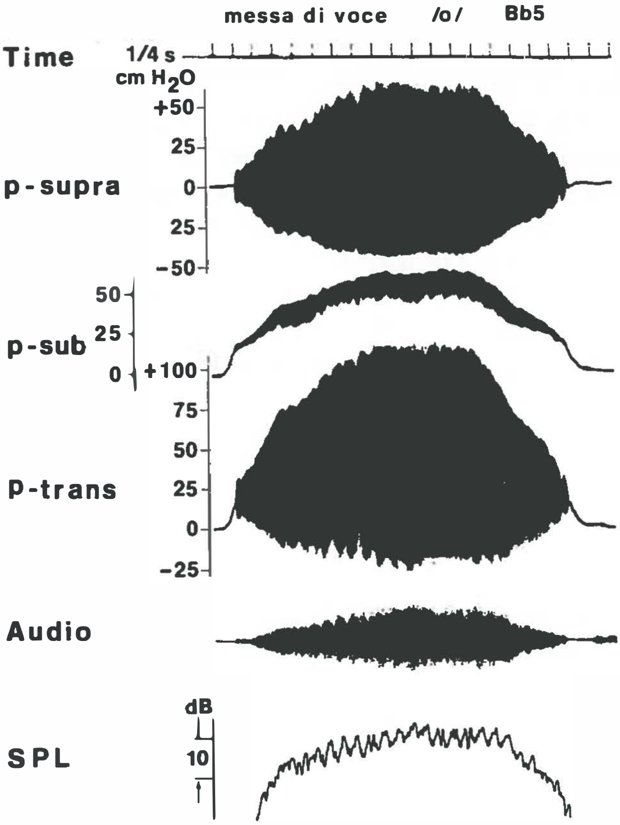


**Figure 3**
A maneuver sung by subject B, a professional opera singer, called *messa di voce*, where the singer begins and ends softly, getting loud in the middle, on a single pitch (in this case *B5-flat*). Note the high values of the pressures, as well as the particularly high ratio of $p_{supra}$ AC to $p_{sub}$ AC. An intensity vibrato, clearly shown in the SPL curve, is also evident in the other signals, particularly at minimum $p_{trans}$ and maximum $p_{supra}$.

fifth of the glottal cycle where the transglottal pressure is quite low, below 5 cm of water pressure. In the data of the stronger singer we not only find considerably higher pressures (note the difference in scale in Figs. 3 and 4), but also an EGG waveform that indicates a closed quotient of around one-third. Nonetheless, when the maximum in the $p_{supra}$ wave occurs early enough in the glottal cycle, just after half-way between points of closure, it is reflected in the $p_{sub}$ wave in spite of the shorter open phase, at which point the transglottal pressure approaches zero.

### Effects of changes due to vibrato

In this voice, however, the vibrato pattern is characterized by a modulation in the phase of $p_{supra}$ with respect to the closing moment. When $p_{supra}$ max occurs around 180° of the cycle (Fig. 4, b), we find the reflection in $p_{sub}$ described above. When it moves to around 270° (Fig. 4, a), however, the closing vocal folds have already reduced the coupling of the spaces above and below the glottis, and rise in $p_{supra}$ is not limited to the instantaneous $p_{sub}$. The result is a higher $p_{supra}$ maximum (the minimum is affected less) and a $p_{trans}$ minimum that drops further below zero, in this case to –10 cm of water pressure. This greater excursion of $p_{supra}$ AC produces an SPL that is 4-5 dB greater than that of the vibrato phase where the $p_{supra}$ maximum occurs earlier in the glottal cycle.

## Discussion

We have selected the phonations presented here because the waveforms present a clear demonstration of moments when the peak supraglottal pressure rises to the value of the instantaneous subglottal pressure, reducing the pressure drop across the glottis to zero, a phenomenon which to our knowledge has not been previously described or anticipated by models. Our further measurements indicate that with tones sung in female ‘head voice’, as well as in male falsetto, the minimum transglottal pressure will commonly approach zero, provided that the fundamental frequency is high enough to be resonated by the first formant.

If the peak in $p_{supra}$ occurs at a moment when the glottis is relatively wide and the coupling of sub- and supra-glottal spaces is high, the $p_{trans}$ does not drop significantly below zero. This means that the maximum SPL of the tone, generally determined by the F0/F1 component in such phonations (Sundberg, 1973) will be a function of the subglottal pressure, since the $p_{supra}$ AC cannot go beyond twice the value of the $p_{sub}$ at the moment of peak $p_{supra}$. If, on the other hand, the peak in $p_{supra}$ occurs when the closing vocal folds have effectively separated the pressures above and below the glottis, then the peak $p_{supra}$ is free to rise beyond $p_{sub}$, creating negative transglottal pressures (in our measurements –10 cm water pressure and beyond) and raising the SPL accordingly. It is worth mentioning in passing that the above implies that the coupling of the sub- and supra-glottal spaces appears to depend on the degree of glottal opening, rather than simply its presence or absence.

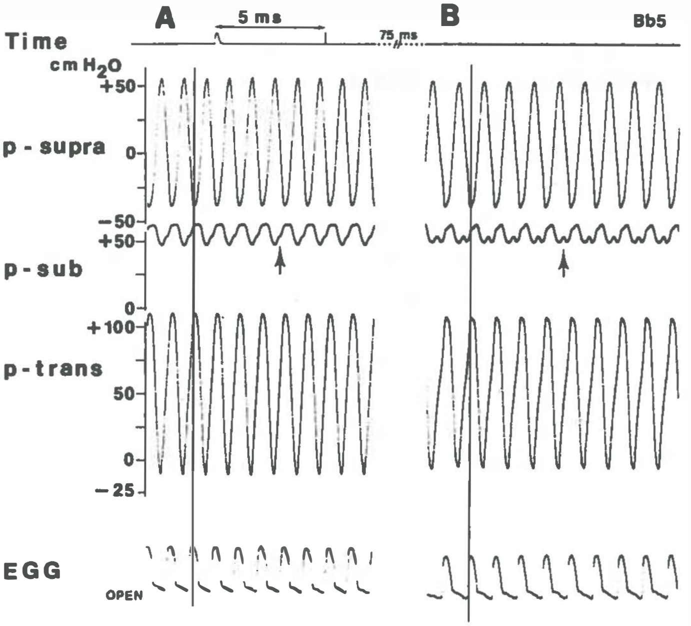


**Figure 4**
Two details from Figure 3 with 75 ms, or one-half a vibrato cycle, between them. Part A, with an F0 of 960 Hz, occurs at a maximum in frequency and SPL of the vibrato cycle, while part B, with F0 of 880 Hz, represents a minimum. The arrow in B indicates a small maximum in $p_{sub}$ reflecting the $p_{supra}$ maximum. Values corresponding to that point are 53, 52 and -2 cm of water pressure for supra-, sub-, and trans-glottal pressures, respectively. The arrow in A indicates the comparable spot, but it occurs later in the glottal cycle, at which point, coupling of sub and supra systems is less complete, allowing the transglottal pressure to drop to -10 cm water pressure. Note the strong EGG signal, revealing a closed quotient of about one-third.

Since it was not measured in this experiment, we can only speculate on what effect the period of reduced $p_{trans}$ will have on airflow. Whether the flow stops is uncertain, but the duration of the period of reduced $p_{trans}$ in the fully open phase in the case of the breathy singer ($p_{trans}$ below 10 cm of water pressure in 40% of the cycle, and below 5 cm of water pressure in 20%), suggests an important effect, in which one would expect the glottal volume velocity waveform to have a large depression somewhere near the center of open phase, the point where the highest flow would be predicted by the linear model based on the glottal area function. It would appear to have an effect in the stronger voice as well, even where the $p_{supra}$ peak comes fairly late in the open phase. Extracting glottal volume velocity waveforms of such phonations is problematic for inverse filtering (Fant, 1982), but a knowledge of the pressure patterns should help in recognizing a correct waveform, such as the one reported recently by Rothenberg (1985).

With regard to singing technique, it is probable that the back pressure phenomenon described here is one of two important components in the 'registration event' (Miller, 1977) that occurs around 700-800 Hz (*F5-G5*) in the female voice, which the Italians call *secondo passaggio*. (The other important component is an adjustment in the intrinsic laryngeal musculature.) The sudden rise in $p_{supra}$ AC characterizing this registration can even be seen in Fig. 1, where the singer moves from *F5* to *G5* and back again. At this point the singer can tune the first formants of not only the close vowels (with the low first formants), but also the open vowels to the fundamental frequency, producing both acoustic enhancement and also a reduction in airflow, when compared with a poorly tuned vowel. Proper tuning along with a correct muscular adjustment in the vocal folds (which may be the more difficult feature of the 'registration event') greatly increases the efficiency of the vocal instrument and contributes to a subjective impression of ease in the upper range, a much desired condition in the trained soprano voice.

The connection of the varying phase of supra- and sub-glottal pressures within the vibrato cycle has important implications for our understanding of vibrato. Subject B in this phonation produced a vibrato of 6 Hz with amplitude modulations of 4-5 dB and frequency modulations of nearly a whole tone. The peaks of amplitude and frequency occurred in phase at a point in the vibrato cycle where the peak supraglottal pressure was relatively late in the glottal cycle (Fig. 4, a), while lower amplitude and frequency were concurrent with the early $p_{supra}$ maximum, the one reflected by the small maximum in $p_{sub}$ (Fig. 4, b). One is hesitant to ascribe cause and effect in a phenomenon as complex as vibrato, but the modulation in pressure patterns at the glottis seems to have precedence over some other possible sources of amplitude vibrato. The peak subglottal pressure is quite steady (Fig. 3), ruling out respiratory pressure modulations as a source. Moving in and out of a formant in the supraglottal vocal tract is a possibility, but in that case one would expect the $p_{supra}$ AC to modulate in both directions, rather than principally on the positive side, a movement we have explained as a function of the phase shift in pressures. Considering one male subject singing in modal voice, Rothenberg, Miller and Molitor (1986) have observed in vibrato a modulation of the airflow approximately 180° out of phase with the frequency modulations. The aspect of vibrato we are considering here appears to be a similar phenomenon in that important modulations in the vibratory pattern of the vocal folds cannot be explained simply as consequences of modulations in fundamental frequency or respiratory breath pressure, nor by the acoustics of the supraglottal vocal tract.

Although Rothenberg (1985), at least partially in response to these data, has developed a theoretical explanation of the dip that our pressure measurements would predict in the glottal volume velocity waveform, it is noteworthy that other models had not anticipated it, in spite of the fact that it appears to be a common feature of high-pitched singing in falsetto and female 'head voice'. This can be explained in part by the reliance of the models on the data from inverse filtering, which in turn gives convincing data primarily in (low-frequency) modal voice. Other aerodynamic data from the glottis have been difficult to obtain. While our method does not allow completely unhindered phonation, it does provide some clear and reproducible data that present a more balanced view of various modes of phonation, suggesting revision, or at least extension, of voice models, particularly if they are to come to terms with female voice and the phase-related modulation within the vibrato cycle.

## Conclusions

On high notes sung in female 'head voice', where F0 approaches F1, the strength of the acoustic wave in the supraglottal vocal tract often reaches the point where the $p_{supra}$ maximum approaches, and even surpasses, the instantaneous $p_{sub}$. We presume, without offering experimental evidence, that this radical reduction of transglottal pressure has a marked effect on the glottal volume velocity waveform, and reduces the average airflow. The position of the $p_{supra}$ maximum in the glottal cycle varies with the vibrato, at least in the case of our professional singer. When the maximum in $p_{supra}$ occurs in a phase of strong coupling with the subglottal space, this maximum will be limited in the amount of the instantaneous $p_{sub}$, limiting in turn the SPL of the radiated sound. If the $p_{supra}$ maximum occurs later, when the closing vocal folds are uncoupling the system, the $p_{supra}$ can rise higher, sending the $p_{trans}$ below zero. Finally, it is suggested that the back-pressure created by the peak in supraglottal pressure during the open phase is an important element in the subjective impression of ease that one gets from proper female singing beyond the second *passaggio* point.

## References

Fant, G. The voice source – acoustic modeling. In *STL-QPSR* 4/1982. Stockholm: Department of Speech Communication and Music Acoustics, Royal Institute of Technology, 1982;28-48.

Miller, D.G. & Schutte, H.K. Characteristic patterns of sub- and supra-glottal pressure variations within the glottal cycle. In: *Transcripts XIIlth Symposium: Care of the Professional Voice.* New York, The Voice Foundation, 1985;70-75.

Miller, R. *English, French, German and Italian Techniques of singing: a study in national tonal preferences and how they relate to functional efficiency.* Metuchen, N.J. Scarecrow Press, 1977;134-135.

Rothenberg, M. *Cosi Fan Tutte*, and What It Means or Nonlinear Source-Tract Acoustic Interaction in the Soprano Voice and Some Implications for the Definition of Vocal Efficiency. In: Harris,S. Baer, T. &Sasaki,C. (Editors). *Laryngeal Function in Phonation and Respiration.* San Diego: A College-Hill Press, Little Brown and Company. 1987;254-269.

Rothenberg, M., Miller, D.G. & Molitor, R. Aerodynamic investigation of sources of vibrato. In *Folia Phoniatrica.* 1986;40:244-260.

Schutte, H.K. *The Efficiency of Voice Production.* Thesis University of Groningen, 1980.

Sundberg, J. Observations on a professional soprano singer. In STL-QPSR 1/1973. Department of Speech Communication and Music Acoustics, Royal Institute of Technology, Stockholm, 1973;14-24.

Vennard, W. *Singing, the Mechanism and the Technique.* Carl Fisher, New York, 1967.

# Chapter 6

# Physical Definition of the ‘Flageolet Register’

# Physical Definition of the ‘Flageolet Register’

Donald G. Miller and Harm K. Schutte

## Summary

The highest ‘register’ of the female singing voice, often called the ‘flageolet register’ (also called ‘flute register’, ‘bell register’, etc., as well as the misleading term ‘whistle register’) is broadly recognized by voice pedagogues, but not generally defined in terms that are adequate for objective description. This article presents a description of characteristic patterns of vocal fold movement and of vocal tract formants that are specific for the register. Measurements are made by electroglottograph, pharyngeally placed wide-band pressure transducers, and an external microphone in professional soprano subjects who are adept in using this register.

## Introduction

In the literature of singing pedagogy the highest segment of the female singing range is often designated as a separate register, distinct from the more frequently sung series of pitches just above the staff (Miller, 1986). In such treatises the concept of register is seldom defined unambiguously, while the varied names given to this highest segment suggest equally varied mechanisms for its production ('bell', 'flute', etc.), and authors do not always agree on its more concrete specifications (for example, whether breath pressure and flow are higher or lower than normal). The infrequent use of this 'register' in the standard repertory, probably reflecting the fact that only a minority of sopranos have the capacity to employ it effectively, helps to keep it from the scrutiny of researchers and some writers of pedagogical treatises, who are often male and thus do not experience it directly. Vagueness and inconsistency notwithstanding, however, the widespread and long-standing recognition of this highest register by singers and voice teachers is evidence of specific perceptual characteristics, for which there may be identifiable underlying physical characteristics.

The practical question of this investigation is not whether the bell/flute/whistle/flageolet register 'exists', but rather what changes occur in voice production in moving between *G5* (788 Hz) and *D6* (1,175 Hz), two pitches which can serve as representatives of their respective registers: *G5* lies well within the series of pitches where the first formant is adjusted to the fundamental frequency, and any major change in production, as the soprano sings above that point, we assume will have occurred by *D6*. Vocal pedagogues will further want to know whether the model we present for this transition is the 'correct' one, or if there is more than one way of executing it, at least an acceptable one. Such normative questions are generally answered 'by ear', that is, by the cumulative preference of listeners in interaction with the sounds offered by singers. Regarding the question of the appropriateness of designating the highest segment of the voice range a separate register, let us declare at the outset that our use of the term register is not restricted to distinct (modes) of the voice source, but can also include certain acoustical phenomena originating in the vocal tract. In this article we present some preliminary conclusions based on physiologic and acoustic data with the intention to stimulate further research on the topic under investigation.

Before proceeding with our own investigation of these questions, it is well to consider the only experimental study of this register that we could find in the literature. Walker (1988) compared pairs of tones on all attainable semitones between *B5* (988 Hz) and *G6* (1,568 Hz) with respect to three parameters; spectral characteristics, air flow, and the ability of auditors to discriminate between similar and dissimilar pairs. The pairs of tones were sung by seven sopranos, all professional singers or advanced graduate-level voice students, who believed they could produce (some of) these pitches in a distinct 'whistle/flute/bell register' as well as in 'upper/head/loft/falsetto', that is, the same register that would be used for pitches an octave below those of the paired tones. Thus the register distinction (in the rest of the article lumped as whistle/head) was determined by the subjects. The study showed that auditors, both expert and naive, were able to hear the distinction intended by the singers, even with sound pressure level (SPL) normalized to a constant level for the tone pairs. The principal general finding with respect to spectral characteristics was that head register tones showed a significantly higher number of harmonics and, similarly, a higher relative strength of overtones to fundamental. (From the published examples, which show partials up to 15 KHz, one concludes that any harmonic that emerged above background noise was counted). Air flow analysis, which was conducted with four of the seven subjects, showed significantly different rates in the paired tones, but in one subject air flow in the whistle register was greater rather than less, and none of the four managed to produce the whistle register at a high enough SPL to match that of the head register. There is no discussion in the article of the physiological or acoustical basis for the differences measured.

The limitations encountered in Walker's study are illustrative of the difficulty of isolating registers as a variable. In his, as well as other studies dealing with chest/falsetto, it is commonly presumed that if a given singer can keep pitch, SPL and vowel constant, then only a change in register (which is defined as a change in the voice source) can account for variance in phonation. The number of factors, as well as their interaction, is actually far more complex than this four-factor, simplified model suggests: the parameter 'constant vowel', for example, can cover a considerable variety of vocal tract articulations; likewise, an enormous variety of intensity detail in (micro) time is lumped into 'constant SPL'. A further problem in such studies is that the pairs of tones where 'register is the variable' are drawn from the range where registers overlap, a range where it is difficult to produce the lower register softly enough and the higher register loudly enough so that the tone pairs are matched in intensity. Even in cases where it is possible to produce pairs with matching intensity the result is seldom natural.

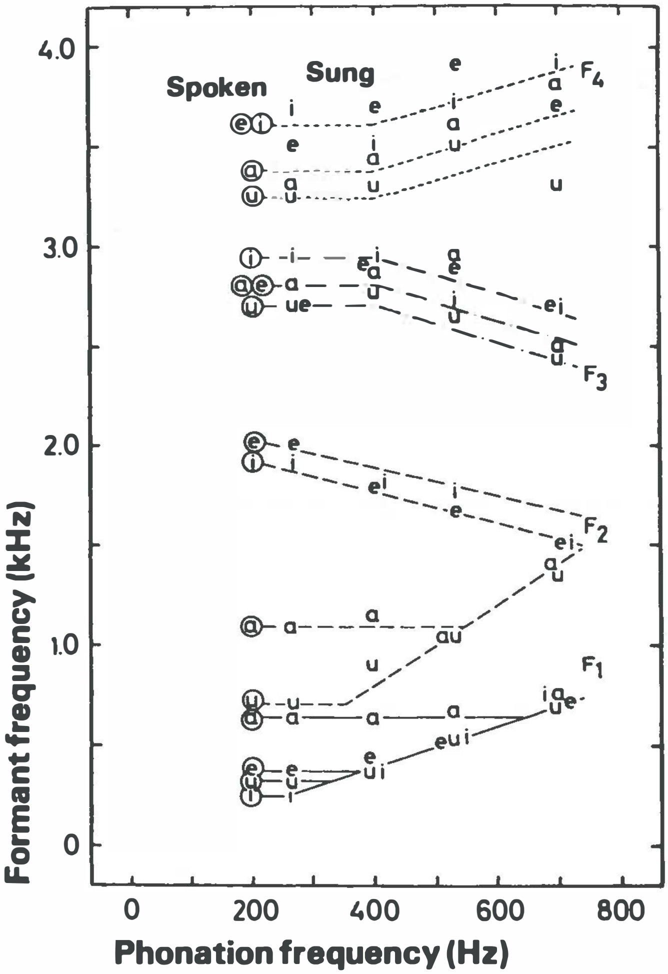


**Figure 1**
Frequencies of the first four formants of vowel indicated, sung by a soprano at various pitches. The circled values pertain to the subject's speech. The first formant is not allowed to fall below the phonation frequency. (Reprinted with permission from Sundberg, 1987)

In this study we have chosen to examine high-pitched phonations where register is not singled out as the variable of interest but rather where vocal maneuvers are executed as singing tasks, performed according to the skills and judgement of our subjects. This results in the close association of registers with segments of the voice range, and we have no 'isoparametric' tone pairs (Large, 1968) in different registers for comparison. In compensation, observing that a skilled singer will generally make a given register transition at the same pitch, given the vowel and intensity level, we avoid asking the singer to do something not ordinarily practiced. The maneuvers seem to us to be reasonably well executed, a judgement confirmed by the informal consensus of a group of Dutch voice teachers. We shall attempt to identify measurable characteristics that distinguish the production of the highest segment of the voice range from that of its lower neighbor (hereafter referred to as 'flageolet' and 'upper' registers, respectively) and then address the question of the appropriateness of designating them separate registers. The variables that get the most attention in this study are vocal fold motion, which is monitored by electroglottography (EGG), and the first two vocal tract formant frequencies, whose measurement at such high F0's has been a problem for voice researchers.

Sundberg (1973) measured these formant frequencies and then argued convincingly that in the range we designate here as upper voice the soprano adjusts the first formant frequency nearly to coincide with the first harmonic (the fundamental), a maneuver that, by his estimate, can add up to 30 dB to the resultant sound. However, his summary of these formant adjustments as the voice ascends in pitch, given in Figure 1, goes no farther than 800 Hz (about *G5*) (Sundberg, 1987). Some questions are left unanswered: Is there a limit beyond which it is no longer feasible for the soprano to raise her first formant? If so, what is the frequency of this limit? Is this related to the singer's perception of a separate register for the highest notes?

The measured phonations in this study should be considered a pilot investigation using as a first step the results from two singer-subjects, both successful and active professional singers in the Netherlands. Subject A, in her early thirties, has enough ease and confidence in producing the highest segment that she risks singing sustained *F6*'s in public. Subject B does not sing so high, but the fact that she can (and does in public) still produce exemplary high notes in her middle fifties is taken as evidence of a sound vocal technique.

### *Results of measurements*

Figure 2 gives an overview of signals from two similar phonations: upward leaps of a major sixth, followed by a return to the original pitch, both sung by subject A on the vowel /ɑ/. The pitch of the upper note in the first phonation is *G5-sharp* (831 Hz), a note well within the segment (about *F5* to *B5-flat*, 698-932 Hz, respectively) of the female voice that we have designated as the

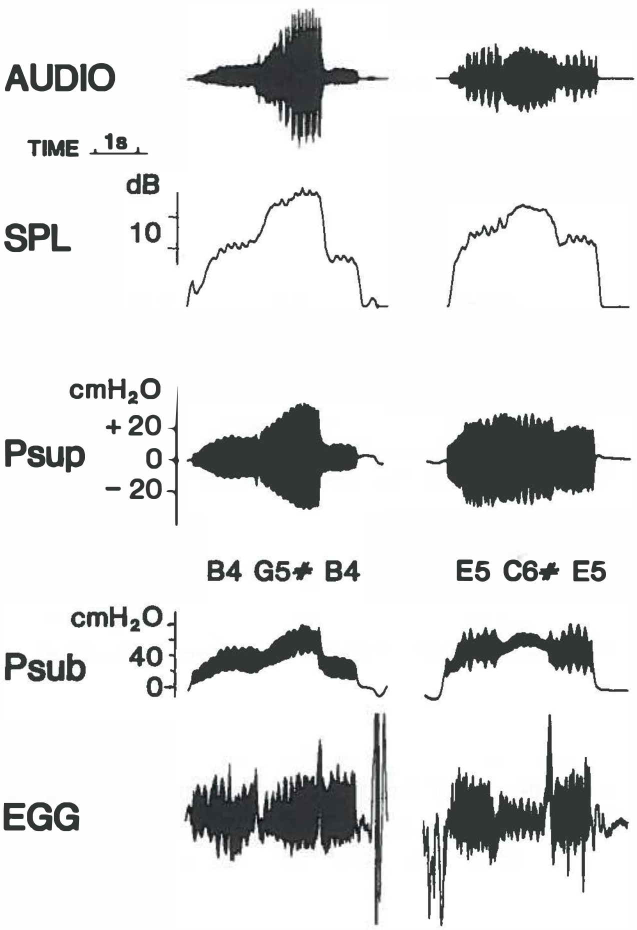


**Figure 2**
Overview of signals from two phonations of a soprano (subject A). The signals are, in descending order, from a microphone 30 cm in front on the subject: sound pressure level (SPL), derived from the microphone signal; supraglottal ($p_{supra}$) and subglottal ($p_{sub}$) pressures, from wide-band pressure transducers on a catheter passed through the posterior commissure of the glottis; and vocal fold contact area (upward is more contact), registered by electroglottograph (EGG). Both phonations encompass the interval of a major sixth.

upper voice. The corresponding note in the second phonation is *C6-sharp* (1,109 Hz), a note within the highest segment of the singing voice, according to our hypothesis. The signals given in the figure, in descending order, are the audio signal, from a microphone in front of the singer; SPL derived from the

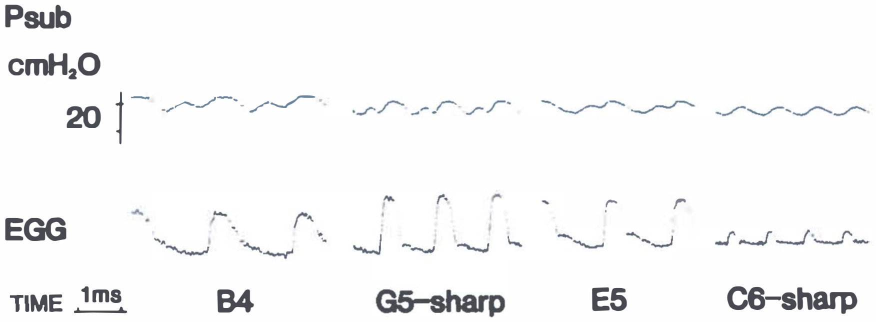


**Figure 3**
Details of the subglottal ($p_{sub}$) and electroglottograph (EGG) signals from the four pitches of phonations in Figure 2. On the highest pitch the small-amplitude, sinusoid modulations of $p_{sub}$ and the marked reduction in vocal fold contact area are indications of reduced vocal fold oscillations and a constantly open glottis.

microphone signal, supraglottal ($p_{supra}$) and subglottal ($p_{sub}$) pressures, from wide-band miniature pressure transducers mounted on a catheter passed via the nose through the glottis; and vocal fold contact area, from an electroglottograph, type Laryngograph.

Aside from the controlled parameter of fundamental frequency there are any number of differences in the details of these two phonations; of particular interest, however, are the limited peak-to-peak modulations of $p_{sub}$ and vocal fold contact area (EGG) on the upper note of the higher phonation. In Figure 3 we see some of these signals in detail: the contour of the signals of $p_{sub}$ and EGG, respectively, do not differ radically among the *B4*, *E5* and *G5-sharp*. On the *C6-sharp*, however, the maximum contact appears markedly diminished, and the fraction of the glottal cycle that shows any (increased) contact is also reduced. The sharply reduced sinusoidal peak-to-peak modulations of psub are consistent with reduced vocal fold motion. The sinusoidal form and small peak-to-peak amplitude of the $p_{sub}$ wave are indications of small vocal fold motion and a continuously open glottis. The subglottal space is apparently coupled with the supraglottal space, and the $p_{sub}$ scarcely shows any influence of the subglottal formants. Together the two signals suggest that the voice source is considerably altered on the highest note. Other recorded phonations of this subject consistently show these sharply reduced peak-to-peak EGG and psub signals in the highest segment of the range.
Figure 4 gives two examples of one-octave chromatic scales beginning at *D5 flat*, first *piano* and then *forte*. Details are shown of the EGG signal, first

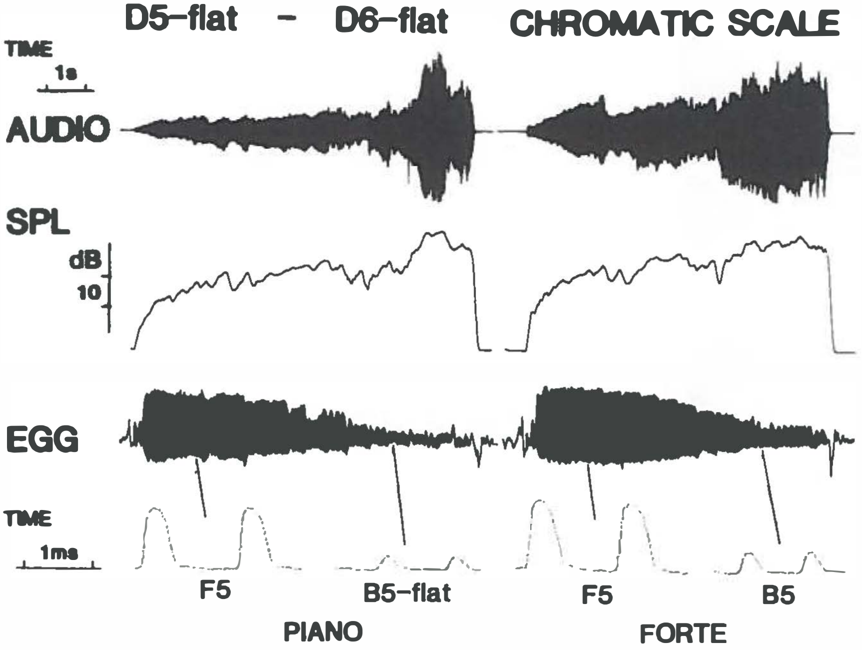


**Figure 4**
Overview of signals from a one-octave chromatic scale ascending from *D5-flat*, sung first *piano* and then *forte* by subject A. Signals as in Fig. 2. Details of the electroglottograph (EGG) signal are taken from the points in the overview indicated by the connecting lines. The higher-frequency details are from the pitches where the EGG peak-to-peak amplitude has reached a stable minimum, after diminishing gradually with rising frequency. SPL = sound pressure level.

taken from the lower portion of the scale and then from the point in the scale where the EGG peak-to-peak amplitude appears to reaches a stable minimum. In both cases this amplitude diminishes with increasing F0 up to the point where the small, stable amplitude is reached. This point varies slightly with intensity, occurring on *B5-flat* in *piano*, and on *B5*, one semitone higher, in *forte*.

Further evidence that the amplitude of the peak-to-peak EGG signal varies not only with F0, but also with intensity of phonation is given in Figure 5, a sustained *B5* with a marked crescendo of ca. 10 dB, produced by subject B. The details of the EGG signal show that the phonation begins with a small sinusoid, adds a sharp point, and finishes with a large, round-topped signal

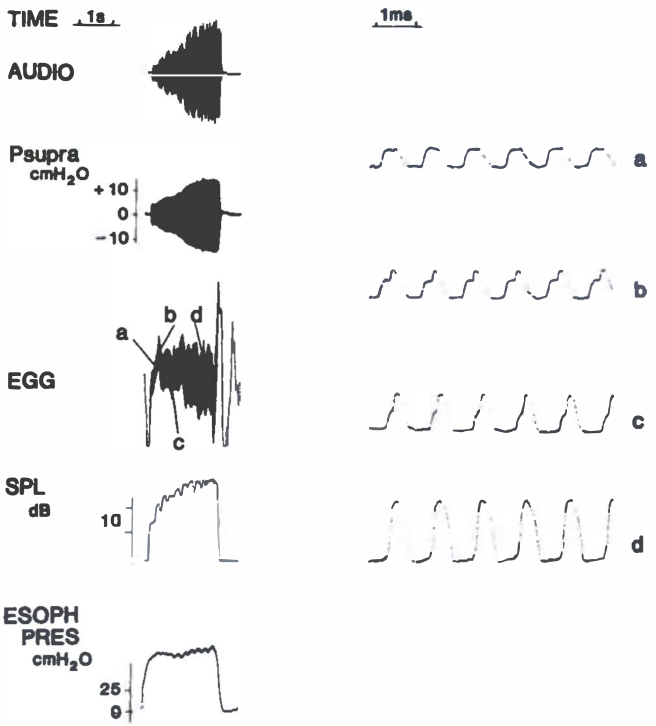


**Figure 5**
The pitch *B5* sung with increasing intensity by subject B. Signals are as in Figure 2, with esophageal pressure (ESOPH PRES) substituted for subglottal pressure. On the right side of the figure and time details of the electroglottograph (EGG) signal, taken from the points indicated in the overview. Note the progression from a small-amplitude sinusoid wave to fuller contact as the sound pressure level (SPL) increases. $P_{supra}$ = supraglottal pressure.

similar to that of subject A (see Figure 3) a few semitones lower. (The $p_{sub}$ signal, in contrast to those of Figure 2, was measured in the esophagus, and for this reason its low peak-to-peak amplitude has no particular significance).

Figure 6 depicts spectrograms of four phonations of subject B: portions of the sustained notes *A5* and *B5*, along with the corresponding ingressive phonations made with the vocal tract articulations of these two sung notes. (Ingressive phonations, made while carefully holding a given vocal tract articulation, produce a non-harmonic voice source and thus a continuous spectrum, clearly revealing the frequencies of at least the first two formants. For details, consult Miller and Schutte, 1990). Comparison of the spectra reveals significant differences between the two notes: (a) On the *A5*, F1 is

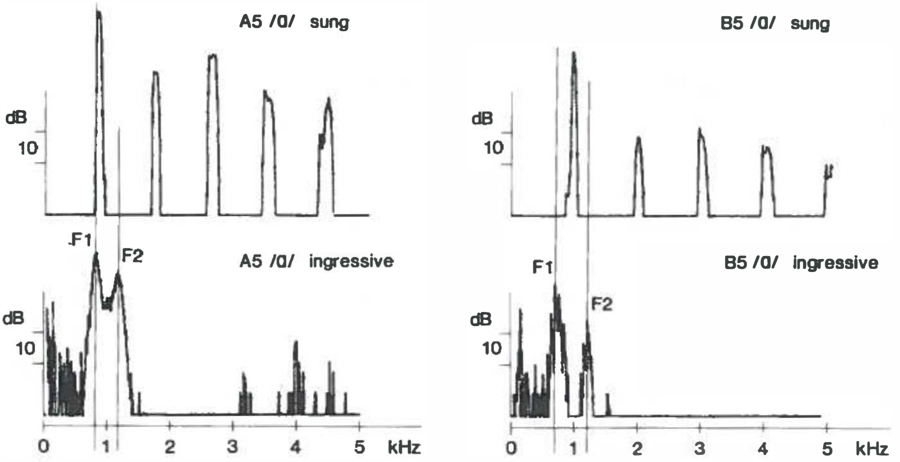


**Figure 6**
Spectrograms of two sung phonations by subject B, together with ingressive phonations while holding the vocal-tract articulations of the two sung notes. The ingressive phonations give a close estimate of the frequencies of the two lowest formants. Note that F1, which matches the phonation frequency on the *A5* (880 Hz), is allowed to fall below the phonation frequency on the *B5* (988 Hz).

apparently tuned to the F0 (allowing for a small error of ca. 50 Hz, which we attribute to the subjects failure precisely to hold the sung articulation), while on the *B5*, F1 is allowed to fall significantly below F0. (b) The marked proximity of F1 and F2 (ca. 325 Hz apart) on the *A5* is relaxed on the *B5* (ca. 475 Hz apart). (c) The ratio of the level of the fundamental to the higher partials is markedly higher in the *B5* (on these two phonations, which were not controlled for total SPL, the difference between the level of the fundamental and the next strongest harmonic is 10 dB greater on the *B5*).

There are two phenomena in the measured data of the highest pitch that attract special attention: the apparent falling of F1 below the F0 and the reduced EGG signal, suggesting reduced vocal-fold motion. We shall consider them in order. Figure 6 presents strong evidence that the highest pitch in which the adjustment of F1 to match F0 is still feasible is reached by subject B somewhere between 880 and 980 Hz. While she was not informed beforehand of the purpose of the experiment, questioning afterwards revealed that, although her own thinking about these notes does not include awareness of a transition between registers, she does feel a definite difference in 'placement' between the two notes.

The position of the second formant is of particular interest. It differs decidedly from that given in the chart of Sundberg, who places F2, for all vowels on the highest notes, near the second harmonic. The measurements presented here, as well as others we have made, reveal a tendency of the two lowest formants to converge in the singer's upper range, especially on the open, back vowels, /ɑ/ and /ɔ/, which to some singers seem to be the only practicable vowels on their highest notes. The vocal tract articulation that produces extreme convergence of F1 and F2, which we see here on the *A5*, apparently requires considerable tension. This particular tension is released on the *B5*, where the effort to raise F1 to the level of F0 (now 980 Hz) is abandoned. The *B5* does not necessarily feel easier to the singer, however, since the higher pitch demands increased tension of another sort.

If F1 indeed rises above F0 on the highest notes, there remains the question of the large loss (up to 30 dB) of SPL that Sundberg calculates would result from failing to match F1 and F0 in the high range. Our data do not show a marked reduction in SPL where F1 falls below F0: subject B can make a strong crescendo on B5 with only a small increase of Psub, and subject A's *forte* chromatic scale (Figure 4) shows a steady crescendo on ascending, at least through the pitch *C6*. On the basis of the data presented here it appears that the strong resonance the first harmonic (the fundamental) finds in the vocal tract is either the second formant, located just above it, or the combined effect of the closely spaced F1 and F2. It should also be noted, however, that in discussing the interaction of F1 and F0 Sundberg is referring to the part of the soprano range that we call here the upper register. The segment that we call flageolet is not considered directly by Sundberg.

Turning to the change in the EGG signal that we measured on the upper notes, we would like to make reference to an investigation by Rothenberg (1988) of the effect of perturbation of vocal tract resonance on vocal fold vibration in two sopranos. By means of a device which momentarily extended the vocal tract, lowering F1 by an estimated 220 Hz, changes in vocal-fold motion (monitored by EGG) could be observed at certain pitches. While in much of the singing range the (involuntary) formant lowering had a negligible effect on the EGG waveform, there were some pitches where the waveform went from a typical falsetto (female 'head-voice') pattern, similar to that in our Figure 5d, to a sinusoid, similar to that in Fig. 5a, when the vocal tract was perturbed. These pitches included the highest tested in both subjects, *G5* and *A5-flat*, as well as, in the case of one subject, all pitches in the range where F1 is tuned to F0 (ca. 600 Hz and higher). Although the precise form of the

EGG wave differs from the sinusoid of Rothenberg's sopranos and our subject B, subject A (who sings with relative ease in flageolet range) consistently produced a minimal EGG pattern beginning at *B5* or *B5-flat*, as indicated in Figure 4. This suggests a lower boundary of the flageolet register at the same frequency as that where subject B let F1 fall below F0. The two phenomena probably have a causal connection, since the force driving the vocal tract oscillations, according to Titze (1983), is theoretically reduced when F1 falls from a frequency above F0 to one below F1.

One further observation on Rothenberg's study strengthens our argument. His study aimed to test 'the sensitivity of a voice to a change in vocal tract resonance', and his method was to employ 'perturbation' of resonance frequencies, rather than a particular change from one frequency to another. However, if one subtracts the presumed amount by which F1 was lowered by (220 Hz) from the *B5/B5-flat* where our data suggest that the flageolet register begins, one arrives very close to the *G5/A5-flat* where both of the sopranos in Rothenberg's study showed a marked reduction in vocal-fold oscillation. It appears that the effect of lengthening the vocal tract was to lower, by about three semitones, the point of transition into the flageolet register.

There is one final point which argues in favor of regarding what we have designated 'flageolet' tones as a separate register. In Figure 5 we have already noted the increased dominance of the first harmonic in the spectrum. This confirms the finding of Walker (1988) and is presumably correlated with the success of his listeners in discriminating between whistle and head registers. In summary, we have made a case for calling the highest notes of the female singing voice a separate register, the lower boundary of which is the pitch where it is no longer expedient to adjust the first formant upward to match the fundamental frequency. In the data examined, this pitch is approximately *B5* (988 Hz), but some variation can be expected due to differences in size and adjustment of the vocal tract. The voice source typically found in this register, characterized by minimal/reduced vocal tract oscillation and no apparent phase of complete closure, appears to be a secondary phenomenon brought about, at least in part, by the reduced driving force on the vocal folds when F1 falls below F0. Our examples show it is possible to sing in the flageolet range with a head source, and other measurements (not shown) indicate that the flageolet source can be used in the head range. Although we do not address the larger question of what constitutes a register, the distinct acoustical adjustment and its characteristic voice source seem adequate justification for applying the term, at least in the usage of singers.

Of the various names used for this register, flageolet seems the most appropriate, not so much because of the suggestion of flute-like sounds (not all sounds in what we have described as flageolet registers are flute-like), but because of the parallel with flageolet tones (harmonics) in stringed instruments. The image of producing the very high notes with a lighter touch, while the string vibrates in a higher and somewhat elusive mode of oscillation, is apt. In the voice, however, the higher mode of oscillation is a feature of the vocal tract, not the voice source.

## Conclusions

The upper register of the female singing voice, defined here as that segment where the first formant closely matches the fundamental frequency on all vowels, ends at a point where it is no longer expedient to adjust the vocal tract for further raising of F1. The pitches above this point are designated here as the flageolet register. Besides this essential acoustical feature, the flageolet register is characterized by reduced vocal fold oscillation, although the further oscillation of the upper register is sometimes used in the vocal-tract articulation of the flageolet register.

Since these conclusions are based on data from only a few subjects, it may be best to regard them as a working hypothesis until a broader database can be accumulated.

## Acknowledgement

This work was supported by a grant from the Netherlands Organization for Scientific Research (NWO) and the assistance of the College of Visual and Performing Arts, Syracuse University, Syracuse, New York, U.S.A. The authors appreciate the assistance of Danny Kuipers and last, but not least, Meindert Goslinga in preparing the manuscript and figures.

## References

Large, J. An acoustical study of isoparametric tones in the female chest and middle registers in singing. *NATS Bulletin*, 1968;25:12-15.

Miller, R. *The Structure of Singing*. Schirmer, New York. 1986, 147-149.

Miller, D.G. and Schutte, H.K. Feedback from spectrum analysis applied to the singing voice. *Journal of Voice*, 1990;4:329-334.

Rothenberg, M. Acoustic reinforcement of vocal fold vibratory behavior in singing. In: Fujimura, O. (ed.). *Voice Physiology: Voice Production, Mechanisms and Functions.* Raven Press, New York. 1988, 379-390.

Sundberg, J. Observations on a professional soprano singer. Speech Transmission Laboratory – *Quarterly Progress and Status Report*, Royal Institute of Technology, Stockholm. 1973;1:14-24.

Sundberg, J. *The Science of the Singing Voice.* Northern Illinois University Press, Dekalb, IL. 1987.

Titze, I. The importance of vocal tract loading in maintaining vocal fold oscillation. In: Askenfelt, A., Felicetti, S., Jansson, E.V., Sundberg, J. (eds.). *Proceedings Stockholm Music Acoustics Conference.* Royal Swedish Academic Music, Stockholm. 1985;1:61-72.

Walker, J.S. An investigation of whistle register in the female voice. *Journal of Voice*, 1988;2:140-150.

# Chapter 7

# Toward a Definition of Male ‘Head’ Register, Passaggio, and ‘Cover’ in Western Operatic Singing

# Toward a Definition of Male ‘Head’ Register, Passaggio, and ‘Cover’ in Western Operatic Singing

Donald G. Miller and Harm K. Schutte

## Abstract

For male singers in the western operatic tradition the upper part of the frequency range contains a series of pitches called passaggio, the transition between ‘chest’ and ‘head’ registers. Although there is no generally accepted objective definition of the distinction between ‘head’ and ‘chest,’ the distinction is widely recognized among singers and voice teachers, as is the term ‘cover,’ a technique widely believed to help in a proper execution of passaggio. On the basis of spectrographic, electroglottographic, and sub- and supraglottal pressure measurements on professional singers' ‘correct’ and ‘incorrect’ singing of passaggio, this study offers a definition of the ‘chest’/’head’ distinction.

## Introduction

From a scientific viewpoint, the jargon of singers is notorious for its vagueness. Since singers, however, seem to know what they mean with such terms as 'forward placement,' 'focus,' and even science-defying 'head resonance,' scientists need to have objective definitions of such concepts in order to facilitate serious empirical and theoretical study of singing. The task of definition, nonetheless, remains a difficult one. Not only is the jargon difficult for outsiders to penetrate, but the singers themselves often disagree as to what they mean with such terms.

To an orderly mind it may seem futile to spend time on such concepts, but one should bear in mind that this vagueness presents less of an obstacle to the practice of vocal pedagogy than it does to voice science. If the pupil learns, by some combination of imitation, suggestion, trial and error, and luck, to 'cover' in the manner approved by the teacher, the fact that 'cover' means something slightly different in another studio is of secondary importance. The term is not meaningless, but imprecise. To avoid traditional terms (and the often emotional arguments that rage around them) by introducing more precisely defined new terms (for example, the terms 'modal' and 'loft' for registration phenomena) is to risk further deepening of the communication gap that separates scientists from the practical world of singing and teaching. The scientist knows what he means, but the singer pays little attention.

An alternative strategy is to give precise, objectively demonstrable descriptions of the phenomena to which the singer's terms refer in the hope that the singer will recognize what he 'means' with the term in the more precise 'scientific' definition. The success of this by no means simple task will depend to a large extent on the degree to which the objective description directly reflects the crucial identifying characteristic in the perception of the singer (or teacher, whose business it is to be more discriminating of such subtleties than the singer himself). Simplicity of concept definition and the possibility of objective measurement are essential in order to keep the terms from being merely theoretical. If the essence of a term can be thus identified and recognized in practice, it is conceivable that teachers (and singers) will welcome the sharpened insight that the definition provides and use the term more precisely.

### *Passaggio*

If a tenor trained in the western operatic tradition (which excludes the abrupt break to falsetto) sings a scale passage between *D4-flat* and *A4-flat* (or a baritone the passage *B3-flat* to *F4*), it is commonly said that he has sung through 'passaggio', the transition between 'chest' and 'head' registers, a range in which the terms 'covered' and 'open' have special relevance. All the words in quotation marks are part of singers' jargon, and in spite of some vagueness, the knowledgeable voice teacher has a pretty good idea of what he means when he uses them.

For a variety of reasons, some readers may object to the label 'head' for the 'register' under consideration in this study. Alternatives to 'head' would be either a different name or leaving the phenomena described here without a convenient label. The retention of 'head' seems to us to offer not only the least of three evils, but also distinct advantages. It locates the present study within the tradition of scientific investigation (Vennard et al., 1970; Large, 1973). Moreover, it reflects the authors' experience that many of the vague entities referred to in the international pedagogical literature, however imperfectly described, have real acoustical and physiological referents. Similar arguments apply to the term 'cover'.

Defining them for the nonsinger and, more to the point, giving a verifiable description of the phenomena they refer to, are tasks which have intrigued voice scientists, especially those who come from the tradition of voice pedagogy, such as Vennard (1970) and Large (1973; 1970). Characteristic of the problem of term definition, however, is the fact that these, as well as other, authors rely on the self-reporting of singers regarding the question of the 'register' in which a given tone is produced. Perceptual studies generally confirm that expert listeners agree with the singer's self-description at a rate well above that of chance (Beard, 1980), but firm definitions which allow one clearly to distinguish 'head' from 'chest' are not achieved. Even the electromyographical study by Vennard et al. (1970) avoids the question of a sharp definition of the line between 'head' and 'chest', leaving some skeptics with the question of whether these two are indeed separate 'registers' (Holien & Schoenaard, 1980).

Rather than attempting a comprehensive description of differences between 'head' and 'chest' registers, a description that would have to include F0-related aspects of subglottal pressure and vocal fold biomechanics, vocal fold oscillation patterns and compensatory (or facilitating) vocal-tract articulatory adjustments, the present authors will attempt to isolate that feature by which singers perceive that the 'head'/'chest' distinction is real, and not merely theoretical.

### *'Register Violation'*

In an extensive explication of registration phenomena in the singing voice Miller (1977) introduces the phrases 'registration event' and 'register violation,' terms which signal directly the singers experience of registers. The 'event' of moving upward (or downward) to a note which seems to the singer (or listener) to be produced in a clearly different way from the previous one, insofar as this difference is not merely the result of differences in pitch, intensity and vowel, is ascribed to 'registration,' a category which is generally thought to pertain to the manner of vibration of the vocal folds. The move from 'chest' to 'falsetto' register is an obvious event of this sort; the move from 'chest' to 'head' is less obvious, in keeping with the lack of clear, verifiable definitions of the 'registers' in question. A 'register violation,' however, in the sense of using an inappropriate register (usually a higher note produced with a register appropriate to a lower segment of the voice) is a concept which every voice teacher and trained singer will recognize (Miller does not bother to define it), even if the appropriateness of a register or the potential danger to vocal health of a 'violation' may be hotly disputed. When illuminated by the readily apparent 'register violation,' the distinction between 'head' and 'chest' can no longer hide behind its multi-dimensional subtlety. Perhaps in the 'violation' it will be possible to measure objectively that feature which clearly signals the distinction to the singer.

## Materials and Methods

The measurements supporting the broad argument of this article come from phonations supplied by one of the authors (D.M.), a bass-baritone with extensive training and professional singing experience, as well as many years of giving lessons to aspiring and performing singers. Since the repeatability of professional singer's phonations depends in part on whether or not they are executed in a habitual mode, it is of relevance that this singer's 'register violation' is not far removed from a mode he formerly used habitually, as he has changes his approach to passaggio more than once in the course of his career, trying both more 'covered' and more 'open' voice production.

The phonations recorded for analysis comprised leaps upward (*E3 - B3 - E4*) and diatonic scale passages (*A3 - E4*) on the open vowels /ɛ/, /ɔ/ and /ɑ/ (as in bed, ball and father) in 'standard' (the singer's current practice), 'open' and 'exaggerated cover' modes at an effort level of unforced forte. The whole series was repeated a semitone higher (*F3 - C4 - F4* leaps and *B3-flat* to *F4* scales) except for the 'open' scales, which were excessively strenuous. The vocal tract articulation of the final note of each phonation was held as accurately as possible and an ingressive phonation, yielding a continuous spectrum for clear identification of the formant frequencies (Miller and Schutte, 1990) in that articulation, was recorded immediately. A multichannel FM instrumentation recorder at a tape speed of 30 inches/s was used to record the following signals: $p_{supra}$, esophageal pressure measured by means of miniature wide-band pressure transducers on a small-diameter catheter introduced through the nares with the tip into the esophagus; the vocal fold contact area, measured by an electroglottographic signal (laryngograph) and the radiated sound. Spectrum analysis was per-formed by a FFT Real Time Spectral Analyzer, Model 4512 (Princeton Applied Research), incorporating a built-in Hanning weighting network. The measuring procedure has been described extensively elsewhere (Schutte and Miller, 1988).

## Results

Since the first formant frequencies of these open vowels in an easy, lower-middle range normally varies between 500 and 650 Hz in this singer, we were particularly interested in the effects of the encounter of the second harmonic (designated H2) with this resonance as the fundamental ascends the scale (on both scale passages, the fundamental begins below 250 and rises above 330 Hz; thus H2 varies through the range <500 to >660 Hz).

Figure 1 shows spectrograms of the pitches *D4* and *E4* in the 'open' production of the vowel /ɔ/. Figure 3 shows spectrograms of *D4*, *E4-flat* and *F4* in the 'standard' production of the same vowel. Each spectrogram, which displays the average distribution of frequency components over a 200-ms 'time slice,' is accompanied (Figure 2 and 4) by a much smaller time segment of the signals from the microphone and $P_{supra}$ (pharyngeal) transducer, selected to illustrate the dominant resonance of the tone, if any is apparent.

Looking first at the 'open' production, the overwhelming dominance of the second harmonic on the E4 (22 dB above the level of the first harmonic, and 10 dB more than the second highest) is very strong evidence that F1 is tuned to H2, all the more because F1 would then have the exceptionally high value (for this singer) of 670 Hz. The strong tuning effect is confirmed by the $p_{supra}$ measurement, which shows a very high peak-to-peak value of 45 cm of water pressure (Fig. 2). Further confirmation of the high value of F1 was obtained from the spectrogram of the ingressive phonation in this articulation.

The signals for the *D4*, by comparison, indicate that H2 (here about 70 Hz lower), although resonated by F1, is not yet high enough in frequency to be maximally affected. H1 is only 9 dB below the level of H2. Both the spectrogram and the audio signal, with its clear four-part form, show a stronger tuning of F2 to H4, even though this tuning is intermittent in the vibrato cycle.

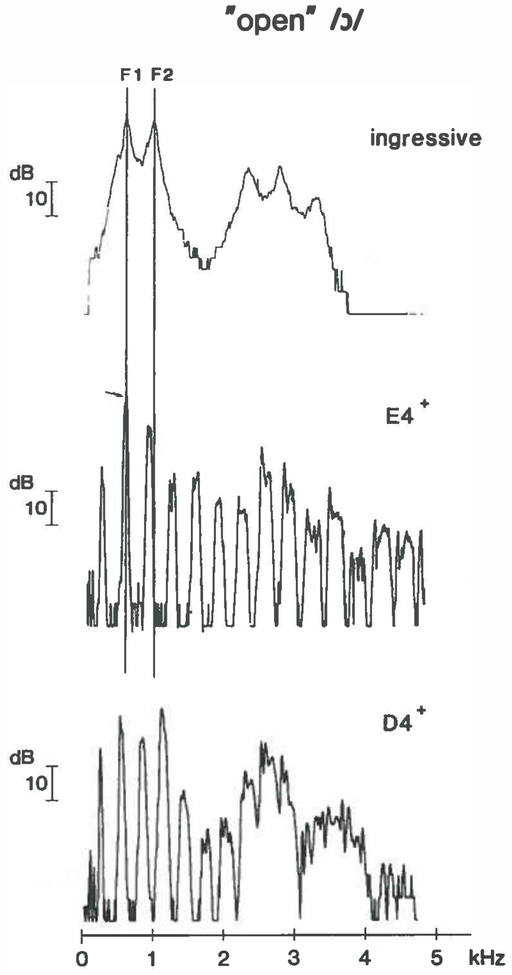


**Figure 1**
Spectrograms taken from the last two tones of the scale passage *A3* to *E4*, sung on the vowel /ɔ/ by a bass-baritone with 'open' articulation, illustrating a 'register violation.' The ingressive glottal source produces a continuous spectrum in the articulation of the *E4*, aiding in locating the formant frequencies. The singer has adjusted F1 to the high value of 670 Hz in order to match the second harmonic of the *E4*, which is strongly resonated (the arrow indicates the upper limit of the harmonic). It is apparent that the upward move of F1 to match H2 on the *E4* has resulted in a lowering of F2, which resonated H4 on the *D4*.

The highest notes of the scale passage in the 'standard' articulation (Figures 3 and 4) show a quite different pattern. Here the impact of F1 on the second harmonic appears to be maximal on the *D4*, judging from the dominance of

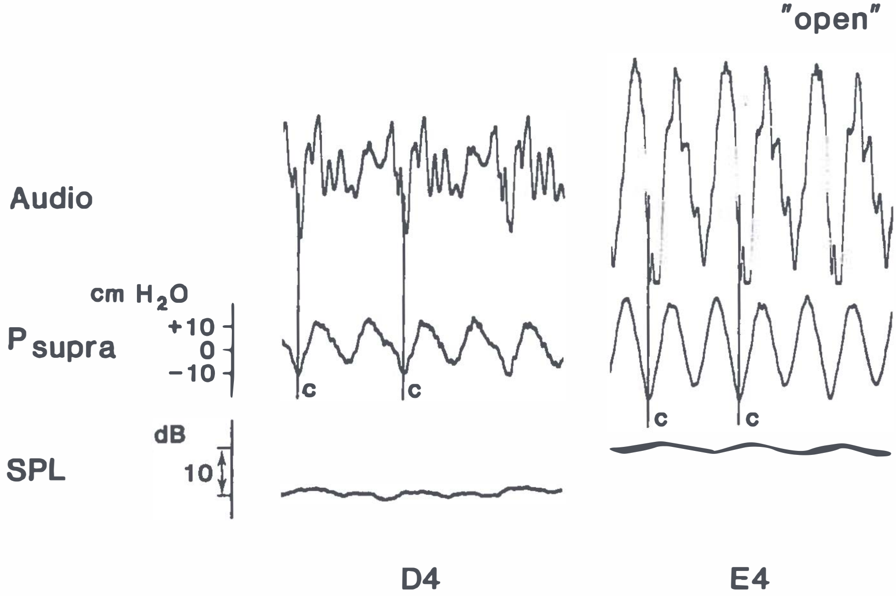


**Figure 2**
This figure depicts the signal from a microphone 30 cm in front of the subject, a pressure signal from a miniature transducer a few centimeters above the glottis, and the SPL at the microphone. The audio signal is adjusted for delay, and the vertical lines marked c indicate two consecutive closures of the glottis. The high peak-to-peak values of $p_{supra}$ on the *E4*, indicating a dominant standing wave from the impact of F1 on the second harmonic, are also reflected in the microphone and SPL signals.

that harmonic in the spectrogram and the high (maximal for the whole phonation) value of the peak-to-peak $p_{supra}$ signal. On the *E4-flat*, a semitone higher (only 25 Hz difference at the frequency of the second harmonic), F1 has evidently fallen below H2, and the first two harmonics are abruptly equal in level. The *E4-flat* becomes a sort of 'passing note' falling between the resonant *D4* and *F4*, which are respectively, 7 and 11 dB higher in sound pressure level (SPL). The nearly equal level of the first four harmonics, the reduced $p_{supra}$ signal showing only a weak F1 resonance, and the (by comparison) amorphous audio signal all indicate that F1 and F2 fall relatively ineffectively between harmonics 1 and 2 (F1) and 3 and 4 (F2). The last note, *F4*, has the highest SPL of the phonation, not only because of the increased subglottal pressure ($p_{sub}$), which in itself does not always produce a greater SPL on the higher pitches, but because of the clear tuning of F2 to the third harmonic, the level of which rises a full 17 decibels above any other harmonic in the spectrum.

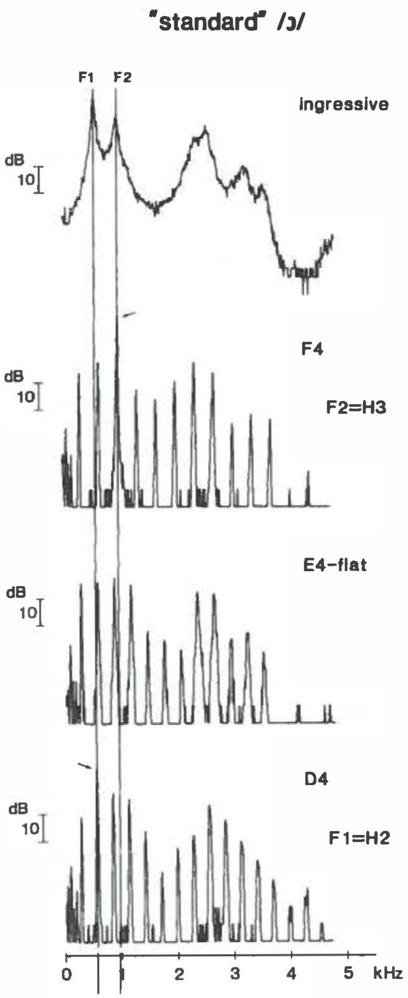


**Figure 3**
Spectrograms taken from the last three tones of the scale passage *B3-flat* to *F4*, sung on the vowel /ɔ/ with articulation that allows the transition from 'chest' to 'head' register. As in Figure 1, the ingressive spectrum reveals the configuration of formants of the highest note. In this case, however, there appears to be little movement of the formants between scale steps, and both the dominant H2 of the *D4* and the H3 of the *F4* align with F1 and F2, respectively.

Because less of the F2 acoustic energy is dissipated in the vocal tract, and because the point at which $p_{supra}$ is measured may result in an underestimation of the true value (see discussion in Miller and Schutte, 1990) the peak-to-peak $p_{supra}$ is unimpressive compared with the tuned F1, say, on the 'open' *E4*;

the dominance of the three-part form, however, is evident both here and in the audio signal.

Sung spectra, nonharmonic (ingressive) spectra, SPLs and $p_{supra}$ curves presented here as evidence of the (lack of) dominance of a particular formant-harmonic combination in the vocal tract, all support and confirm one another. Of the four, however, it is the $p_{supra}$ signals which offer the strongest, most direct evidence of standing waves in the vocal tract, and these are further corroborated by the microphone (audio) signals, which were maintained at constant gain for Figures 2 and 4. The sung spectra and SPLs are further effects of these same signals, while the ingressive spectra, which may be subject to uncertainty regarding the singer's ability to reproduce the sung articulation, are the least essential to the demonstration.

## Discussion

While singers and voice teachers may vary in their judgment of our 'standard' example of negotiating passaggio, the authors expect that the great majority of such professionals trained in the western operatic tradition will easily identify the 'open' example, upon hearing it, as an attempt to extend the 'chest' register too high for the structure of this particular voice, a 'fault' we have designated a 'register violation.' A 'correct' execution of the scale passage calls for a (ideally, well-disguised) 'registration event' which will put the voice in a 'legitimate head register,' '*voce piena in testa*,' or whatever one chooses to name that adjustment which brings the male singer beyond the upper frequency limits of his shouting voice.

What characterizes the widely recognized upper limit of the 'forced call'? While many authors (Miller, 1977, p. 104) have attempted to locate this limit in the balance of intrinsic muscles in the larynx (with the 'register violation' considered an static over-balance of thyroarytenoid over cricothyroid), our investigations indicate that the boundary is set at the point where F1, firmly linked to the second harmonic, reaches its upper limit. With the vocal folds committed to the 'chest register' pattern of oscillation, with its long (>50%) closed phase, the voice finds a very strong natural resonance where the first formant and second harmonic coincide. In most men's voices this occurs when the fundamental is in the neighborhood of *E4* (330Hz). The shouter (or untrained singer) is able to force his voice a few semitones higher by raising his larynx and lifting F1 to follow H2. The muscular strain he feels in so

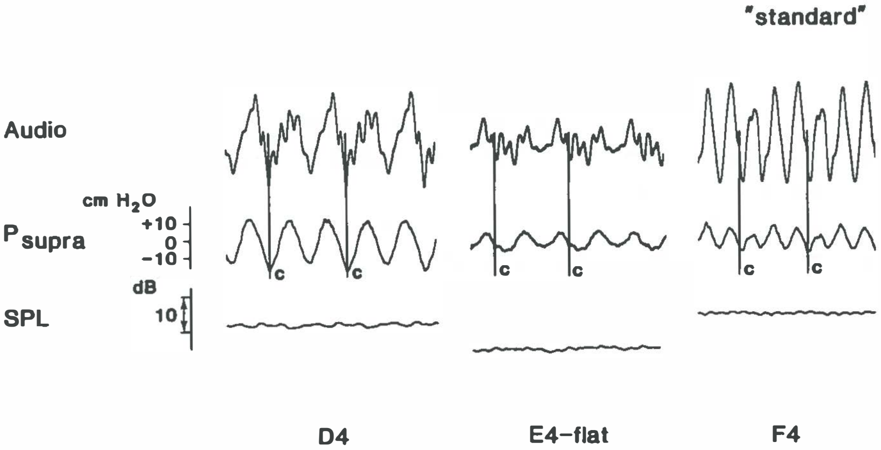


**Figure 4**
Signals, similar to those in Figure 2, taken from the three sung tones whose spectra appear in Figure 3. The dominance of the standing waves of H2 and H3 in the *D4* and *F4*, respectively, is apparent in both the vocal tract ($p_{supra}$) and the radiated sound (audio). The rising of H2 above F1 in the *E4-flat* results in a large loss of resonance, reflected in SPL.

doing may be perceived as approaching a maximal 'stretching' of the vocal folds, but actually he has extended to the limit his means for raising F1. In this adjustment the voice is, as it were, locked on a particular resonance; with the resonance at its upper frequency limit, the voice cannot move higher without 'breaking' into a falsetto adjustment. (This instance of dependence of vocal-fold oscillation on a favorable resonance is a complex phenomenon whose explanation goes beyond the limits of this paper. Here we merely assert it.)

In contradistinction to this, as the trained singer ascends the scale in passaggio he anticipates the trap of the 'register violation' and avoids it by allowing F1 to fall below H2 before the two become locked together with the help of an elevated larynx position. This description yields a first approximation of a definition of the distinction between 'chest' and 'head' registers in the male singing voice, at least for open vowels sung at relatively high intensity: the 'chest' register comprises the lower portion of the range, extending to the highest pitch where the first formant matches the second harmonic; the 'head' register comprises that portion of the range where F1 falls below H2. Thus the change in the position of the first formant, from (just) above to below the second harmonic, is the primary 'registration event' which all male voices must deal with in the western operatic tradition, with its demands for extended singing range.

The terms ‘open’ and ‘covered’ also derive (Miller, 1977, pp. 81ff, 137ff) from this ‘registration event.’ When used disapprovingly, the most characteristic meaning of the term ‘open’ (*offen*, *aperta*) is to describe the sound of the elevated F1 ‘register violation’ as in our ‘open’ example (Figure 1). Those who approvingly use the term ‘covered’ (*gedeckt*, *coperta*) might define it as the ‘correct’ counterpart of the (too) open sound on the pitches which properly belong in ‘head’ register, as defined above.

Both terms, ‘open’ and ‘covered,’ naturally develop extended meanings, referring to articulations of the vocal tract with higher and lower (open-vowel) F1 frequencies, respectively. Even in that majority of teachers and singers who agree about the more obvious ‘register violations,’ the desirable degree of ‘cover’ is a matter of much dispute.

A particular technical approach, encountered most frequently among low male voices and rejected by many teachers as ‘too covered’ has at least the practical advantage of resolving the passaggio problem (in our definition: letting F1 fall below H2) low enough in an ascending scale that the danger of the ‘register violation’ is minimized. Whether the cost, in terms of vowel distortion, unnatural sound, etc., of this convenient solution justifies its use is another matter.

The definitions and measurements presented here have all been acoustic, in accordance with the authors' assumption that the singer's control mechanism for these ‘events’ is primarily auditory. Mention should be made, however, of certain physiological correlates of the acoustic parameters under discussion. The following summary is based on Fant's *Acoustic Theory of Speech Production* (Fant, 1970): The vocal tract can be modeled as two cavities connected by a movable constriction between the tongue and the opposing side of the vocal tract. In the formant range of the open vowels considered here, the back cavity is ‘affiliated’ with F1, and the front cavity with F2. Extending the back cavity (the equivalent of lowering the larynx) has the effect of lowering all the formants, but especially F1. Narrowing the mouth opening or protruding the lips will also lower all the formant frequencies, but especially F2, because of its affiliation with the cavity most directly affected. Opposite effects on the formants can be produced by raising the larynx or increasing the mouth opening, respectively. A third basic means of changing the first two formants is to change (advance or back) the position of the tongue constriction, a move which lowers F1 while raising F2, or vice versa.

Thus the principal means for achieving the acoustic effect of 'cover,' which we have described as a lowering of F1, are (1) lowering the larynx; (2) diminishing the mouth opening/ protruding the lips (e.g. modifying the vowel /ɑ/ toward /ɔ/ or (3) 'closing' a forward vowel (e.g. modifying the vowel /ɛ/ toward /e/). 'Open' singing is produced by opposite measures. In the complex real world we would expect to find these three maneuvers combined with one another as well as with other measures (e.g., the raising of the velum, which generally accompanies larynx lowering). Although these techniques can be consciously employed, in practice they are more often the result of unconscious adjustments in response to the singer's anticipation of a desired sound. In the phonations presented here these physiological adjustments were not measured, although it is evident that they occurred.

Our preliminary definition of 'chest' register for the high F1 vowels is fairly simple; the acoustic profile of 'head' register, on the other hand, admits a variety of forms, most of which will only be briefly mentioned in this study. When, on open vowels, the male singer's fundamental frequency is higher than the point where F1 easily matches H2, he has the following principal options: (1) to change the oscillation patterns of his vocal folds to that of 'falsetto,' with short or absent closed phase of the glottis (falsetto); (2) to reduce the closed phase of the glottis in a soft phonation which preserves the possibility of a smooth (step-free) crescendo to a loud phonation (voix mixte); (3) to extend the range a few semitones further in fairly soft phonation by allowing F1 (and usually the larynx) to rise (*voce finta*); (4) to extend the range a few semitones in loud phonation, allowing F1 and the larynx to rise (male belt); (5) to allow F1 to fall below H2, while F2 produces a dominant resonance on H3 (for back vowels /ɑ/ and /ɔ/) or on H4 (for front vowels /æ/, /ɛ/ and /œ/) (F2-tuning); (6) to allow F1 to fall below H2, while F2 falls low enough to produce, in proximity with F1, a dominant H2 (dark cover); (7) to allow F1 to fall below H2, while the singer's formant (2400 - 3200 Hz) constitutes the dominant resonance (singer's formant dominance).

The first option, falsetto, is not generally considered an acceptable, legitimate alternative in the western operatic tradition. Like voix mixte (which is an important resource of the legitimate tenor), it is distinguished from (robust) 'chest' register primarily by an adjustment in vocal fold function. In both of these the shorter (or absent) closed phase of the glottis favors the dominance of the first harmonic (Miller, 1987), and the singer tends to lower F1 to engage this harmonic, even on the open vowels. *Voce finta* (feigned voice) and male belt (a term derived from a similarly named, more widely used female register function) are the mild and strenous forms, respectively, of the 'register violation.'

The last three options are all instances of what we have called robust 'head' register. 'Dark cover' is a refuge chiefly of voices whose upper extension is expected to be limited. Its spectrograms, generally with dominant second harmonics, look suspiciously like those of chest register, supporting the doubts in the minds of some as to whether this should be classified as 'head' register (see discussion in Miller, 1986). F2-tuning, to which we shall return later, constitutes the staple of the operatic tenor or (high) baritone: in combination with a well-developed singer's formant it is the basis of the full, unforced high notes the public expects from such voices. Singer's formant dominance completes the list, occurring on high notes where F2-tuning is either unfeasible or not employed.

This array of options may seem unduly complicated to anyone who has not spent years trying to master the perils of passaggio, but it is merely a sketch of what we consider to be the most determinative clues that guide the singer's (and teacher's) perception. Much of importance in a more complete description of the 'head'/'chest' distinction is passed over. To begin with, our discussion is limited to open vowels and, for the most part, to loud phonation. This simplification is in part because it appears to us that the passaggio problem, from a singer's viewpoint, is in its essence a problem of open vowels. Separate attention will have to be given to the question of what happens to close vowels in the upper range. These have low first formants, which fall uneventfully below H2 in the less challenging middle range of the voice. Furthermore, other barriers in the higher range besides the F1-H2 encounter are important: an obvious one is the limited ability of the vocal folds to accommodate the necessary stretching and thinning, while less obvious is the hindrance presented by the relatively fixed subglottal resonance (Titze, 1988). We also see clear evidence of subtle adjustments in vocal fold function and $p_{sub}$. These, as well as the articulatory adjustments of 'cover' (for a more complete discussion see Hertegård et al (1990), may occupy the foreground of the singer's (teacher's) attention, while the clues we have called determinative may be taken for granted. Passaggio is indeed complex; it is not for nothing that excellent tenors are rare.

An operatic tenor needs to be able to cross the threshold between 'chest' and

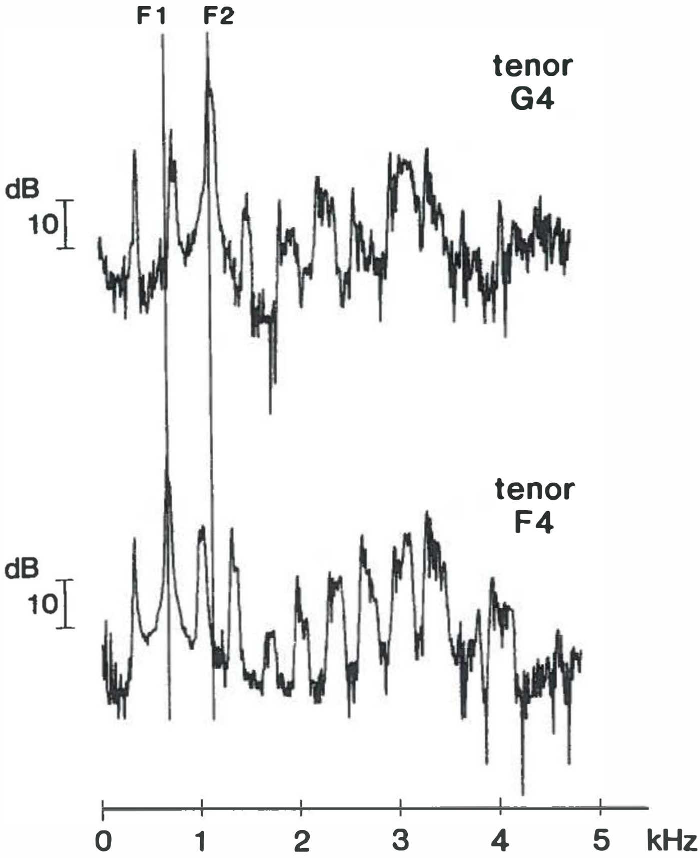


**Figure 5**
Spectrograms from the pitches *F4* and *G4* taken from a *B3-flat* to *B4-flat* scale, sung by a tenor on the vowel /ɑ/. Here the move from 'chest' (F1=H2) to 'head' (F2=H3) is accomplished in one scale step without any need to move the formants.

(robust) 'head' register not just occasionally, but repeatedly, almost routinely. Aside from the widely recognized fact that this 'registration event' occurs at a higher pitch than it does for the baritone (Miller, 1977; p. 129), are there other factors that distinguish these voice categories in the move to 'head' register?

The spectrograms in Figure 5 are taken from the pitches *F4* and *G4*, neighboring tones in a *B-flat* scale, sung on the vowel /ɑ/ by an operatic tenor who

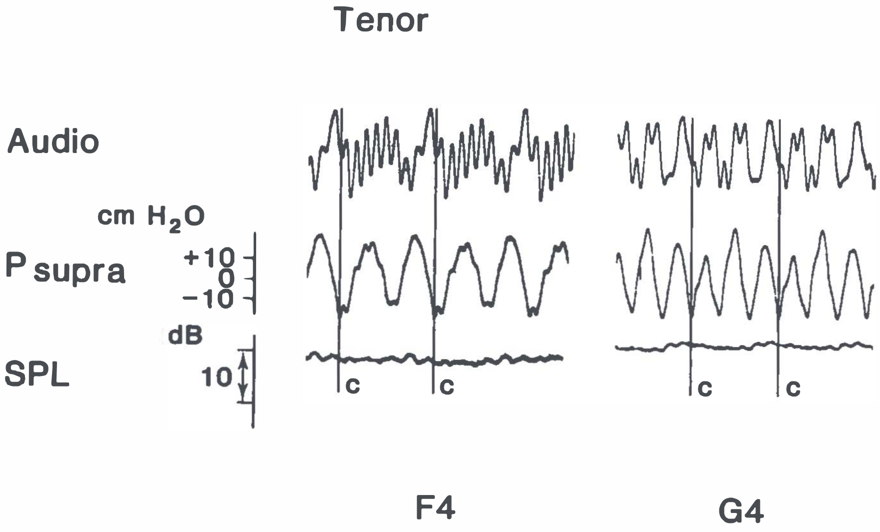


**Figure 6**
Signals, similar to those in Figure 2, taken from the sung tones whose spectra are shown in Figure 5. The standing wave ($p_{supra}$) in the vocal tract has nearly the same amplitude in F2 dominance as in F1 dominance.

is making an international career with roles such as those in *Tosca* and *Madame Butterfly*. They were recorded under the same circumstances as those of Figures 1 and 2. Even a superficial comparison of this example of passaggio with that in Figure 2 reveals prominent differences.

One of these is that the tenor shows no weakly resonated scale step between the two with clear F1 and F2 dominance, whereas on the pitch *E4-flat* the bass-baritone, while already in 'head' register according to our definition, has not yet reached a point where F2 resonance is the decisive component in the sound. Another point of difference is in the respective $p_{supra}$ signals: in moving from F1 to F2 dominance, the peak-to-peak amplitude, indicating the magnitude of the standing wave in the singer's vocal tract, hardly changes in the tenor (Figure 6), whereas in the bass-baritone it is reduced by a factor of two. While the factors behind this difference in the two voices invite further investigation, one can say at the very least that the difference suggests that, in comparison with the bass-baritone, the tenor is more at home in 'head' register. (Our 'standard' baritone example, however, should not be construed as implying that such a voice cannot go directly from 'chest' to 'F2-tuning' on successive scale-steps in passaggio.)

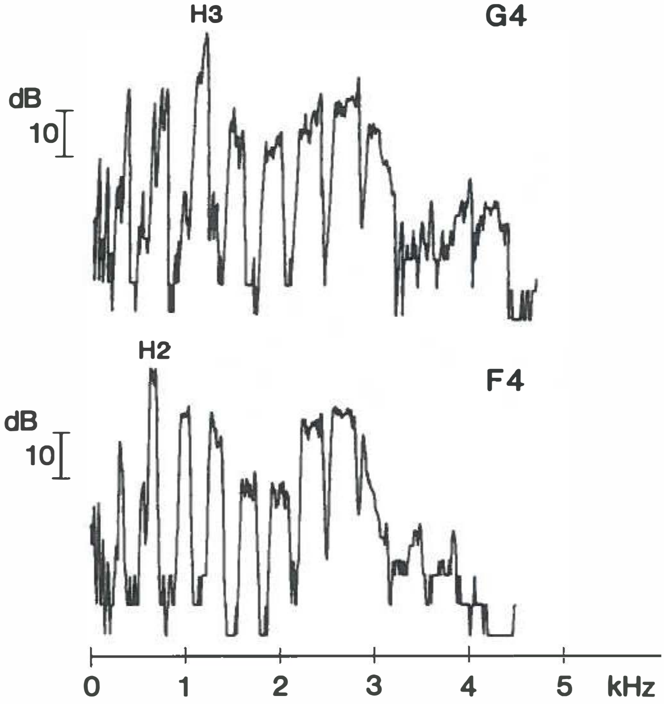


**Figure 7**
Spectrograms taken from sustained vowels /ɑ/ at two unaccompanied moments from the recitative *'Tombe degl' avi miei'* (Lucia di Lammermoor, G. Donizetti), as recorded by tenor Placido Domingo. As with our experimental subject, the *F4* and *G4* show F1 and F2 dominance, respectively, and belong in 'chest' and 'head 'registers as defined in the text.

The tenor's clear move from 'chest' to 'head' register with F2-tuning raises a point about the use in passaggio of 'cover' which, in our formulation, lowers F1 in order to facilitate this move. Among voice teachers there is disagreement about the degree to which 'cover' should be actively applied, that is, whether at a given point in passaggio one should (consciously or habitually) make the move to 'head' register occur by lowering the larynx, rounding the lips, etc. Without attempting to resolve this controversy, we would merely point out the theoretical possibility, as well as the practical example given here, of a transition without active 'cover.' When the tenor moves upwards from *F4* to *G4* his second harmonic moves approximately from 700 to 800 Hz, and his third harmonic from 1,050 to 1,200 Hz. With his first two formant frequencies on the vowel /ɑ/ at (just above) 700 Hz and 1,200 Hz, he can go from (chest) F1-tuning to (head) F2-tuning without any modification of the vowel.

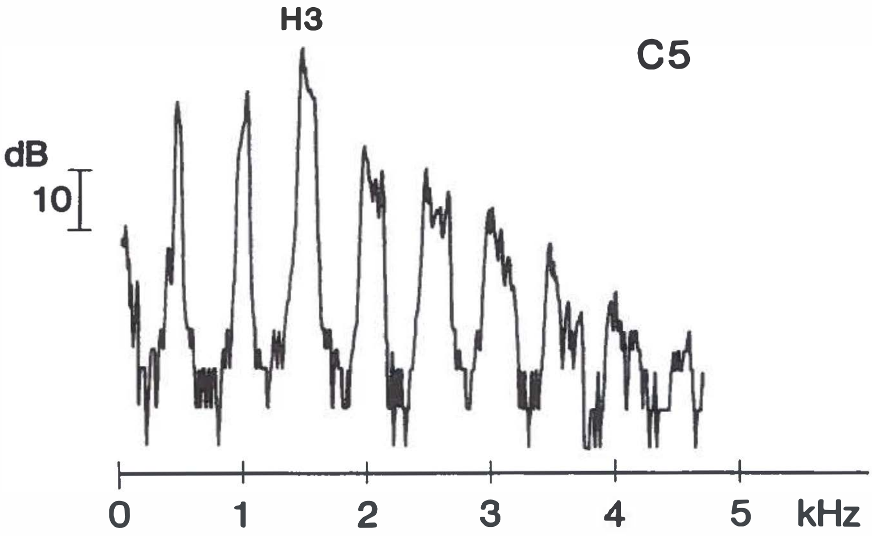


**Figure 8**
Spectrogram taken from the final *C5* of the aria *'Ah, mes amis' (La Figlia del Regimento*, G. Donizizetti), recorded by tenor Luciano Pavarotti. The dominant third harmonic results from the impact of the second formant.

A moment's reflection tells us that not all such transitions can be accomplished so neatly and that the singer sometimes may well have to make a conscious decision to 'go over', with appropriate modification of vocal tract articulation. If F2-tuning is an important part of the picture, however, then a generalized 'cover' with both larynx lowering and lip rounding will not offer a satisfactory solution, since the matching of F2 to H3 or H4 will often require a raising of the second formant. The build of the instrument (the size and balance of front and back cavities), the vowel with its possible modifications, and the fundamental frequency (including the frequency excursions of the harmonics due to vibrato) will all be important factors in determining where the first two formants can be most effectively positioned.

Nonetheless in the male upper range the number of satisfactory solutions is limited, and one tenor will tend to resemble another with respect to the way F1 and F2 fall among the first five partials. Figure 7 shows, by way of illustration, spectra of *F4* and *G4* on the vowel /ɑ/, taken from recordings of the tenor Placido Domingo. One can see that these exhibit the same 'chest' and 'head' patterns as the spectra of our measured subject.

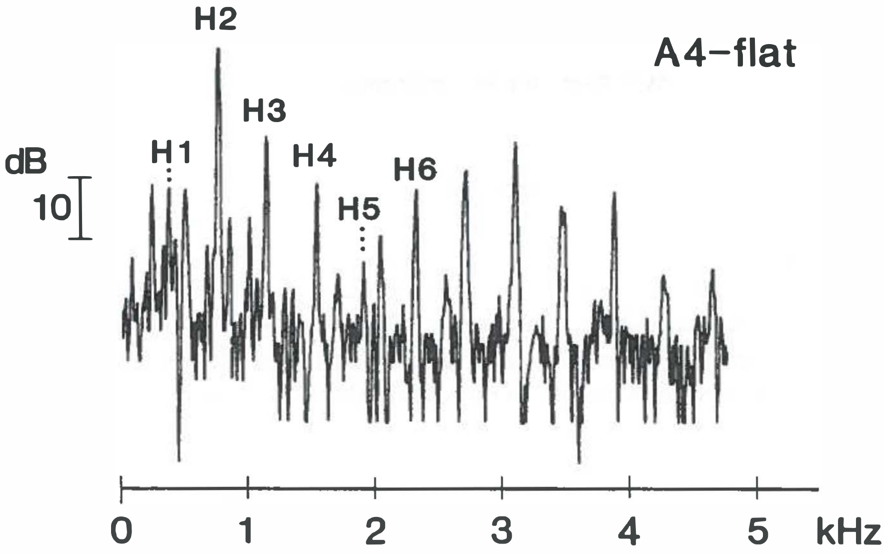


**Figure 9**
Spectrogram taken from the final held note (*A4-flat*, vowel /ɔ/) of the duet 'You're nothing without me' from the cast recording of the musical *City of Angels*. The first six harmonics of the voice have been indicated against the orchestral background. The clear dominance of H2 confirms the perceptual impression that the first formant rises above 800 Hz to resonate this harmonic.

As the fundamental frequency rises beyond *G4*, a point will be reached where it can be problematical for the singer to get the second formant high enough to tune to the third harmonic, which reaches approximately 1,500 Hz on the tenor's high *C*. Figure 8, a spectrogram of one of the several which Luciano Pavarotti sings of this pitch in *La Figlia del Regimento*, shows that it nevertheless can be done.

A final illustration from the recorded literature brings us back to our point of departure, the 'register violation.' Because of the high value placed on understandable diction and 'natural' sound, the 'covered' sound is generally shunned in musical theatre and popular styles of singing, sometimes even in instances where the pitch is high enough that the 'register violation' is almost painful to a listener with 'classical' training.

Figure 9 is a spectrogram taken from a *G4* of the vowel /ɔ/, the final note of the duet, '*You're nothing without me*,' from the original cast recording of the Broadway musical *City of Angels*. The clear pattern of F1=H2 tuning can be seen against the orchestral background in the spectrogram. Although the singer, judging from perception, is a natural baritone, he evidently raises his first formant to nearly 800 Hz for this long note (7 seconds), which is repeated in a similar number elsewhere in the show.

The sophisticated amplification system used in this show helps him to be heard against the orchestra, even without the cultivated 'singer's formant' of his unamplified operatic colleague (Sundberg, 1977), but it is nonetheless a loud, energetic voice production, illustrating well what we have referred to as the 'male belt.' Regarding the potential danger of such production, presumably executed with elevated larynx, to the health of the voice, it is interesting to consider that the show is repeated eight times weekly and, since it is a popular success, over an indefinite period.

## Concluding remarks

The explanation offered here for passaggio, 'cover,' etc. is largely in agreement with the findings of some previous investigations. For example, Hertegård et al (1990) described the lowering of F1, as well as a number of the physiological features of 'cover.' Sundberg and Askenfelt (1983) identified two acoustic consequences of larynx raising – higher formant frequencies and a reduced level of the first harmonic – which are also apparent in our 'register violation' (Figure 1). Titze (1985) described the (un)favorable effect that F1 can have on vocal fold oscillation when it is in the vicinity of one of the lower harmonics, a phenomenon we also find where F1 encounters H2. What is new in the argument here is the observation of the tendency of the first formant to 'lock' on the second harmonic at the top of the 'chest' register, as well as the identification of the dropping of F1 below H2 as determinative for 'cover' and robust 'head' register.

While not denying a secondary role to adjustments in intrinsic musculature of the larynx, our argument bypasses the common theoretical explanation of the passaggio from 'chest' to 'head' as an (ideally gradual) reduction in contraction of the vocalis muscles combined with increased contraction of the cricothyroids. Aside from the fact that electromyographic data is inconclusive from the most extensive experimental investigation this question (Vennard, 1970; Hirano, 1970), we submit that the 'registration event' of passaggio is primarily about a change in the position of F1 with respect to H2, rather than a change in the balance of the intrinsic muscles of the larynx. Likewise, the essence of 'cover', in a sense which is also acceptable to those who are averse to a generalized darkening of the sound, is positioning F1 below H2, rather than larynx lowering, lip rounding, or even lowering F1 per se.

A 'proof' of the general validity of what we consider the determinative mechanism of the 'registration event' of passaggio is not offered here. All of our phonations and their measurements, however, are repeatable, as are, of course, those from recordings. Validation will occur if investigators and singers, especially those who have the opportunity to confirm with spectrum analysis what they hear with their ears, find these definitions in accordance with their experience.

## Acknowledgements

This work was supported (partially) by a grant from the Netherlands Organization for Scientific Research (NWO). The authors appreciate the assistance of Ella Blokzijl, and especially Meindert Goslinga for his meticulous preparation of the figures.

## References

Beard, C. Recognition of Chest, Head, and Falsetto Isoparametric Tones. *NATS Bulletin* 1980;36:8-14.

Fant, C.G.M. *Acoustic Theory of Speech Production*. The Hague/Paris, Mouton & Co., 1970:63-90

Hertegård, S., Gauffin, J., Sundberg, J. Open and Covered Singing as Studied by Means of Fiberoptics, Inverse Filtering, and Spectral Analysis. *Journal of Voice* 1990;4(3): 220-230.

Hirano, M., Vennard, W.D., Ohala, J.J. Regulation of Register, Pitch and Intensity of Voice: An Electromyographic Investigation of Intrinsic Laryngeal Muscles. *Folia Phoniatrica* 1970;22:1-20.

Hollien, H., Schoenhard, C. The Riddle of the 'Middle' Register. In: Titze IR, Scherer RC, Eds. *Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control*, Denver, Colorado, Denver Center for the Performing Arts, 1985:256-272.

Large, J.W., Baird, E., Jenkins, T. Studies of the Male High Voice Mechanisms, pp. 109 - 115, In: Lawrence VL, Weinberg B, Eds. *Transcr IXth Symp Care Prof Voice*, New York: The Voice Foundation, 1980:109-15.

Large, J.W.. Acoustic Study of Register Equalization in Singing. *Folia Phoniatrica* 1973;25:39-61.

Miller, D.G.. Some Observations on the Soprano Voice. *NATS Bulletin* 1987;43(5):12-15.

Miller, D.G., Schutte, H.K.. Characteristic Patterns of Sub- and Supra-glottal Pressure Variations within the Glottal Cycle. In: Lawrence VL, Ed. *Transcripts of the XIIIth Symposium: Care of the Professional Voice.* New York: The Voice Foundation, 1984;1:70-75.

Miller, D.G., Schutte, H.K.. Formant Tuning in a Professional Baritone. *Journal of Voice* 1990;4(3):231-237.

Miller, D.G., Schutte, H.K.. Feedback from Spectrum Analysis Applied to the Singing Voice. *Journal of Voice* 1990;4(4):329-334.

Miller, R. *English, French, German and Italian Techniques of Singing: A Study in National Tonal Preferences and How They Relate to Functional Efficiency.* Metuchen, NJ: The Scarecrow Press, Inc., 1977.

Miller, R. *The Structure of Singing: System and Art in Vocal Technique.* New York: Schirmer Books, 1986:116ff

Sundberg, J. The Acoustics of the Singing Voice. *Scientific American* 1977; 236(3):82-92.

Sundberg, J., Askenfelt, A.G. Larynx Height and Voice Source: a Relationship?. In: Bless DM, Abbs JH, Eds. *Vocal Fold Physiology: Contemporary Research & Clinical Issues*, San Diego, CA: College-Hill Press, 1983:307-316.

Titze, I.R. The Importance of Vocal Tract Loading in Maintaining Vocal Fold Oscillation. In: Askenfelt A, Feli-cetti S, Jansson EV, Sundberg J, Eds. *Proceedings Stockholm Music and Acoustics Conference.* Stockholm: Royal Swedish Academy of Music. 1985:61-72.

Titze, I.R. A Framework for the Study of Vocal Registers. *Journal of Voice* 1988;2(3):183-194.

Vennard, W.D., Hirano, M., Ohala, J.J. Chest, Head, and Falsetto. *NATS Bulletin* 1970;27:30-37.

# Chapter 8

# Soft Phonation in the Male Singing Voice

# Soft Phonation in the Male Singing Voice

Donald G. Miller, Harm K. Schutte, and James Doing*

## Abstract

Sustained high notes, diminishing gradually from the loudest to the softest phonation in a maneuver called *messa di voce*, are examined in two contrasting professional tenor voices. Signals of the sound pressure level, electroglottograph, and mean esophageal pressure are recorded, and similar maneuvers by the same subjects are examined stroboscopically. The lyric voice is found to make a gradual diminuendo while maintaining nearly constant posture of the vocal tract together with a phase of complete closure in the glottal cycle. The robust voice, by contrast, passes abruptly from a production of high subglottal pressure and a high closed quotient to one of low pressure and incomplete closure, and the transition is marked by a sudden opening of the previously constricted laryngeal collar. It is proposed that the mode of soft voice production demonstrated by the robust voice be recognized as a distinct register of the singing voice.

---

* School of Music, University of Wisconsin-Madison, USA

## Introduction

From the enormous range of sounds the human voice is capable of producing, nearly all of which are capable of evoking an emotional response in the listener, the 'classical' singing voice makes use of only a small fraction. There are several good reasons for this self-limitation, four of which we will consider here. The first of these that comes to mind is the fact that this technique evolved in the context of the solo voice with orchestral accompaniment in relatively large spaces, generally demanding a high intensity level in order for the voice to be heard. Weak voices and conversational intensity are not appropriate for these circumstances.

Other reasons are perhaps less obvious. One has to do with the distinction between near-field and far-field communication (Tembrock, 1996): the singing voice is related to a sort of calling mode, shared by other species, in which the recognizable identity of the unseen sender is a large part of the message. This requirement implies that high and low, soft and loud, must all resemble in some essential ways the high-intensity production of the far-field mode. Another reason is that the sound produced must conform to and execute the patterns of music, resembling other musical instruments with respect to pitch, rhythm, 'line,' etc. A final requirement is that the sound itself must strike most ears in an appropriate audience as 'beautiful,' or at the least, not ugly.

These requirements greatly restrict the sounds that can be used by the singer, and at first glance they might be expected to reduce severely the range of emotion expressed in singing. This turns out, however, not to be a serious problem. In the first place, the singer does not address the audience alone, and the total music, of which the voice is only a part, is very effective in conveying nuances of a great variety of emotion. In the second place, the very intensity of typical classical sound production suggests a high degree of (emotional) arousal, which, in an ideal case, is then shared by the audience. The singer gives up most of the potential variety of vocal sounds – especially those of the near-field – but in return he gains an apparent emotional excitement.

In this musical and emotional context, the sounds at the low end of the intensity range occupy a peculiar place. In the conventions of music, they must suggest to the audience the delicate sounds that would be appropriate to express certain types of emotion, but they must do this without lapses in vocal identity or level of emotional arousal.

The soft 'high note' is where this paradox is realized par excellence. Since high notes are ordinarily loud, requiring high breath pressure, the degree of softness is perceived as a measure of willed restraint and therefore of an alternative sort of arousal. Without displaying his effort, the singer lets us feel a sort of defiance of gravity. Appearing to maintain the fullness of the instrument on a 'weightless' high note, the singer compels our attention and wonder.

## Pianissimo and *mezza voce*

The purpose of this study is to describe the essential features of what we consider a distinct register of the male singing voice. The name that we assign to this register is *mezza voce.* The argument supporting these suggestions will be easier to follow if a few terms are clearly defined from the start. In the singing voice we apply the term register to a set of tones that are perceived by either listener or singer as being of the same character, and distinct from some other group of tones of another character, with the two groups separated by a discontinuity in the way they are produced. (The discontinuity between the 'natural registers,' chest and falsetto, is the model for other, finer differences in register recognized by singers; the art of the singer is to bridge the discontinuities between registers.) *Mezza voce* and *messa di voce* are specialized terms in singing that are often confused, and not only by those who do not know Italian. The former refers to a quality of sound, and the latter to a singing maneuver of some complexity, sometimes called a swell-tone. The terms will be described more concretely in the course of the article.

The musical designation for the soft high notes examined in this study is ordinarily pianissimo, the extreme permissible level at the opposite end of the continuum from forte or fortissimo. *Mezza voce* (half voice, as opposed to *voce piena*, or full voice) also designates a soft sound; however, it also raises a question. Unlike, say, *con sordino* (with mute) for stringed instruments, the term *mezza voce* carries no simple indication of just what the singer must do differently. The existence of such a term, independent from pianissimo, would seem to imply that the two are not identical and that *mezza voce* would call for a different *quality* of voice, rather than merely the opposite of loud. While such a distinction may or may not be intended by composers employing the term, there appear to be singers who in practice make all their softest sounds with the same basic voice production, regardless of whether the designation is *mezza voce* or simply pianissimo. Considering the difficulty of producing an acceptable soft high note, it may be that such singers have no alternative.

An added challenge to the task of sustaining a high soft tone is that of producing a continuous *diminuendo* to such a tone from forte, as in the second half of a perfect *messa di voce* (a singing maneuver in which a sustained tone is gradually swelled from *pp* to *ff*, and then back to *pp*). Regarding the execution of the *messa di voce*, detailed descriptions by the most renowned singing teachers of the nineteenth century (Stark, 1999) reveal an important difference of opinion. While there was general agreement that the changes throughout the exercise should be gradual – thus disguising any discontinuity – Manuel Garcia taught that the maneuver went from falsetto to chest register and back again, and Francesco Lamperti famously proclaimed that 'the piano should in all respects, with the exception of intensity, resemble the *forte*.' In any case the question is complicated by the fact of variation among the types of voices, the difficulty of providing an objective description of the invisible registration events in the larynx, and the lack of precision in the terms used for describing these events.

The present study will attempt to shed some light on this problem area, as well as on the distinctness of *mezza voce* voice production, by a careful examination of the diminuendo portion of the *messa di voce* by two dissimilar singers.

## Method and results

Phonations of an extended *diminuendo* on a (non-extreme) high note from fortissimo down to the softest artistically responsible pianissimo were recorded from two highly accomplished professional tenors. Recorded simultaneously were a microphone signal at a distance of 30 cm, signals from an electroglottograph (EGG, presumed to follow vocal-fold contact area), and (mean) pressure signals from an esophageal balloon, reflecting subglottal pressure ($P_{sub}$). A variety of other phonations were recorded from the subjects of the experiment, and among these were samples of what they considered falsetto and *mezza voce*, as well as mean subglottal pressures measured along the contours of a phonetogram, that is, at maximal and minimal sound pressure level (SPL) for each semitone within the singer's F0 range. Their larynges were also viewed stroboscopically as they performed similar *diminuendi* to those displayed in the figures.

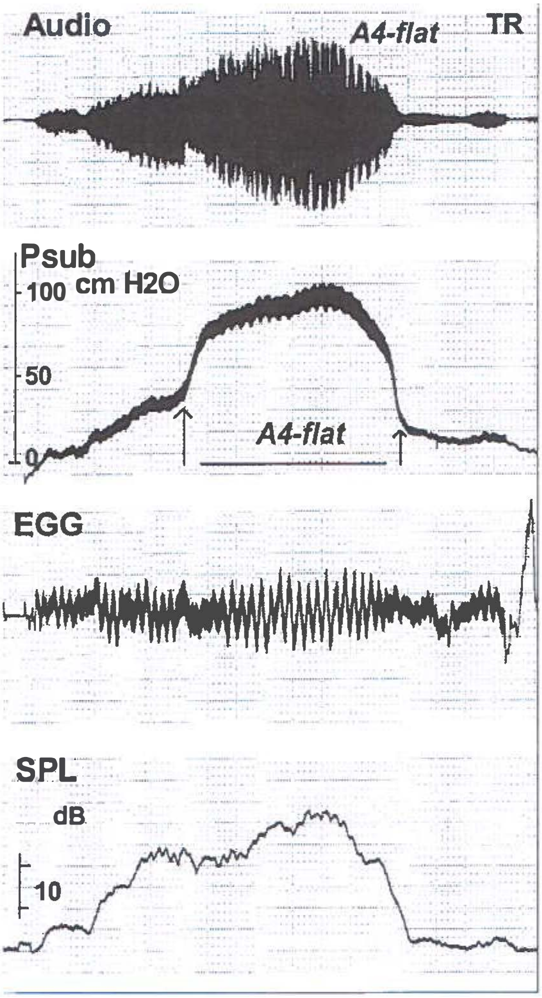


**Figure 1**
Overview of phonation with *messa di voce* on *A4-flat* by a robust tenor (TR). Signals shown are: microphone at 30 cm distance, esophageal pressure, electroglottograph, and sound pressure level. The *A4-flat* (nominally 415 Hz, but here given slightly lower), indicated between the arrows, is approached and left by arpeggio (*A3-flat, C4, E4-flat, A4-flat,* and reverse). The subglottal pressure rises to above 100 cm water pressure in fortissimo and falls relatively abruptly to ca. 15 cm water pressure in pianissimo at the end of the high note.

The subjects were chosen for the contrast between the types of voice they represent. One ('TR') sings the robust Italian operatic repertory internationally, while the other ('TL', one of the authors) is most at home in the characteris-

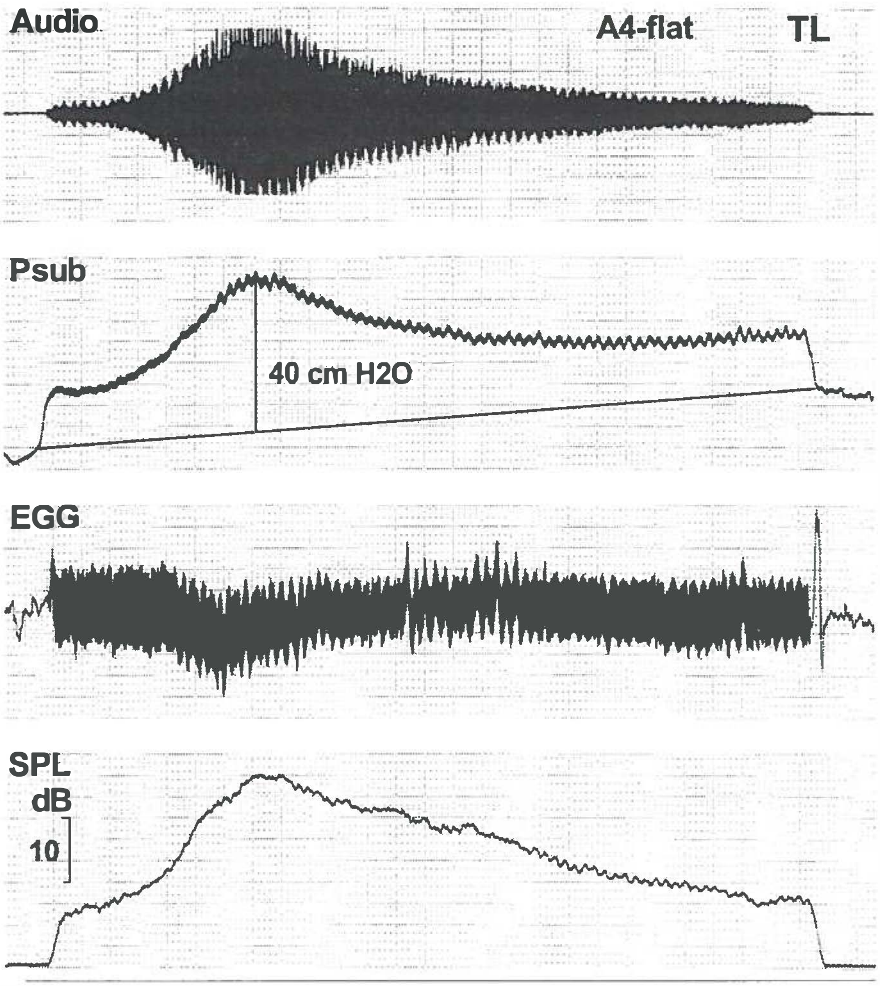
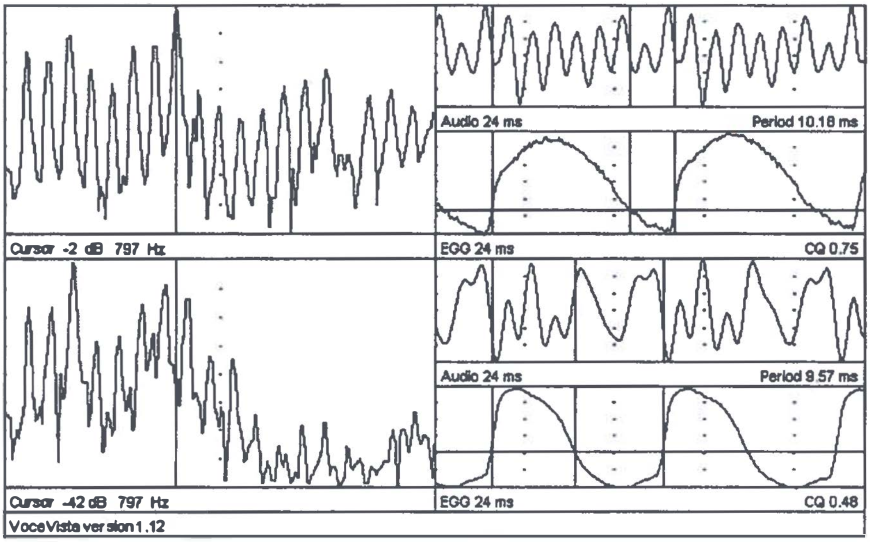


**Figure 2**
Overview of phonation with *messa di voce* on *A4-flat* by a light lyric tenor (TL). Signals are the same as in Figure 1. The crescendo to *ff* is intentionally steeper than the long diminuendo, which stretches out over 8 s of gradual decline in $p_{sub}$ and SPL. Maximal pressure in *ff* does not exceed 40 cm water pressure, while beginning and ending pressures, measured from the sharp rise and fall at the extremes of the phonation, are on the order of 15 cm water pressure. (The apparent rise in the zero level over the course of the phonation is due to the change in lung volume.)

tically 'lyric' repertory of Bach and Mozart. Both were in their middle to late thirties at the time of the recordings.



**Figures 3a-c**

Time details and spectra taken from the phonations in Figures 1 and 2, comparing the robust tenor (above) with the light lyric tenor (below), singing a *messa di voce* on the pitch *A4-flat* (nominally 415 Hz), vowel /ɑ/. Signals shown are power spectrum (left) and microphone and electroglottograph (right).

N.B. The write-outs are made from a slow playback, and the values displayed must be corrected by multiplying the frequencies (left) by four, and dividing the time (right) by four.

**Figure 3a**

The robust tenor (above) produces a fortissimo tone with a very high closed quotient (estimated at 75%) and a dominant standing wave at the frequency of the singer's formant (H7, nearly 3.2 kHz). The lyric tenor's *ff* has a CQ estimated at 48%, and the dominant standing wave is at the frequency of H3 (resonated by the second formant), ca. 1250 Hz. His singer-formant resonance is noticeable in the three maxima in the closed phase of the audio. Note also the strong harmonics of TR above 4 kHz, where TL produces very little sound.

Of the several instances of such *diminuendi* recorded from each singer, selection was made of one that best exemplified a successful example from a singer's point of view: a large, gradual drop in perceived loudness without the intrusion of apparent discontinuity, all while maintaining a single vowel (/ɑ/) with a quality of sound acceptable for use in performance (see Figures 1 to 3).

Of the examples displayed in the figures, subject TR's diminuendo (on an *A4-flat*, 415 Hz – Figure 1) possesses the greater contrast between loud and soft,

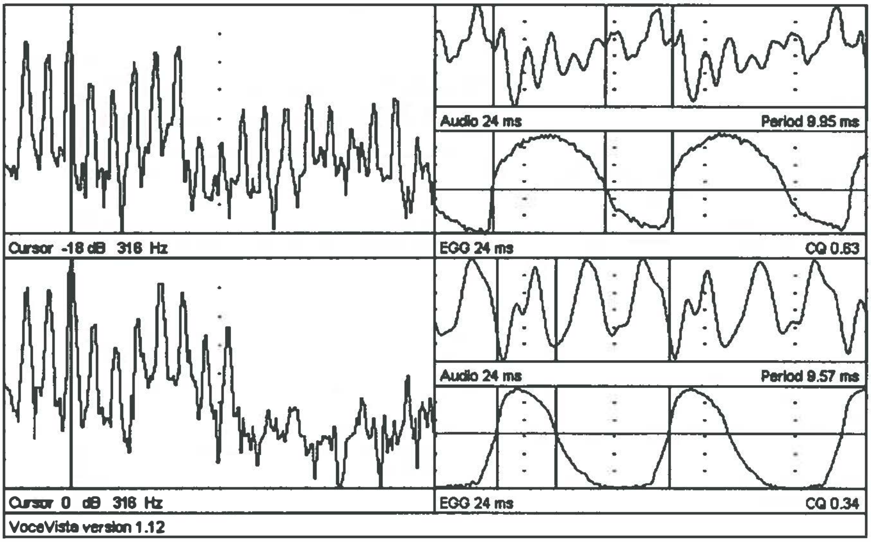


**Figure 3b**
TR's singer's formant resonance is considerably weaker at this intermediate point in the diminuendo, while his closed quotient is still high, though no longer extreme. With reduced CQ, TL retains his F2-dominated resonance, while the singer-formant component is hardly apparent in the audio signal.

both in terms of SPL and spectral composition. His fortissimo sound is dominated by the singer's formant, which disappears altogether in pianissimo. His subglottal pressure ($p_{sub}$) reaches the high level of 100 cm of water at the loudest point of his *messa di voce*, but falls rapidly to ca. 15 cm in pianissimo. His effort to achieve a smooth reduction of intensity might well be effective under performance circumstances, but it is not successful in all respects, and careful listening reveals a discontinuity in the sound at the point where the subglottal pressure drops precipitously. The rising slope of the EGG signal at the closing of the glottis, quite steep in forte, is shallower in pianissimo, suggesting that this closing is relatively gradual as well as incomplete. This finding was confirmed stroboscopically.

The lyric tenor, subject TL (Figure 2), shows a relatively small maximal $p_{sub}$, which he reduces smoothly in the approach to pianissimo. In fortissimo his singer's formant is prominent, but not dominant like that of the more robust tenor. The level of the singer's formant is progressively reduced during the diminuendo, but it remains in the contour of the spectrum nearly to the end of the phonation. His EGG signal shows a considerably smaller closed quo-

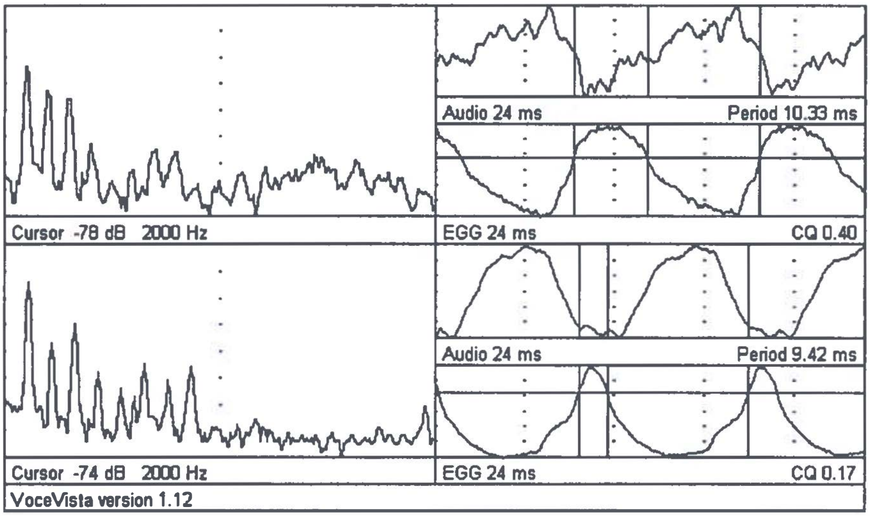


**Figure 3c**

In *pp* the first harmonic (the fundamental frequency) has become dominant in both voices. Although stroboscopy showed a lack of complete closure of the glottis of the robust tenor, his EGG continues to show the pattern of deep contact of the vocal folds, along with a much flatter closing slope. The lyric tenor's closing slope is also flattened, but he still manages full closure (confirmed by stroboscopy) with a particularly small closed quotient, suggesting contact limited to a thin portion of the vocal folds.

tient than that of TR. The CQ's of both diminish as the sound becomes softer (see Figure 3), but in the case of TL the glottal cycle retains a closed moment, even in the softest phonation. This complete closure was also confirmed stroboscopically.

Stroboscopic examination of the maneuver revealed an additional major difference between the two tenors. Both singers noticeably narrowed the opening to the 'laryngeal collar' in producing the forte sound, but this opening was greatly enlarged for the *pp* of subject TR, while it eased only slightly for TL.

## Discussion

Our data are consistent with the finding of Yanagisawa et al (1989) that correlates the constriction of the laryngeal collar with the production of the singer's formant. This constriction appears to be greatest in the forte production of the 'lighter' of the two voices, and it relaxes only slightly in his *dimin-*

*uendo*. In the robust tenor, where the contrast between *ff* and *pp* is most pronounced, there appears to be a categorical difference in the articulation of the laryngeal collar – a marked opening of the constriction at the point of the abrupt move to *pp*. Such an opening of the laryngeal collar was generally found in a study by Dmitriev (1957), which included lateral x-ray examination of a number of Russian male opera singers.

A second categorical difference between the production of sounds at the soft extreme of the intensity range is in the completeness of glottal closure. In their softest phonation both subjects showed a glottis with firmly closed arytenoid cartilages (the posterior portion of the glottis), but in TR a region toward the anterior portion of the glottis remained open in the most-closed phase. Subject TL, by contrast, showed a glottal configuration of neatly parallel vocal folds with an instant, however brief, of closure. It is important to note that the location of the open segment of the glottis was not at the posterior end, as is typically the case with breathy voice (Södersten and Lindestad, 1990), but in the membranous portion. We have observed an opening at this same location in the softest phonations of other professional male singers as well.

The EGG signals of soft phonation also reveal a mode of voice production in TR that differs from both his own forte and from the *pp* production of some other singers, including subject TL. The high closed quotient of TR's fortissimo is reduced in the diminuendo, but even after the incomplete closure has appeared his EGG indicates a relatively long most-closed phase in the glottal cycle (see Figure 3). TL's closed quotient, by contrast, scarcely exceeds 50% when maximal, and can be reduced below 20% while still maintaining a phase of full closure. These differences between the two tenors suggest fundamental differences in structure of voice, the way of handling it, or both.

## *Messa di voce* in 'falsetto' function

In order better to understand the differences between the voices in this study, it is helpful to compare a similar maneuver executed by a professional female singer. Figure 4 gives similar signals to those of our first two figures, but here display a soprano's *messa di voce* on the pitch *B5-flat* (932 Hz). In this case the crescendo and diminuendo are quite smooth, and there appear to be no problems in the continuous variation of subglottal pressure, sound pressure level, and even the closed quotient. A speculative explanation for the relative ease with which a female singer can accomplish the *messa di voce* is the following: Regardless of whether male or female, the 'falsetto' vibratory pattern, in which only the upper margins of the vocal folds participate, while the inferior parts remain stationary (Titze, 1994), facilitates control over fine adduction of the glottis, as compared with the 'chest' vibratory pattern. The augmentation and reduction of glottal resistance (and thus the regulation of Psub) in 'falsetto' can be implemented by (smoothly) varying the magnitude of contact of the vocal folds downward (caudally) from the thin upper margins. The 'chest' vibratory pattern differs essentially from this: because the contact between the inferior portions of the vocal folds is a fundamental feature of the 'chest' pattern, step-less reduction of vocal-fold contact to the small magnitude achieved between the thin upper margins is ordinarily not a possibility.

The subjects in this study could distinguish clearly in their own voices between a 'falsetto' production and a 'chest' vibratory pattern on a soft tone. Subject TL, however, showed the unusual ability of being able to reduce gradually the amplitude of his sound to a point where he could pass smoothly to falsetto. But even in this register his vocal folds were observed to make contact in parallel fashion along the midline. Because of this ability, his diminuendo to pianissimo has a much closer resemblance to the step-less example of the soprano than does that of subject TR. Furthermore, the retention of a closed phase of the glottis in *pp* makes it possible to produce, at a reduced level, the singer's formant, which requires an abrupt cessation of glottal flow to generate high-frequency spectral components.

The diminuendo of subject TL seems to fulfill the frequently quoted requirement of Lamperti (see above) that the pianissimo sound resemble the forte in all respects except intensity. What are we to make of the robust tenor, who fails the pedagogical ideal in at least two ways – an overly abrupt diminuendo and a *pp* sound that differs in character from the forte?

The practical answer to this question is found in the simple fact that subject TR uses the pianissimo production illustrated in the figure on leading operatic stages of the world with impunity. Furthermore, many of his colleagues, tenors singing this same repertory, use a similar soft sound, lacking in the singer's formant component. A more sensible question would be whether TR might achieve the sort of diminuendo prescribed by pedagogues if he were trained differently. The answer to this question is necessarily speculative, but our speculation can be guided by experience in observing and measuring a

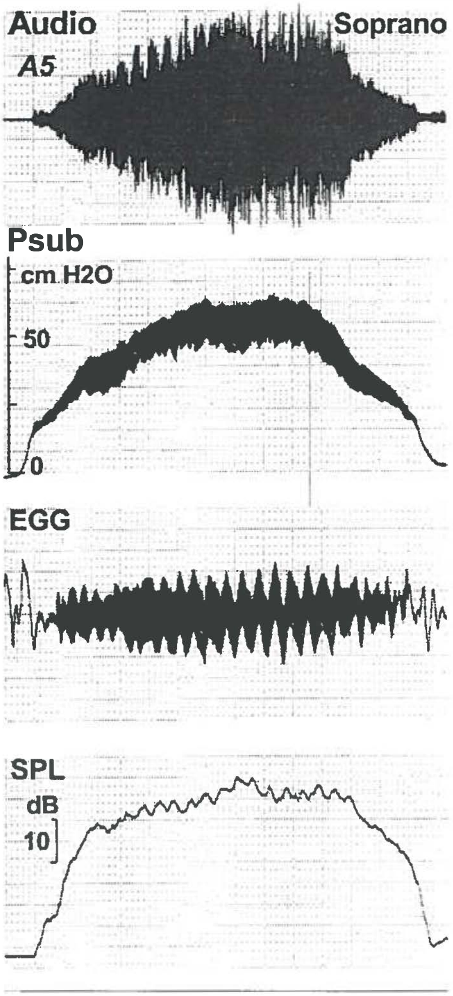


**Figure 4**
*Messa di voce* on *B5-flat* by a lyric soprano. Signals are the same as in Figures 1 and 2. Mean subglottal pressure reaches 55 cm water pressure in *ff*, and falls to ca. 15 cm water pressure in *pp*. The smooth and even rise and fall in Psub and SPL suggests that the maneuver is easier to accomplish in a female voice that in our male subjects.

few similar singing voices. The fact that subject TR can routinely achieve subglottal pressures above 100 cm of water pressure and closed quotients in the range 67-75% suggests that his is a different species of voice than that of TL, who scarcely exceeds 40 cm water pressure and 55% for these parameters. It seems quite plausible that vocal folds that allow such high closed quotients and subglottal pressures may not lend themselves to the stretching and thinning of the contact edges that can be observed in the case of TL. In addition to this factor, it is possible that the configuration of muscular contractions – inclusive of those promoting laryngeal stability – necessary to produce and maintain such high subglottal pressures, does not permit a gradual reduction down to the ca. 15 cm water pressure typical of *pp* voice production. The singer may have no choice but to leap, as stealthily as he can, from a condition of high $p_{sub}$ and CQ to one of low pressure and a glottis with incomplete closure.

The use of such soft production is far less problematical when there is no need to connect directly with the contrasting mode of loud singing, as there is in the *messa di voce*. Subject TR uses this phonation routinely, and it is likely that a similar phonational mode, including low pressure, incomplete glottal closure, and a relaxed laryngeal collar, is used by other robust voices for soft singing. Consider the following response of Luciano Pavarotti, when asked whether *mezza voce* was 'always easy' for him:

> 'To sing "Parigi, O cara" on the stage, it took me one year. The first year I always sang it forte. Singing piano it is very easy for this piece to be flat. Then I said to myself, "I cannot sing piano, why?" One night I said to myself: "Let me pretend to have lost my voice." It was there just like that... incredible, natural pianissimo... same position... same diaphragm.' (Hines, 1982, p. 222.)

Without the help of stroboscopy and electroglottography, translating Pavarotti's statement about pianissimo into a factual description is hindered by the typical discrepancy between what singers actually do and what they say they do. On the other hand, it is not difficult to find places where Pavarotti, like many other tenors who sing the heavier operatic roles, uses soft sounds that lack the singer's formant, much as subject TR's pianissimo in the example shown here. It seems at least plausible that this is the sound of incomplete closure of the glottis, a voice production that Pavarotti has hit upon by 'pretending to have lost his voice.' Of course one cannot be sure of this without more direct evidence. In the light of our measurements of TR and Dmitriev's fluoroscopic observations regarding the laryngeal collar, however, it would seem to be an hypothesis worth exploring.

Subject TL's execution of the *messa di voce* is clearly more in accordance with the traditional prescription, whether one prefers the version of Garcia or that of Lamperti, simply because discontinuity is avoided as the voice production moves from fortissimo to pianissimo. The imperfect version of TR results in what can be regarded as a separate register. This is distinguished both from full voice, in ways already described, and from falsetto, in that the vibratory pattern of the vocal folds maintains the high-amplitude, large-CQ EGG signal, even when the glottal cycle no longer has a completely closed phase.

If we are correct in supposing that at least some robust male voices are by nature incapable of smoothly bridging the gap between *ff* and *pp*, this still need not imply that the type of pianissimo used here by TR is merely a virtue created by necessity. On the contrary, if it is accepted by the listener as a manifestation of the singer's vocal identity (in contrast to falsetto, which is typically rejected in the 'classical' tradition), such voice production offers a solution to the problem of reaching a larger audience with emotions appropriate to intimate space. *Mezza voce* could be the term used to designate such production, whose identifying physical characteristics are incomplete glottal closure in the membranous portion of the glottis, a relaxed laryngeal collar, and subglottal pressures on the order of 15 cm of water pressure. Even singers capable of producing a textbook pianissimo might alternatively use such a *mezza voce* deliberately for expressive purposes.

## Summary and conclusions

Two contrasting tenor voices, one robust and the other lyric, are examined in their execution of a *messa di voce*, particularly the *diminuendo* from the loudest to the softest phonation. The lyric voice, in accordance with traditional prescriptions for the maneuver, is able to make the *diminuendo* smoothly and with only minimal changes in the configuration of the glottis and vocal tract. The robust voice, which shows considerably higher subglottal pressures and closed quotients in forte than does the lyric voice, makes a more abrupt transition to a soft sound characterized by incomplete closure of the membranous glottis and a markedly relaxed collar of the larynx. It is suggested that the two subjects represent fundamentally different types of voice, having different capabilities and limitations. It is further suggested that the combination of incomplete membranous glottal closure and widened laryngeal collar in soft phonation be regarded as distinguishing characteristics of a separate register in the male singing voice, distinct from both falsetto and full voice (in its 'chest' and 'head' manifestations). *Mezza voce* is an appropriate designation for this register.

Considering that only two subjects are investigated in this study, there can hardly be any conclusions drawn about the general classification of voices. What might be conclusions in a study with more subjects are therefore presented as hypotheses. Publication may thus seem premature, but there are good reasons not to delay until more evidence has been gathered. One of these is that singers at this level of excellence are rarely available for complex measurements. Our experience with singers, although limited, suggests that these subjects, apart from their unique capabilities, are also representative of recognized voice types. A second reason is that the measurements require specialized equipment, expertise, and a certain amount of luck to be carried out successfully. The measurement of subglottal pressure in combination with other parameters is therefore unlikely to be accomplished in large numbers, even disregarding the problem of finding an adequate number of qualified singer-subjects. A third reason is that laryngoscopy of singers, although common enough, has not, in the years since Garcia's detailed observations with what must have been fairly primitive equipment, been very productive of useful general information on the singing voice, even with the historical addition of stroboscopy. The great variety of throats, the invasiveness of the procedure, and the fleeting aspect of what is seen (even when videotaped) all contribute to make it exceedingly difficult to build cumulative knowledge regarding critical visual features of the singing voice. Looking for the subtle organic lesions of singers is challenge enough, and familiarity with varied and complex vocal-fold movement patterns of singing is more than can be expected of laryngologists (Sulter et al, 1996). Once a few specific, recognizable features have been identified, however, it may be possible to test the general validity of our proposed description of *mezza voce*.

## Reference List

Dmitriev LB. *Rentgenologicheskoye issledovaniye stroyenia prispossoblenia golossovogo apparata u pevtsov*. Leningrad: Dissertation, 1957.

Hines J. *Great Singers on Great Singing*. New York: Doubleday & Company Inc, 1982.

Södersten M, Lindestad P-Å. Glottal closure and perceived breathiness during phonation in normally speaking subjects. *Journal of Speech and Hearing Research* 1990;33:601-611.

Stark J. Bel Canto: *A History of Vocal Pedagogy*. Toronto: University of Toronto Press, 1999.

Sulter AM, Schutte HK, Miller DG. Standardized laryngeal videostroboscopic rating: differences between untrained and trained male and female subjects, and effects of varying sound intensity, fundamental frequency, and age. *Journal of Voice* 1996;10(2):175-189.

Tembrock G. *Akustische Kommunikation bei Säugetieren: die Stimmen der Säugetiere und ihre Bedeutung*. Darmstadt: Wissenschaftliche Buchgesellschaft, 1996.

Titze IR. *Principles of voice production*. Englewood Cliffs, New Jersey USA: Prentice-Hall, 1994.

Yanagisawa E, Estill JA, Kmucha ST, Leder SB. The contribution of aryepiglottic constriction to 'ringing' voice quality. A videolaryngoscopic study with acoustic analysis. *Journal of Voice* 1989;3(4):342-350.

# Chapter 9

# Measurement of Characteristic Leap Interval between Chest and Falsetto Registers

# Measurement of Characteristic Leap Interval between Chest and Falsetto Registers

Donald G. Miller, Jan G. Švec*, and Harm K. Schutte

A markedly smaller time constant distinguishes a chest-falsetto leap from the more usual execution of a sung interval by muscular adjustments in the length and tension of the vocal folds. The features of such a chest-falsetto leap are examined in detail with respect to F0, peak-to-peak amplitude of the vocal-fold contact area signal (EGG), and the closed quotient. A method is proposed to standardize and quantify this chest-falsetto leap in the characteristic leap interval (CLI), a measure of the separation between the natural registers in a given singing voice. The measure is applied to a varied group of experienced singers. Preliminary results include a suggested dimorphic pattern with respect to sex, with female voices exhibiting smaller CLI's and less individual diversity than male voices.

* Centre for Communication Disorders, Medical Healthcom, s.r.o.: Prague 8, the Czech Republic

## Introduction

In a classical experiment made famous through his film, 'The Vibrating Larynx,' (Van den Berg et al, 1960) Janwillem van den Berg shows how an excised larynx will sustain a slowly rising fundamental frequency as the longitudinal tension on the vocal folds is gradually increased, and then leap abruptly upward to a higher frequency, at which point the slow rise in F0 resumes. The abrupt leap in frequency is convincingly explained in the film as a change in the vibratory pattern of the vocal folds from a chest to a falsetto mode. The laryngeal event in the film has been diagrammed (see Figure 1) by Švec et al (1999).

Contributing to the persuasiveness of the demonstration are such details as these: although the 'voice' sounds strange without a vocal tract, the pitches emitted fall within the normal male range; furthermore, the interval of the leap to falsetto, an octave, seems realistic. The preparation of the larynx is such that the spontaneous leap occurs in much the same way as it might in a living yodler. The difference is of course that a living yodler has a large measure of control over the musical interval covered by the leap, while in the excised larynx there is nothing to determine this other than the biomechanical properties of the vocal folds themselves, as arranged in the experimental setup. It would be informative (if it were only possible) to produce a similar spontaneous leap in a living larynx, without the active control of the subject.

The fundamental distinction between the chest and falsetto vibratory patterns is well known. Distinguished by deeper vertical contact between the vocal folds (Titze, 1994), the chest register has a characteristic vertical phase difference, along with a relatively larger closed quotient. The chest register is also presumed to have higher ratio of thyroarytenoid to cricothyroid activation (Van den Berg, 1968). In the polar contrast commonly used to describe the factor of registration in the singing voice, chest represents the 'heavy' element and falsetto the 'light.' The discontinuity between the two 'natural registers' can be conceived as a distance that varies according to the individual voice.

Forty years ago, when Van den Berg made his film, he was not so much breaking new ground as he was making visible what was already understood. A question that is seldom addressed in the literature, however, is that of the magnitude of such a chest-falsetto leap. The present authors have recently reported on the divergent results of three subjects, two male and one female, making random leaps from chest to falsetto (Švec et al, 1999). There was a considerable spread of values for the intervals attained, but each of the subjects stayed within his or her own particular range of intervals. One clear finding resulting from the exercise was that the interval covered by an undirected chest-falsetto leap is not the same in all persons. There appeared to be a given interval, or range of chest-falsetto leap intervals, that goes with a given set of vocal folds.

In the case of the professional singing voice it is reasonable to suppose that the relationship between the chest and falsetto registers – such factors as the relative robustness of each and the degree of dissimilarity between them – may be important in voice classification (Martienssen-Lohmann, 1981). Expert voice teachers presumably hear such factors and take them into account in assessing a voice, even if these factors are not always explicitly discussed in the literature. Smoothing and disguising register transitions is a widely acknowledged principal goal of singing instruction, and the degree of discontinuity between chest and falsetto in voices that must cross this primary register transition in practice can hardly be irrelevant. It is likely that the magnitude of the interval traversed between the registers can be related to this degree of discontinuity.

The goal of this study is to explore the problem of identifying what may be considered a characteristic leap interval (CLI) for chest-falsetto leaps in a given voice. There are a number of obstacles to be overcome in specifying such an interval. Most of these have to do with the fact that both registers in question have a wide range, and, unlike the excised larynx in Van den Berg's experiment, the living singer can consciously direct a rapid movement from one pitch to a variety of other pitches. The ideal CLI should fulfill the following conditions:

1 A clear change of register should occur between the two pitches.
2 In other respects, the larynx and vocal tract must remain, as much as possible, in the same setting for both pitches.
3 The terminal pitch should not be targeted by the subject, but arise spontaneously from the laryngeal adjustments of the chest phonation, only in the falsetto vibratory mode.
4 In order to be comparable among subjects, the interval should be measured at comparable intensity and a comparable point in the range of the various subjects.
5 The tuning of the vocal tract should be removed as a factor in determining the falsetto frequency (as explained below).

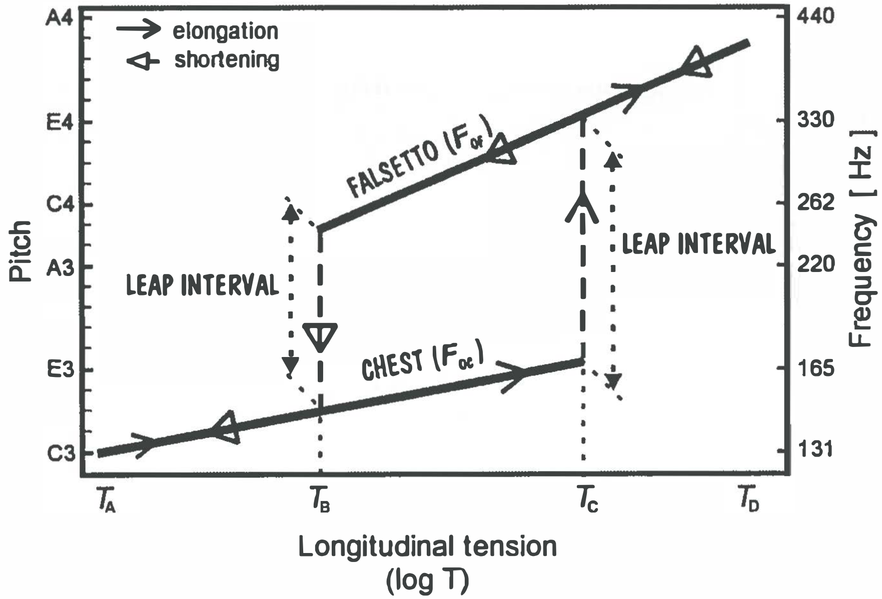


**Figure 1**
Schematic illustration of a chest-falsetto leap in an excised larynx. In the experiment, longitudinal tension is increased along the x-axis, resulting in a gradual rise in F0. At point Tc there is an abrupt upward leap of one octave as the vibratory pattern shifts from 'chest' to 'falsetto', after which the gradual rise in F0 continues. Point Ts is where the abrupt leap back to 'chest' occurs upon reduction of the longitudinal tension.

This list presents a large number of factors to keep under control simultaneously. In order to manage them, the experiment will be divided into a preliminary section and the controlled study itself, with both sections making their contribution to information about the behavior of registers.

## Prologue to the experiment: distinguishing a chest-falsetto leap from pitch changes in an ordinary singing mode

The usual way to change pitch in the singing voice is through adjustments in the intrinsic musculature of the larynx. Such adjustments involve muscular movements that lengthen or shorten structures in the vocal folds and as such require measurable amounts of time. Sundberg has measured the maximum speed at which singers are capable of moving between pitches and found this to be 40-80 ms (Sundberg, 1979). Alipour and Titze have found confirmation of Sundberg's results in an experiment of quite another kind, one in which the

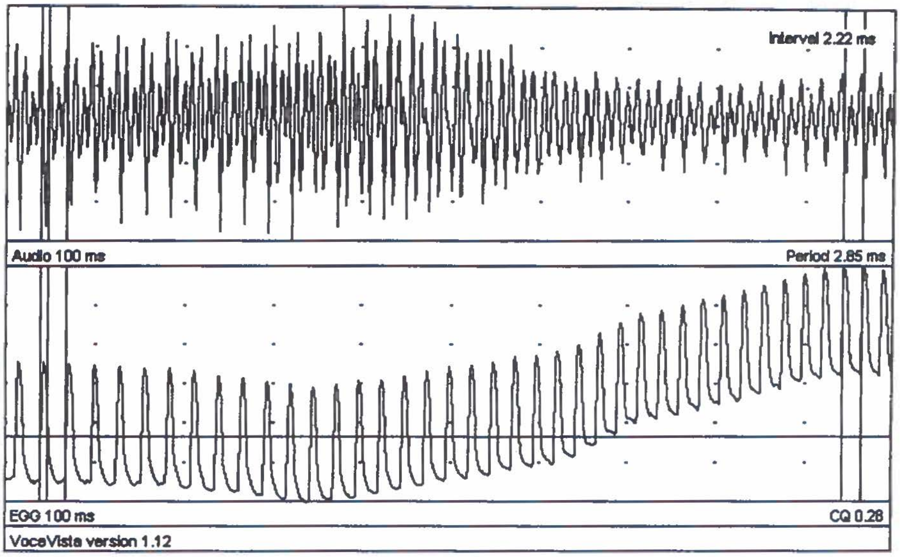


**Figure 2a**

Rapid upward transition from *F4* to *B4-flat* sung by mezzo-soprano (subject F-6). The two panels show 100 ms of time signals of the microphone (audio) and electroglottograph (EGG), with the microphone signal corrected for time delay. From the three vertical cursors on the left, located in the portion of the phonation at the pitch *F4*, come the readings of the 'Period' (2.85 ms, thus 350 Hz) and 'CQ' (closed quotient, at 28%). The cursors on the right are placed at the beginnings of consecutive cycles of the *B4-flat* ('Interval' 2.22 ms, thus 450 Hz).

response time of intrinsic muscles of canine larynges to electrical stimulation was measured (Alipour and Titze, 1999). We informally repeated Sundberg's experiment and ascertained that our subjects accomplished fast pitch changes in times comparable to those he measured (see Figure 2).

In the realm of musical performance the completion of a pitch change within one-half of a vibrato cycle (that is, 70-100 ms, assuming a vibrato frequency of 5-7 Hz) is sufficient to produce a virtually instantaneous move from pitch to pitch. The singer can leave the lower pitch at the bottom of a vibrato cycle and arrive at the upper pitch at the following frequency maximum. Since it is hardly necessary to change F0 even this rapidly in order to achieve the effect of moving smartly between pitches, one can see that executing pitch jumps at still greater (that is, maximum) speed does not ordinarily arise in the practice of singing.

A minute examination of some of the chest-falsetto leaps of the subjects selected for this experiment revealed that these intervals were often accom-

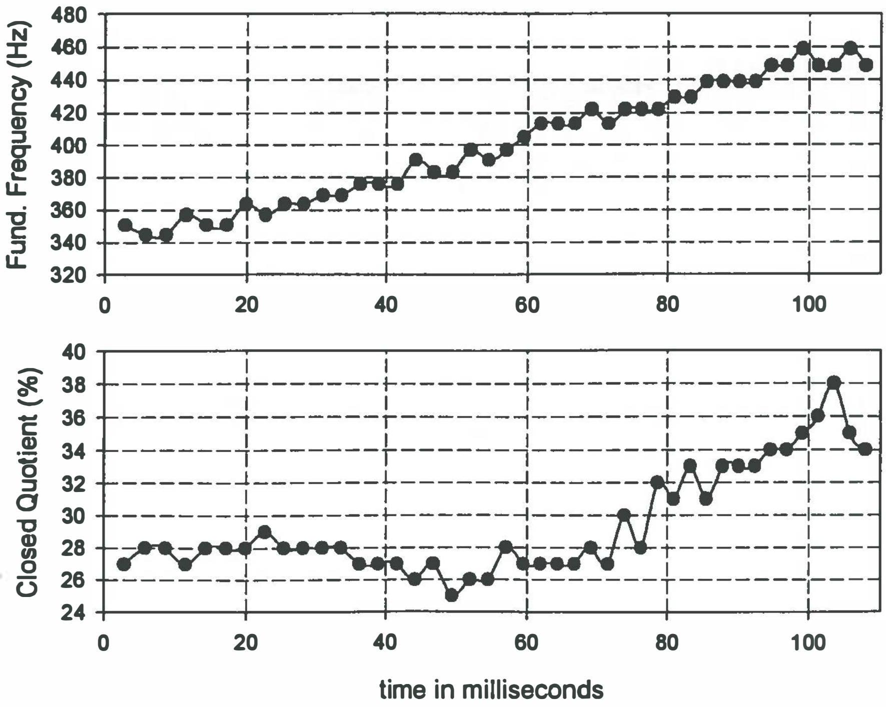


**Figure 2b**
Cycle-by-cycle graphic display of the time segment of Figure 2a, with the upper panel giving F0, and the lower panel the closed quotient. The rise in frequency is evenly distributed over about 70 ms. The closed quotient rises with the transition upward within a single register (here middle register, with a 'falsetto' vibratory pattern).

plished in only a few glottal cycles, or less than 20 ms. Two of these are shown here in Figures 3 and 5. These examples are sufficient to demonstrate that such a leap is qualitatively different from a fast pitch change within a single register, where the increase in F0 is distributed more or less evenly over a large number of glottal cycles, as in Figure 2.

If chest-falsetto leaps can occur without the relatively slow movements of lengthening the vocal folds described by Sundberg (1979), what is then the mechanical basis of the change? Without attempting to answer this question fully, we observe that it is in the EGG signal that we find evidence of a minimal adjustment that can trigger the abrupt leap in frequency. This signal allows us to follow the closed quotient (CQ), the percentage of the glottal cycle in which the vocal folds are in contact, preventing the flow of air through the glottis. In the example presented in Fig. 3 the CQ diminishes for

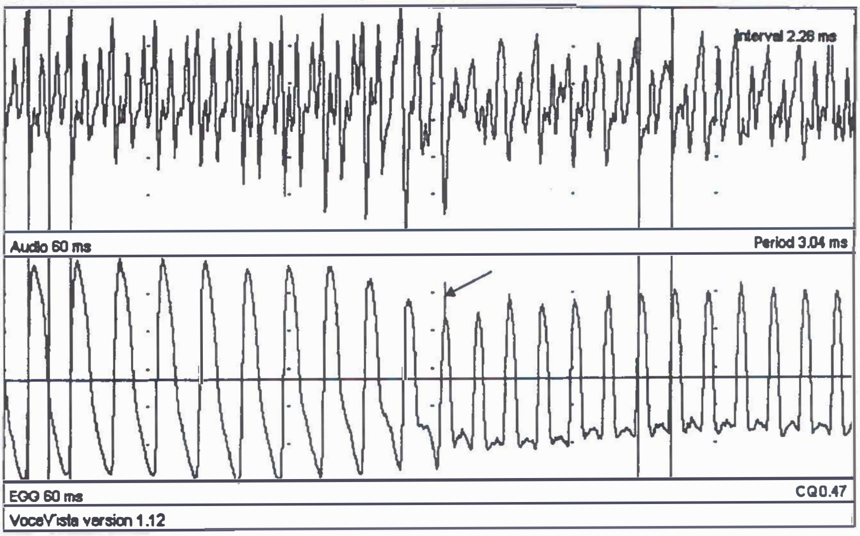


**Figure 3a**
Chest-falsetto leap of a soprano (subject F-3) from *E4* (330 Hz) to *A4* (440 Hz). The two panels show 60 ms of time signals of the microphone (audio) and electroglottograph (EGG), with the microphone signal corrected for time delay. The first 'falsetto' cycle in the EGG is marked with an arrow. The reduced peak-to-peak amplitude of the EGG signal in 'falsetto' is evident. From the three vertical cursors on the left, located in the *E4* portion of the phonation, come the readings of the 'Period' (3.03 ms, thus 330 Hz) and 'CQ' (closed quotient, at 47% in 'chest'). The cursors on the right are placed at the beginnings of consecutive cycles of the *A4* ('Interval' 2.28 ms, thus 439 Hz).

the last six cycles of the chest register portion, while rising only slightly in F0. The vocal folds are evidently abducted up to the point where the deeper vocal-fold contact of the chest register is suddenly lost, and the glottal vibration then rapidly reorganizes on the basis of the reduced mass, and thus higher F0, of the margins of the folds.

Although instances can be found where the falsetto vibratory pattern is established immediately, there is more often a brief delay before the new F0 is stable. One possible consequence of such transient instability is the appearance of subharmonic frequencies, a phenomenon that we have investigated in Švec et al (1996). Another possibility is a short interval of time without a clear F0 pattern. Especially in male voice, the contact between the vocal folds in the voice source is quite weak upon leaving the chest register, and the falsetto F0 pattern, not supported by a tuned vocal tract (Rothenberg, 1988), is not immediately stabilized (see Figure 5). The presence of a brief delay in the

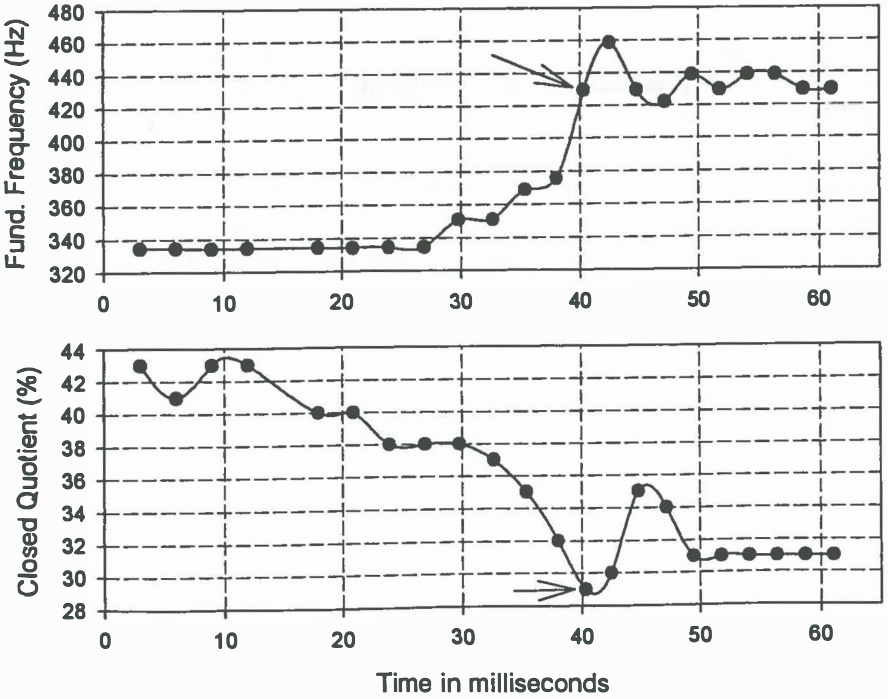


**Figure 3b**
Cycle-by-cycle graphic display of the time segment of Figure 3a, with the upper panel giving F0, and the lower panel the closed quotient. The arrow marks the first 'falsetto' cycle, as in 3a. The abrupt leap is preceded by a slight rise in F0, as well as by a marked reduction in closed quotient. The CQ stabilizes at a lower value in 'falsetto.'

establishment of the falsetto F0, however, will not be regarded as disqualifying a given token from being classified as a chest-falsetto leap.

Before moving on to the experiment itself, two further varieties of the chest-falsetto leap should be mentioned. One of these is a combination of the two basic types, containing both an upward gliding of pitch within at least one of the registers and an abrupt leap across registers. The second variety resembles an actual yodel, in that the falsetto pitch is supported by a tuned vocal tract resonance, constraining the weak falsetto vibration to an F0 'preferred' by the vocal tract.

This influence of the vocal tract in constraining the falsetto pitch clearly interferes with our goal of allowing the structure of the vocal folds to determine the CLI. It is to avoid the yodel pattern of exact and predetermined musical intervals that in the experiment itself we exclude close vowels, whose relatively low first formants exert a strong attraction for F0 in falsetto. This precaution does not prevent two of our female singers from finding a yodel relationship at the fifth and fourth, slightly larger intervals than those that occurred on their involuntary breaks, where there was no evidence of vocal-tract tuning. We presume, however, that it is only within certain limits that the chest-falsetto leap can be constrained by vocal tract tuning. Thus even a tuned interval can still tell us something about the CLI.

A summary of the types of sung intervals encountered in this study includes the following:

1 F0 transitions within a single register, with the increment distributed over a longer (up to 100 ms) time segment (see Figure 2).
2 Abrupt (<20 ms) 'leaps' from one stable F0 in chest register to another in falsetto, with or without a short intervening period of unstable F0 (see Figures 3 and 5).
3 Transitions that combine these types, containing both leaps and more gradual transitions.
4 Leaps in which the falsetto F0 is constrained in part by vocal-tract resonance.

Having arrived at these distinctions by minute analysis of selected examples of sung intervals, we are now ready to proceed further in our search for the characteristic leap intervals of our subjects.

## Method

The experiment was set up with a view of isolating the factor of register (here lumped with the closely connected F0 and airflow) from other variables in the singing voice. Subglottal pressure, which ordinarily rises considerably in singing upward intervals (Rubin et al, 1967), was seen as the key independent variable in maintaining constant conditions across registers. Keeping this factor constant during phonation during the phonation of speech is not ordinarily practiced, but singers do learn to produce steady-state pitch and loudness. Therefore experienced singers – not necessarily professionals – were selected as subjects.

In order indirectly to monitor and register subglottal pressure, the pressure in the esophagus was measured via a transnasal catheter to an esophageal balloon (Schutte, 1980). Other signals registered were those from a microphone 30 cm in front of the subject and an electroglottograph (type Laryngograph). The combination of audio and EGG provided the basis for determining the type, as well as the degree of spontaneity, of the leap. For selected subjects airflow was also recorded.

In order for the measured interval to be comparable among subjects it needed to be normalized. Informal trials showed that both pitch and loudness of the initial tone may affect the magnitude of the interval, with greater loudness and lower initial pitch tending to increase the interval of the leap. The relatively large individual differences found in chest-falsetto leaps, as well as the general male-female F0 difference, make it problematical to identify a frequency and intensity in each voice that can serve as a common referential point of departure for comparison with other voices. In order to achieve some degree of normalization a protocol was devised with the aim of establishing where the chest-falsetto leap was most natural and comfortable for each of the subjects.

## Protocol

The registration of signals was preceded by a short practice session in which the maneuver was explained and the subject's attention directed toward self-monitoring of subglottal pressure. The objective of the practice session was for the subject to find an approximate interval where leaps between stable chest and falsetto phonations could be made at constant subglottal pressure and without changing the vocal tract articulation (the vowel). The minimal changes that induced the leaps were not specified, but it was suggested to the subjects that airflow be increased to the point where the leap occurred spontaneously. When the experimenters judged that maneuver was adequately understood and executed (or alternatively, that further improvements were not to be expected over the short term), the esophageal balloon was introduced, and the registration of signals proper was begun. These began with the subject's counting to fifteen, giving an indication of the pitch of the habitual speaking voice. The maneuver proper consisted of a continuous phonation, beginning and ending in chest register, with two leaps to falsetto and back: C(hest)-F(alsetto)-C-F-C (see Figure 4). Each of the five segments was sustained from one-half to one second. The transitions, accomplished with constant vowel and minimal (input) change in the source, were instantaneous, at least in intent. Instantaneity of the leaps, constancy of subglottal pressure, and stability of F0s were criteria for identifying the most successful tokens.

Most of our subjects had no trouble distinguishing the registers during the maneuver, but occasionally the EGG had to be consulted in female voices. On the basis of the training session, the majority of subjects could approximate constant subglottal pressure adequately by proprioceptive monitoring. Occasionally the experimenters, who could follow the signal of esophageal pressure, intervened to adjust the interval of the leap and improve constancy of the subglottal pressure.

The protocol proper consisted of measurements of nine phonations for the vowel /ɑ/: first leaping from the given (chest) pitch at comfortable intensity, then softer and louder; from a lower pitch at the three intensity levels; from a higher pitch at the three intensity levels. The higher and lower pitches were one to three semitones removed from the first pitch given. (The point here was to find the range where the maneuver worked best, not to explore the whole range where it was possible.) At any point in the protocol the pitch could be adjusted if that appeared to facilitate the execution of the maneuver. Sometimes the pitch and – more often – the intensity would drift, through repetitions, closer to the more comfortable starting point. This was not opposed by the experimenters, since drift toward the comfortable and natural was seen as a move in the direction of normalization. The series of nine phonations was then repeated for the open front vowel /æ/.

Our subjects, who were chosen with a view of obtaining a variety of voice types, both male and female, included both trained and untrained, although all had extensive experience with singing. On this basis their singing voices could be classified with some confidence. All three authors participated as subjects.

## Results

The basic maneuver is complex, requiring both skill from the subject and judgment from the experimenters. Some of our subjects proved to be more adept at the maneuver than others. Let us first look at an example (Figures 4 and 5) of a phonation that was considered successfully executed. In this instance we have the additional registration of airflow, revealing further information about the mechanics of the maneuver.

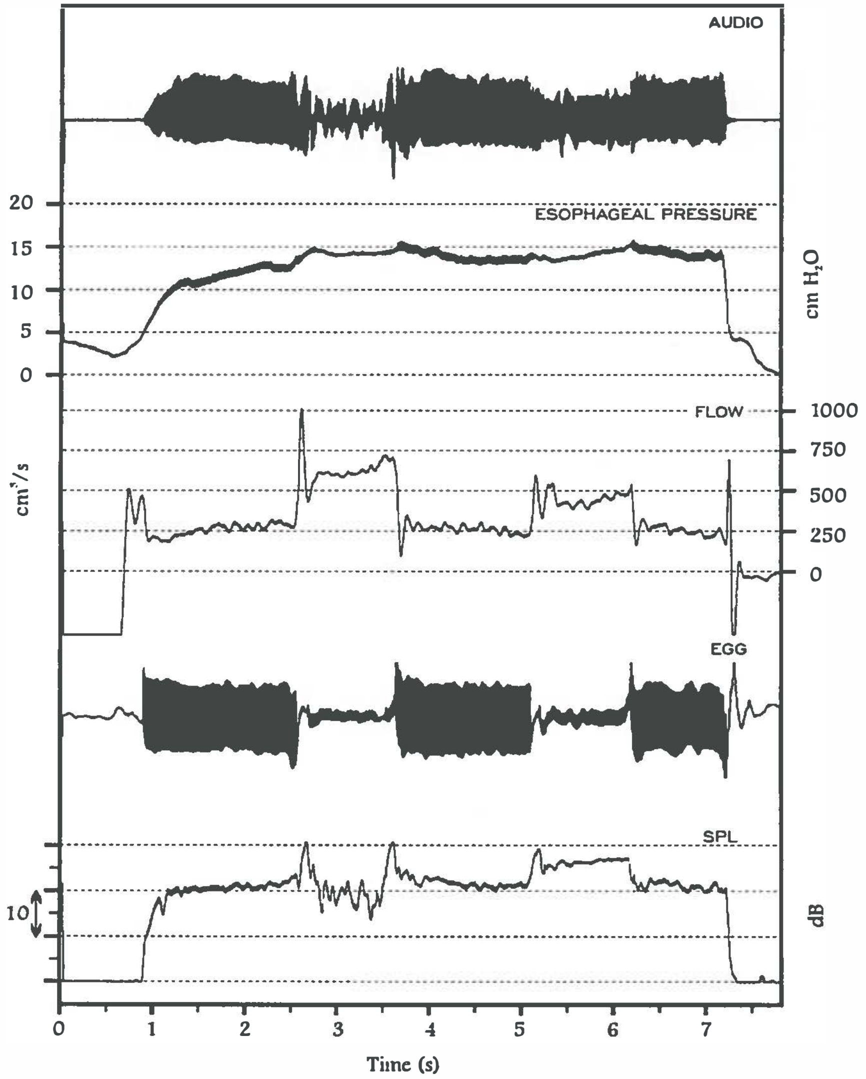


**Figure 4**
Overview of complete phonation consisting of five 'steady-state' segments (C[hest]-F[alsetto]-C-F-C) connected by 'instantaneous' leaps, while maintaining 'constant' subglottal pressure. Subject is M-1, a bass. Signals, in descending order, are: microphone (at 30 cm distance); esophageal pressure (regarded as equivalent to subglottal pressure in this experiment); airflow rate (Lilly pneumotachograph); electroglottograph (EGG), sound pressure level. The factor of register is best followed in the EGG signal, with its sharp contrasts in peak-to-peak amplitude between chest and falsetto. Note the stability of the esophageal pressure, once the level has been established in the first falsetto segment.

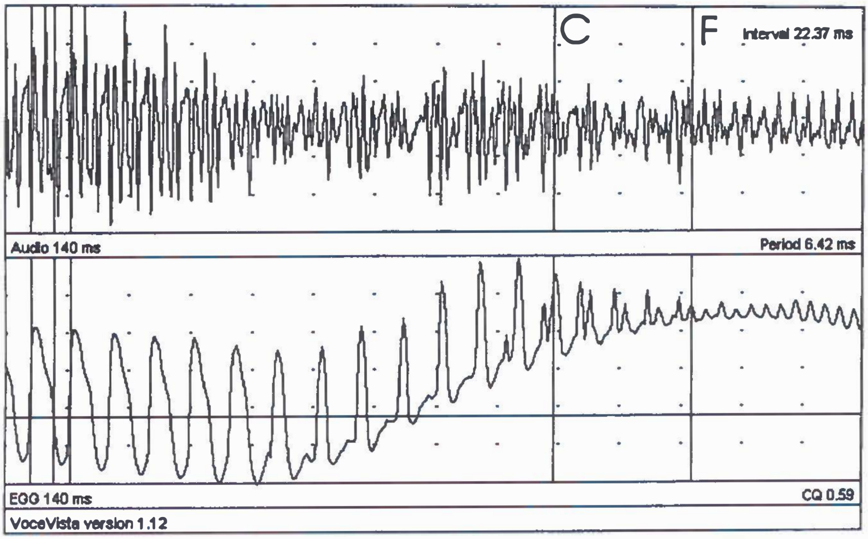


**Figure 5a.**
Detailed time segment of 140 ms of the second chest to falsetto leap in Fig. 4, showing microphone (audio) and EGG signals. Approaching the leap, the closed quotient drops rapidly – while the F0 remains constant – from nearly 60% in the first displayed cycle to less than 25% in the center of the figure. The F0 of the 'chest' vibratory pattern prevails up to the cursor marked C. The falsetto F0 is already established at the point marked by the cursor F, although the contact between the vocal folds is much reduced, with no full closure of the glottis. Beginning in the center of the figure, a second wave appears to arise in the EGG, introducing a 22 ms period of F0 instability in the 16-semitone transition from *E3-flat* (156 Hz) to G4 (392 Hz).

Steadiness of subglottal pressure was our principal control for determining the phonational setting whose minimal change produced the characteristic leap interval. There is a certain settling of this pressure during the initial (chest) segment through the first leap to falsetto. In this case, as with most of the subjects, the second leap to falsetto was considered to reflect most accurately the desired setting. An abrupt, brief surge in the airflow typically accompanies the moment of the leap to falsetto. This movement is reversed on the return to chest register.

Table 1 summarizes the data of our first eleven subjects, including our best estimate of the interval of the characteristic spontaneous leap. Although the sample is not large enough to make any firm generalizations, there is a clear dimorphism in the male and female data. With one exception, the female

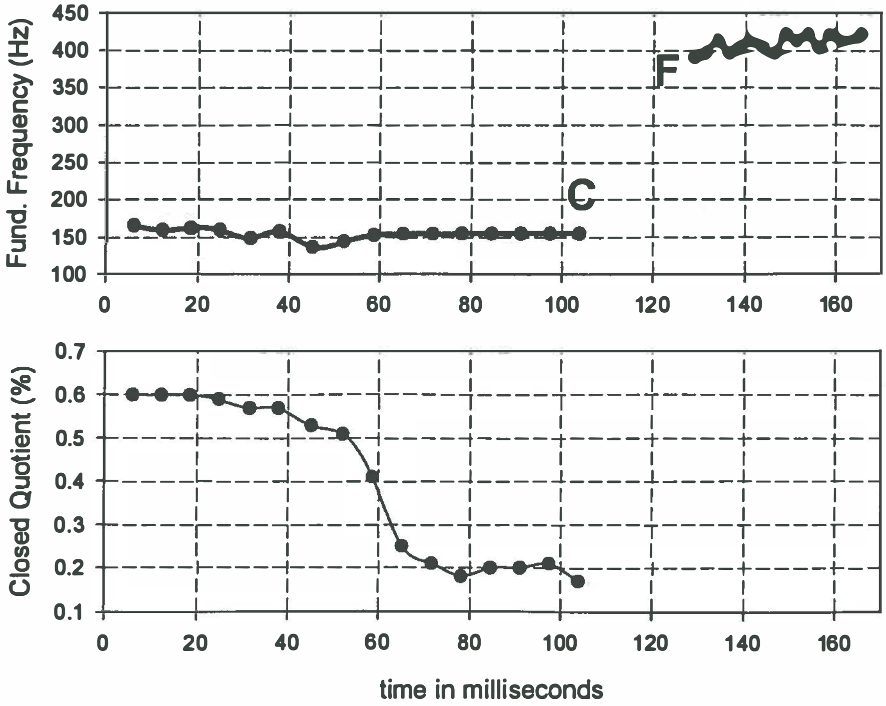


**Figure 5b.**
Cycle-by-cycle graphic display of the time segment of Figure 5a, with the upper panel giving F0, and the lower panel the closed quotient (CQ). The sharp decline in the CQ reflects the loss of contact between the inferior parts of the vocal folds (Titze, 1994). Once the falsetto pattern prevails, there is no closed phase in the glottal cycle; thus no closed quotient is given.

voices have an interval of 7 semitones (a perfect fifth) or less, and three of the six have an interval of 5 semitones. The characteristic interval of the men's voices is on the order of an octave, with bass-type voices having larger intervals than tenor types.

In one case (F-4) the success in executing the maneuver was judged doubtful by the experimenters, and subsequent detailed examination of the data revealed that she was the only subject who failed to produce a chest-falsetto leap that met the standard of the small time constant, as discussed above. The other anomaly was subject M-3, who could not produce the break at the comfortable level of subglottal pressure; it was convincingly spontaneous, however, when it appeared at a higher pitch and intensity.

**Table 1**
Measured intervals selected as characteristic for leaps between chest and falsetto register for six female and five male subjects, all experienced singers. With the exception of subject F-4, who failed to jump across registers, leaps are virtually instantaneous and achieved with minimal adjustments of the voice source between the segments of sustained pitch. Intervals of the leaps are given in semitones, with the chest (initial) and falsetto (upper) pitches identified. Subglottal pressure, measured indirectly by esophageal balloon, is kept (nearly) constant throughout the maneuver.

| subject | voice | interval (st) | pressure ($cmH_2O$) | leap from | leap to |
|---|---|---|---|---|---|
| F-1 | mezzosoprano | 7 | 13 | D4b | A4b |
| F-2 | high mezzosoprano | 6 | 12 | B3b | E4 |
| F-3 | soprano | 3 | 11 | D4 | F4 |
| F-4 | high soprano | 12 [7][1] | 9 | B3 | B4 |
| F-5 | high mezzosoprano | 5 | 9 [12][2] | E4b | A4b |
| F-6 | low mezzosoprano | 5 | 13 | D4b | G4b |
| M-1 | bass | 16 | 14 | E3b | G4 |
| M-2 | bass-baritone | 12 | 14 | D3b | D4b |
| M-3 | tenor | 5 | 25 | D4b | A4b |
| M-4 | tenor | 8 | 15 | B3b | G4 |
| M-5 | tenor | 7 | 10 | A3b | E4b |

[1] This subject generally leapt an exact octave, except from the higher pitch, when she changed to a perfect fifth.

[2] This subject consistently used higher pressures for the falsetto portions; thus pressures for both chest and falsetto are given here.

Not all subjects were equally successful in producing a break with an 'unintended' goal pitch. This task appeared to be more difficult for the female voices, with the result in some cases that the interval recorded was clearly an intentional one. This was accepted if the leap was sufficiently abrupt and the subglottal pressures in the two registers were nearly equal. In two female subjects (F-2 and F-3) a moderate initial instability in the falsetto F0 was taken

**Table 2**
Additional data of the protocol from which Table 1 has been selected, showing the full range of intervals, pitches and pressures measured. The columns under 'Speech' refer to the subjects' counting, a preliminary sampling of the speaking voice.

| Subject | Speech | | Leaps from | | | Interval | Pressure |
|---|---|---|---|---|---|---|---|
| | Pitch | Pressure ($cmH_2O$) | med | low | high | (semitones) | ($cmH_2O$) |
| F-1 | B3b | 6 | D4b | B3b | E4 | 7-8 | 7-16 |
| F-2 | A3 | 8 | B3 | A3 | D4 | 5-7 | 6-17 |
| F-3 | B3b | 7 | D4 | B3 | E4b | 3-6 | 10-15 |
| F-4 | A3 | 7 | B3b | A3 | D4 | 7-14 | 8-11 |
| F-5 | A3 | 8 | E4b | D4b | G4b | 2-7 | 4-21 |
| F-6 | F3 | 3 | D4b | B3b | E4b | 4-8 | 3-14 |
| M-1 | G2 | 8 | D3 | B2 | E3 | 14-19 | 8-21 |
| M-2 | G2 | 12 | D3b | B2b | E3b | 10-16 | 8-27 |
| M-3 | D3 | 5 | D4b | B3b | E4b | 4-9 | 18-31 |
| M-4 | C3 | 5 | B3b | A3b | C4 | 7-9 | 9-20 |
| M-5 | C3 | – | G3b | E3 | A3b | 6-14 | 9-26 |

as evidence that the goal pitch was not predetermined, and therefore more convincingly authentic.
The smaller interval of the female leaps went together with a restricted F0 range where they could be produced at comfortable effort. In the male voices there was a larger range of chest-register pitches from which a comfortable leap into falsetto could be made.

Familiarity with the chest register as a singing mode seemed to help the women in executing the task. The one exception to the smaller leap interval among the females, subject F-4 (who produced both perfect fifths and octaves, but without the small time constant), sang with an easy high soprano, and was unaccustomed to using chest register except in speech.

The male voice (M-3) whose values were farthest removed from the others, an untrained tenor, could produce the break only in a narrow and relatively high range, which required considerably higher pressure and flow than the other subjects. In his case, however, the interval was deemed authentic, since he had no control over the goal pitch.

In most cases there was a tendency for the interval to be larger for a lower initial F0 and for greater effort. This tendency, however, was subject to considerable individual variation.

## Discussion

The findings of the study – the distribution of various values around a male-female dimorphism, as well as exceptions that fall outside these clusters of data points – have implications with respect to the matter of 'registration' as a feature of the structure of a voice. In contrast to the literature of singing pedagogy, discussions of register in the voice science literature usually treat the characteristics of chest and falsetto as a general matter (sometimes with different terminology), paying little, if any, attention to male-female differences (Titze, 1994; Hollien, 1974). If the distance between the registers, as we have formulated the matter, is markedly greater in male voices than in female, this difference can have important consequences for both theory and practice of the singing voice, where the smoothing of register transitions is an item of high priority.

Along with the interval of the spontaneous break, the difference in amplitude of vocal fold contact can help to illustrate the distance between the registers (Figures 3 and 5). If we compare the chest-falsetto leaps of two of our subjects with very clear EGG signals (considered an indicator of vocal fold contact area), we notice that the peak-to-peak amplitude of the signal is radically diminished in the falsetto of the male subject, while in the female subject the change is largely in the *form* of the curve. It is also clear that the reduction in the closed quotient that accompanies the leap is greater in the male voice, where it drops to zero. (The absence of a fully closed phase in the EGG of falsetto was typical for our male subjects.)

This clear male-female difference in the amplitude of the EGG signal at the leaps is in marked contrast to the results of a study by Roubeau et al (1991). These authors measured register breaks occurring on glide tones, and found no significant gender difference between the ratios of amplitude of the EGG signals in the chest-falsetto transition on upward breaks. This difference in results is probably related to the difference in the tasks that produced the breaks. Our (singer) subjects had an assignment that brought to bear more of the skills of singing: finding a point of balance in the voice source which could shift back and forth, using minimal adjustments, between steady-state productions of both registers.

The variation *within* the gender groups with respect to the structure of registration also has important consequences for the management of the singing voice. If the *messa di voce* (swelltone) is considered to go from falsetto to chest to falsetto, as some authorities maintain (García, 1982), then it might be expected that voices with a smaller 'distance between registers' would have a marked advantage in such a maneuver. This might help explain why many voices seem incapable of such a smooth transition in spite of excellent training. Furthermore, the tenor voices measured in this study, characterized by smaller CLI's, seemed to show a predisposition to greater continuity in the chest-falsetto transition. This last observation, however, was not tested objectively in the experiment.

Finally, the finding that some voices, like subject M-3, fall rather far from the typical cluster of values for this indicator of 'registration,' can help us to appreciate the many ways in which voices can vary, and perhaps to avoid misguided efforts to achieve the impossible in training.

A difficulty with the experiment lies in the fact that, to a certain extent, an experienced ear is required of the experimenters in order to judge the authenticity of the leap interval. Since the magnitude of this interval is, in most cases, at least to some degree dependent upon the initial pitch, the fact that the experimenters gave that pitch implies a certain circularity in the experiment. That objection notwithstanding, the narrow range of pitches and intensities where the leaps 'worked best' (Table 1) emerged rather easily in the majority of cases. The quite limited spread of subglottal pressures for the most successful leaps (as compared with the range of pressures tried) is evidence that 'comfortable intensity' is a comparable condition among subjects. To a large extent, the criteria of successful execution of the maneuver (quasi-constant $p_{sub}$, categorical chest-falsetto change, stable signals within a given register) sufficiently narrowed the range of results.

## Conclusions

Our results give evidence that the relationship between the primary registers of the voice source in a given voice reflects biomechanical properties of the vocal folds. This relationship can be characterized, in part, by the magnitude of a measurable frequency ratio between the primary registers. This characteristic leap interval, an index of 'registration,' suggests a dimorphic distribution of male and female voices, as well as individual variation within the gender groups. The implications of this index for the training and employment of a given voice are not explored further here, but it is expected that it will prove useful in the classification of singing voices.

The dimension 'registration,' abstracted from the ('light' and 'heavy') concrete registers found in the overwhelming majority of voices, both male and female, is more difficult to conceive than the equally abstract dimensions frequency and intensity. Nonetheless, it allows, and even invites, a degree of quantification, which can contribute to the objective documentation of the singing voice.

## Reference List

Alipour-Haghighi F, Titze IR. Active and Passive Characteristics of the Canine Cricothyroid Muscles. *Journal of Voice* 1999;13(1):1-10.

García M. *Hints on Singing.* New York: The Joseph Patelson Music House, 1982. Notes: translated from the French by Beata García, new and revised edition of 1894.

Hollien H. On vocal registers. *Journal of Phonetics* 1974;2:125-143.

Martienssen-Lohmann F. *Der wissende Sänger: Gesangslexikon in Skizzen.* 3 ed. Zürich: Atlantis Musikbuch-Verlag, 1981.

Rothenberg M. Acoustic Reinforcement of Vocal Fold Vibratory Behavior in Singing. In: Fujimura O, editor. *Vocal Physiology: Voice Production, Mechanisms and Functions.* New York: Raven Press, 1988:379-389.

Roubeau B, Chevrie-Muller C, Arabia C. Control of Laryngeal Vibration in Register Change. In: Gauffin J, Hammarberg B, editors. *Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms.* San Diego: Singular Publishing Group, 1991: 279-286.

Rubin HJ, LeCover M, Vennard WD. Vocal intensity: subglottic pressure and air flow relationships in singers. *Folia Phoniatrica* 1967;19:393-413.

Schutte HK. *The Efficiency of Voice Production.* Thesis University of Groningen, 1980.

Sundberg J. Maximum speed of pitch changes in singers and untrained subjects. *Journal of Phonetics* 1979;7:71-79.

Švec JG, Schutte HK, Miller DG. A subharmonic vibratory pattern in normal vocal folds. *Journal of Speech and Hearing Research* 1996;39:135-143.

Švec JG, Schutte HK, Miller DG. On pitch jumps between chest and falsetto registers in voice: Data from living and excised human larynges. *Journal of the Acoustical Society of America* 1999;106(3(1)):1523-1531.

Titze IR. *Principles of voice production.* Englewood Cliffs, New Jersey USA: Prentice-Hall, 1994.

Van den Berg Jw. Register Problems. In: Krauss M, editor. *Sound Production in Man*. New York: The New York Academy of Sciences, 1968:129-134.

Van den Berg Jw., Vennard WD, Burger D, Shervanian CC. Voice Production: *The Vibrating Larynx* (Instructional Film). Stichting Film en Wetenschap, 1960.

# Chapter 10

# Comparison of Vocal Tract Formants in Singing and Nonperiodic Phonation

# Comparison of Vocal Tract Formants in Singing and Nonperiodic Phonation

Donald G. Miller, Arend M. Sulter[1], Harm K. Schutte and Rienhart F. Wolf[2]

## Abstract

The skilled use of nonperiodic phonation techniques in combination with spectrum analysis has been proposed by the present authors as a practical method for locating formant frequencies in the singing voice. This study addresses the question of the degree of similarity between sung phonations and their nonperiodic imitations, with respect to both frequency of the first two formants as well as posture of the vocal tract. Using magnetic resonance imaging (MRI), linear predictive coding (LPC) and spectrum analysis, two types of non-periodic phonation (ingressive and vocal fry) are compared with singing phonations to determine the degree of similarity/difference in acoustic and spatial dimensions of the vocal tract when these phonation types are used to approximate the postures of singing. In comparing phonation types, the close similarity in acoustic data in combination with the relative dissimilarity in spatial data indicates that the accurate imitations are not primarily the result of imitating the singing postures, but have instead an aural basis.

---

[1] ENT-Clinic, University Hospital Groningen
[2] Department of Radiology, University Hospital Groningen

## Introduction

It is common for singers trained in the Western operatic tradition to conceive of the vocal instrument as consisting of three basic components: the respiratory apparatus, the oscillating vocal folds, and the vocal tract. The vocal tract, with its highly variable posture, is directly responsible, not only for the text in singing, but also for what is often vaguely described as 'resonance,' an important factor in the strength and quality of a given voice. The effectiveness of this 'resonance' presumably depends upon the configuration of the resonances (the formants, especially F1 and F2) of the vocal tract, such that certain harmonics of the voice source are well amplified. At the relatively low frequencies of speech, the close spacing of the harmonics, which are integral multiples of the fundamental frequency, assures that the formants are never very far (in frequency) from a harmonic; however, at the higher frequencies of singing, especially that of higher pitched voices, the proximity of formants to (widely spaced) harmonics is a prominent, and even a critical, consideration in the singer's 'resonance.'

The same wide spacing of harmonics, which makes the locations of the formants along the frequency spectrum critical, makes it difficult for the singer (and scientist) precisely to determine formant frequencies of sung tones. Furthermore, the problem is aggravated as F0 (and thus the distance between harmonics) increases. An informal method exists, however, for dealing with this problem: if the singer is able to maintain, or reproduce, the specific posture of the vocal tract used for a given combination of vowel and F0, while substituting a nonperiodic sound source, the resultant power spectrum can reveal the formant frequencies with a high degree of precision. Subsequent comparison of this (continuous) spectrum with the (line) spectrum of the sung tone can reveal the relationship between even minute modifications of the vowel and changes in the 'resonance,' interpreted as the total impact of the various formants on proximate harmonics of the voice source. (Although the third, fourth and fifth formants have an important role in forming the 'singer's formant', it is generally one or both of the two lowest formants which carry the major part of the sound pressure of a sung tone. The controlled adjustments in singing of these two formants, which are the focus of this study, have been described as 'formant tuning' [Sundberg, 1977; Miller and Schutte, 1990a; Schutte and Miller, 1986]).

The authors, who have made extensive use of this informal method, reported on it in this journal in 1990 (Miller and Schutte, 1990b). While it continues to have an important place in our applications to singing pedagogy, an obvious question remains regarding the degree of similarity between the vocal tract while singing and the vocal tract while attempting to reproduce the singing vocal-tract posture. Our practical experience with the reproducibility of formant frequencies suggested that such singing and non-singing postures might be highly similar and even, in the best cases, indistinguishable. When an opportunity arose to compare them by means of magnetic resonance imaging, it became possible to test this hypothesis.

Three methods were explored of producing a continuous spectrum from a vocal tract in the posture of singing: vocal fry, ingressive phonation, and a vibrator applied to the exterior of the neck. The first two are techniques of non-periodic phonation. Adequate vocal fry for this purpose can be produced by most subjects with only a few minutes' practice. It needs amplification to achieve comparable intensity with a sung tone. Ingressive phonation, a sort of vocal fry in reverse, usually requires more practice, but has the advantage of producing a sound with an intensity near that of (loud) speech. The use of the vibrator, intended as a control, has the advantage of not requiring a type of phonation differing from that of singing. The authors had not used it extensively because of the weak sound and consequent indistinct spectrum that it generated in some subjects, presumably because the walls of their necks damped the sound.

The obvious disadvantage of all three procedures is that the non-periodic phonation cannot occur or, in the case of the vibrator, be measured (because of its relative weakness) simultaneously with the singing. In each case the posture of the vocal tract used in singing a particular vowel and pitch must be maintained and/or imitated in the alternative procedure.

## Methods

Magnetic resonance imaging (MRI) has both advantages and disadvantages when compared with other techniques for viewing bodily structures not available to the unaided eye. Compared with lateral X-ray photography, which historically has been the source of much of our knowledge of the dimensions and positioning of the less accessible portions of the vocal tract, MRI gives clearer pictures of the soft tissues, presents no potential danger to the subject, and can locate points in three dimensions. Disadvantages of the procedure, apart from its cost, are the relatively long exposure time, high level of noise, and restriction to the supine position of the subject. These disadvantages, however, seemed manageable: the apparatus available delivered pictures with a good resolution in 30 s, a duration within the capability of our singer-subject; phonations for acoustic analysis could be produced sequentially, rather than simultaneously with the MRI procedure; informal comparison of singing and non-periodic phonation in the supine position revealed that the formant frequencies did not differ by more than ~ 5% from those produced in the customary upright position (Sulter et al., 1992).

In order to test the degree of spectral and postural similarity between repetitions of sung phonations, as well as the similarity of phonations using non-periodic techniques (ingressive and vocal fry phonation, henceforth termed collectively 'imitations'), the following procedure was used. The single subject, a bass-baritone with many years of professional singing and voice training experience (one of the authors), was placed in the position to be used for the MR image acquisition. He then recorded (at a microphone distance of 20 cm) a sustained, sung /i/, followed by imitations of that phonation, first in ingressive, then vocal fry, with each preceded by short repetitions of the sung vowel as an aid to auditory recollection. This series was immediately followed by the same procedure for the vowels /u/, /ɑ/ and /ɛ/. Thus, 12 samples of acoustic data were recorded (three phonation types times four vowels). Then the subject, positioned in the scanner with the vocal tract in the magnetic isocenter, repeated the procedure, sustaining each of the twelve phonations, plus a silent simulation of each vowel, for 30 s. Because of the noise generated by MRI, no acoustic recordings were made during image acquisition. The entire procedure was repeated three times ('sessions') in the course of 2.5 h, rounded off with a fourth and final acoustic session.

The choice of vowels – two each of open, close, front, and back – was motivated by the intention to include the basic vowel types. The nominal pitch of the sung phonations was always *D3* (~150 Hz.), chosen because it is comfortable for the subject to sustain (yet sufficiently removed from his habitual speaking pitch to assure a singing mode), and because at that relatively high density of harmonics the inclination to 'tune' the formants is presumably not strong.

## Results and discussion

The measured data of interest for this study are of two basic kinds: spatial data regarding the dimensions of the vocal tract, and acoustic data (frequency domain), from the various phonations. Both can be displayed in figures, although comparing the two kinds on the basis of formant frequency values will require a more strict quantitative treatment.

### *Acoustic data*

Figure 1 presents spectral displays of the recorded phonations, arranged according to vowel. For each vowel a selected vocal fry and an ingressive phonation is arranged vertically over a sung phonation. A vertical line drawn through each of the two lowest formants in the ingressive example allows a comparison with formant positions in the other two phonation types. As a first approximation, it can be seen that the non-periodic phonations give reasonably consistent approximations of the position of F1 and F2 in the sung phonations.

It must be borne in mind, however, that the problems in determining the formant frequencies (as well as the potential yield in 'resonance') of a sung tone grow sharply with increasing F0. In a careful study designed expressly to compare the accuracy of formant frequency estimation by automated linear predictive coding (LPC) with that by skilled spectrum analysis, Monsen and Engebretson (1983) found both methods accurate to within ±60 Hz in the F0 range 100-300, but decreasing greatly in accuracy at 350 Hz and higher. This would imply that highly accurate formant measurements of sung tones cannot necessarily be expected even at our chosen F0 of 150 Hz.

The accuracy with which the ear discriminates differences in formant frequencies appears to be greater than that claimed for the measurement techniques used in the Monsen and Engebretson study (1983). Flanagan (1955) found difference limens of ~ 21 Hz for the first formant and ~ 50 Hz for F2. Translating these values into musical terms, it seems plausible to hypothesize that skilled singers could develop a sensitivity threshold of a semitone difference, at least in ranges where that difference would have noticeable consequences in resonance. For this reason we use this frequency difference (± 6%) as a practical criterion of accuracy in locating F1 and F2. Thus, in the present study, we would consider a repetition or an imitation (with a different phonation type) 'accurate' if the formant frequencies fall within the semitone range. (It will be noted that this sets a higher standard for accuracy than the ±60 Hz of the Monsen and Engebretsen study, which is greater then a semitone in most of the F1 and F2 range.)

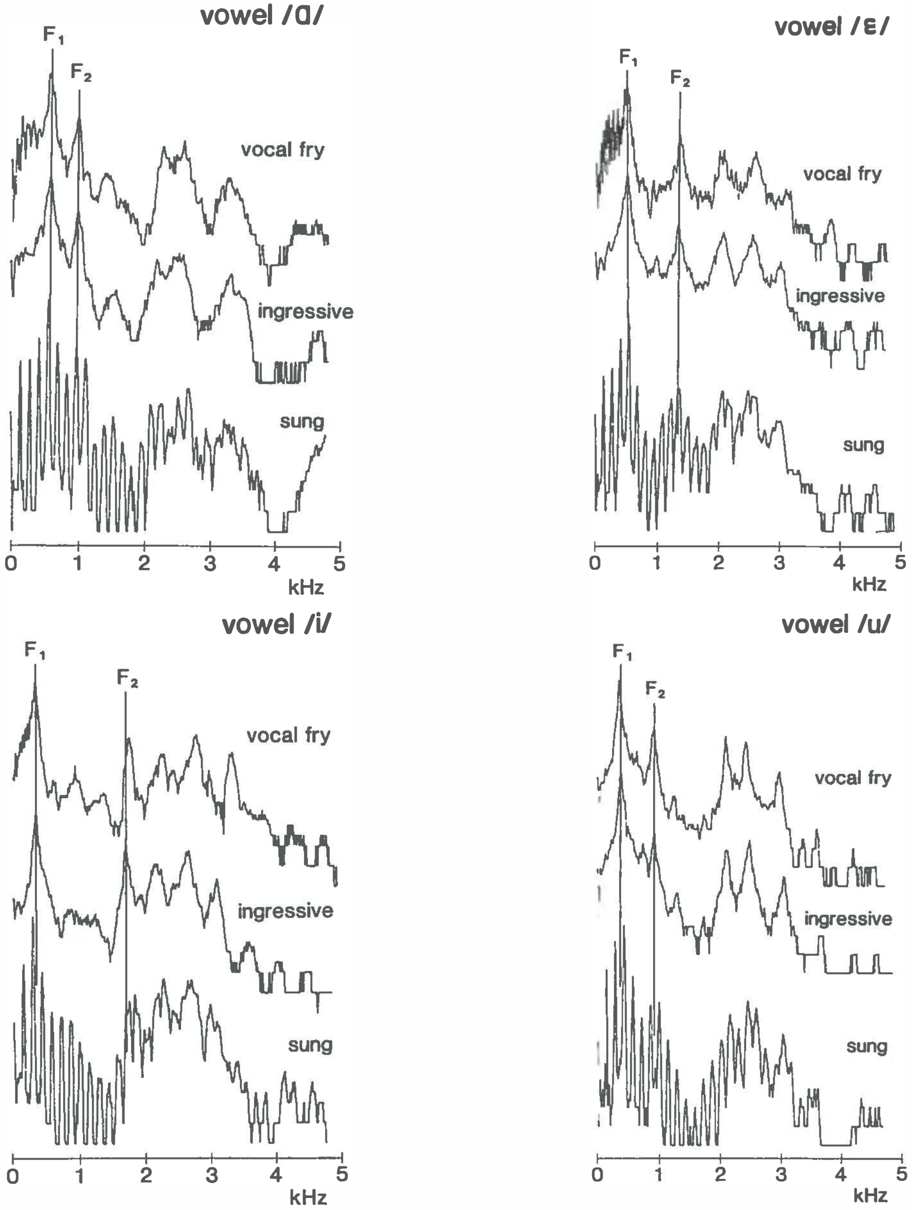


**Figure 1**

Spectral display of sung phonations (nominal pitch D3) of four vowels, shown together with their non-periodic 'imitations.' Vertical lines drawn through the first two formants allows comparison with formant locations in phonations of the three different types.

**Table 1**
Mean frequencies, with standard deviations, of first formants of four vowels, according to phonation type, as measured by linear predictive coding.

| Phonation type | Vowel /i/ | Vowel /u/ | Vowel /ɑ/ | Vowel /ɛ/ |
|---|---|---|---|---|
| Sung | 308 ± 11 Hz | 386 ± 48 Hz | 591 ± 10 Hz | 568 ± 44 Hz |
| | 4% | 12% | 2% | 13% |
| Ingressive | 366 ± 20 Hz | 391 ± 28 Hz | 605 ± 22 Hz | 538 ± 27 Hz |
| | 5% | 7% | 4% | 5% |
| Vocal fry | 358 ± 17 Hz | 394 ± 20 Hz | 597 ± 21 Hz | 530 ± 16 Hz |
| | 5% | 5% | 4% | 3% |

Mean value ± SD, percentage deviation.

For a strict quantitative comparison of formant frequencies in repeated, as well as imitated, phonations, it is necessary to assign values to the formants in sung phonations. We do this with the aid of an LPC program (from the ILS package), while not forgetting the reservations concerning its accuracy mentioned above.

Tables 1 and 2 list the mean formant frequencies for F1 and F2, respectively, of the four vowels and three phonation types, as measured by the LPC program. The SDs give an indication of the scatter of data in repeating the measurements over the four recording sessions. For the nonperiodic phonations these measurements were checked for gross error by locating spectral peaks with a cursor in spectral displays illustrated in Figure 1. Using ± 1 semitone (6%) as a standard of accuracy, among the repeated measurements there are only two cases of (apparent) conspicuous inaccuracy [F1 of sung /u/ (12%) and /ɛ/ (13%] and two marginally inaccurate [/u/ F1 ingressive and /u/ F2 vocal fry (each 7%)]. Excluding the two exceptionally high values, the mean percentage deviation in repetitions comes to 4.1%, well within the narrow standard of accuracy we adopted, and even smaller than the values measured by Flanagan (1955).

Inspection of the spectra of the four vowels suggests an explanation of the two anomalous values. In both /i/ and /ɑ/ the LPC program locates F1 almost

**Table 2**
Mean frequencies, with standard deviations, of second formants of four vowels, according to phonation type, as measured by linear predictive coding.

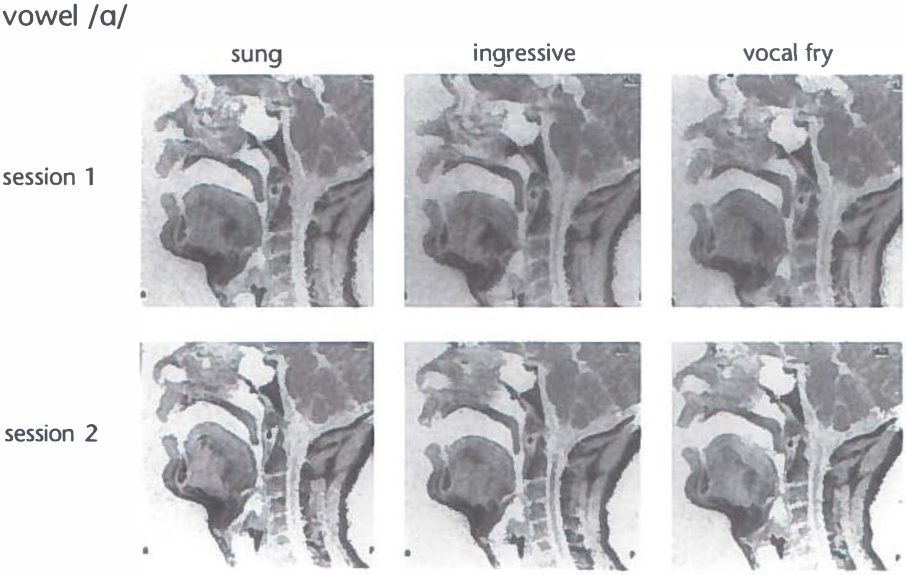
| Phonation type | Vowel /i/ | Vowel /u/ | Vowel /ɑ/ | Vowel /ɛ/ |
|---|---|---|---|---|
| Sung | 1,791 ± 47 Hz | 1,009 ± 43 Hz | 1,091 ± 3310 Hz | 1,446 ± 25 Hz |
| | 3% | 4% | 32% | 2% |
| Ingressive | 1,732 ± 72 Hz | 978 ± 38 Hz | 1,034 ± 59 Hz | 1,431 ± 45 Hz |
| | 4% | 47% | 6% | 3% |
| Vocal fry | 1,758 ± 91 Hz | 894 ± 59 Hz | 1,069 ± 3021 Hz | 1,445 ± 38 Hz |
| | 5% | 7% | 3% | 3% |

Mean value ± SD, percentage deviation.

precisely on a dominant harmonic, flanked symmetrically by weaker harmonics, resulting in exceptionally low standard deviations. In /u/ and /ɛ/, on the other hand, the F1 configuration of harmonics is strongly asymmetrical. It seems more likely that the large error – nearly twice that of any other measurements – is the result of limitations of LPC in measuring harmonic sound (consider the ± 60 Hz error in the Monsen and Engebretson study), rather than exceptional inconsistency on the part of the subject. An alternative explanation, which we find less persuasive, would be that the subject uses the dominant harmonics of F1 in /i/ and /ɑ/ as a guide for precise repetition.

The lower level of accuracy in the repetitions of /u/ was not unexpected. This has been a ‘problem vowel’ for our subject, who reports never having found a consistent, satisfactory pattern for singing the vowel. Reduced accuracy of repetition was also evident in the greater scatter of the MRI data for /u/. F2 of /u/ (where vocal fry has a distinctly lower F2 than the other types) is one of only two places in the data where the spread between phonation types exceeds the 6% standard of accuracy.

The other instance of excessive spread between phonation types presents a more challenging anomaly: the discrepancy between F1 in the sung /i/ and its two non-periodic imitations, which are in close agreement. The difference is on the order of 60 Hz, within the standard of accuracy established in the



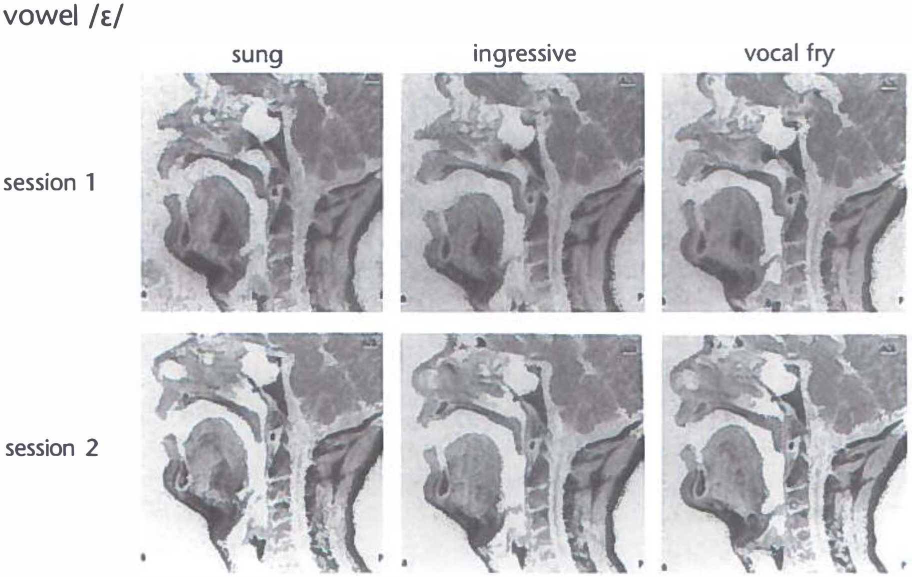
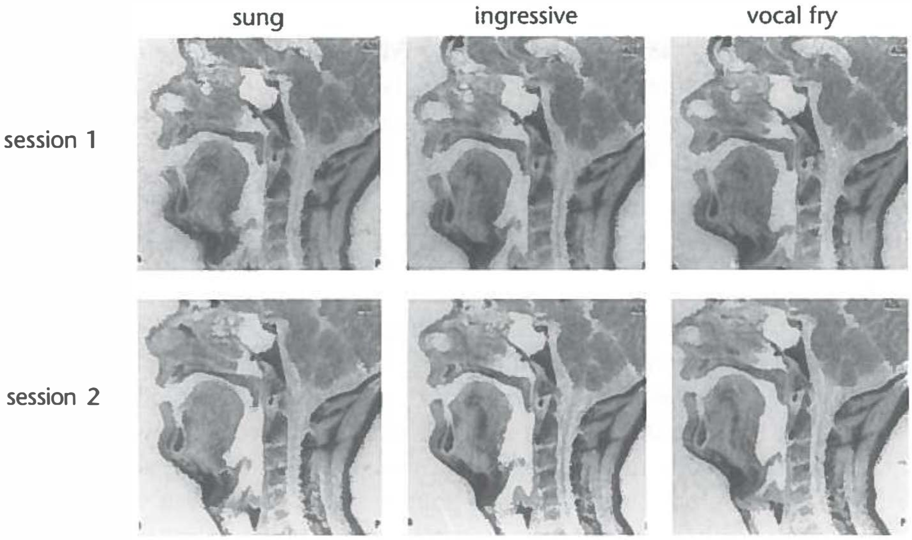


**Figure 2**
Mid-sagittal views taken from 30-s MRI exposures of phonating vocal tract. For each vowel, phonation types can be compared with one another, as well as with similar phonations from a different session (*continues on opposite page*).

vowel /i/



vowel /u/

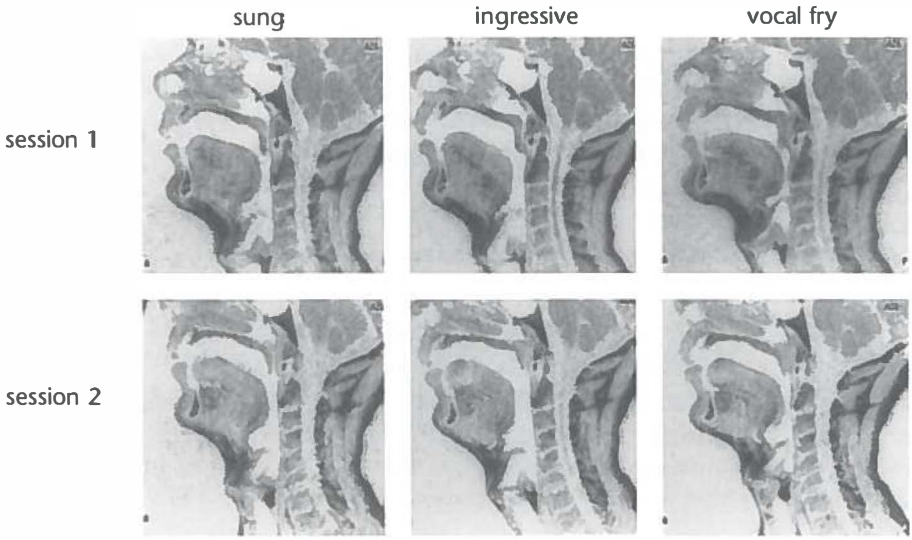


**Figure 2** (*continued*)
Midsaggital views taken from 30-s MRI exposure of phonating vocal tract.

Monsen and Engebretson (1992) study, but in excess of three semitones. Repeated informal trials showed this to be a predictable systematic error on the part of the subject, an error which would have to be taken into account in any application of the (non-periodic imitative) method to determine formant frequencies. We shall return to this discrepancy in our discussion of the acoustic and spatial data taken together.

### *Spatial data*

Figure 2 shows the mid-sagittal slice taken from a 30s exposure of the three phonation types in a given vowel and session, arranged horizontally. Vertically below these are the corresponding results from one of the remaining sessions. Both were selected as those displaying the highest degree of visual clarity. Contrary to our hypothesis concerning the similarity of vocal tract postures for similar formant frequencies, the differences between phonation types are readily distinguishable on visual examination. However limited the resolution presented here, it is apparent that, for a given vowel, similarity is noticeably greater among repetitions of the same phonation type than it is between phonation types.

The calculation of formant frequencies from the spatial dimensions of the vocal tract is an elaborate undertaking. Working with MRI, Baer et. al. published such a study for the 'corner vowels' /u/, /ɑ/, /æ/ and /i/ in 1991. By varying the plane of the imaging, they were able to establish values for three-dimensional slices of the whole length of the vocal tract, thus arriving at measured vocal tract area functions for two subjects, both including and excluding the piriform sinuses.

Adopting a previously developed algorithm, they calculated formant values and compared them with those analysed from acoustic measurements. The discrepancies between the calculated (spatial) and measured (acoustic) values of the first two formants exceeded substantially the standard of precision adopted in the present study. For the vowels /ɑ/, /æ/ and /i/ the mean discrepancy was 10.4% for F1 and 12.5% for F2. For /u/ (which appeared to be a problem vowel in their study as well as the present one) the discrepancies were 19% and 54% for F1 and F2, respectively.

In the present study, where spatial data is limited to (two-dimensional) mid-sagittal MR images, there can be no question of accurately calculating formant frequencies, let alone achieving the 6% standard of precision adopted for the acoustic data. Some key dimensions of the vocal tract, however, can be extracted from the mid-sagittal images for quantitative treatment. We present two of these below, together with the rationale for considering them key dimensions.

The first of these, which also presents the strongest support for the hypothesis that similar formant frequencies reflect similar vocal tract postures among the phonation types, is the dimension vocal tract length. In the case of the vowel /ɛ/, which has no evident tongue constriction, the resonant behavior of the vocal tract ideally resembles that of a pipe closed at one end, and the formant frequencies have a nearly linear relationship to the vocal tract length (Fant, 1970). Table 3 gives the values of this dimension, measured along the midline of the vocal tract, for the phonation types singing, ingressive and vocal fry in all three sessions. In spite of the differences among the phonation types in direction and magnitude of air pressure on the larynx, the variation is only ± 4%, comparable to the high acoustic standard of precision.

The other dimension examined here is that of the tongue constriction, i.e., the distance between the tongue and the opposite side of the vocal tract at its narrowest point in the mid-saggital section. We consider the cases of /i/ and /ɑ/, vowels where the 'back cavity' of the vocal tract is presumed to behave like a so-called Helmholtz resonator (1970). The dimension 'constriction area' occurs in the basic formula for the lowest resonance of such a resonator:

$$F = \frac{c}{2\pi}\sqrt{\frac{S}{l \cdot V}}$$

where $F$ is the resonant frequency, $c$ the speed of sound, $S$ the area of constriction, $l$ the effective length of the constriction, and $V$ the enclosed volume of the resonant cavity.

Since the constriction area has a nearly linear relationship to the square of the constriction width (Baer et. al., 1991), it is apparent from Eq. 1 that a doubling of this dimension could be expected to have a pronounced effect on the resonance frequency. The values for the tongue constriction for all the vowels are displayed graphically in Figure 3. The scatter of values is quite large. This could be due in part to the limit of resolution in the images, which was estimated at ± 1 mm (the width of a single pixel), but that factor cannot account for the major part of the scatter.

**Table 3**
Values for vocal tract length, measured from MR images, arranged by phonation type and session, for the vowel /ɛ/. The differences fall within a 4% range, and there is no systematic variation by phonation type.

| | Session | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Singing | 188.3 | 190.0 | 191.1 |
| Ingressive | 188.8 | 189.4 | 196.8 |
| Vocal fry | 188.2 | 195.8 | 189.5 |

Considering the data for the constriction width as a function of phonation type, it becomes apparent that there is a systematic basis for differences in this dimension: in all four vowels ingressive phonation results in a consistently greater constriction width. A Tukey post hoc test indicates a significant difference ($p < 0.05$) between the ingressive phonation type and both singing and vocal fry.

### *Data considered together*

Comparing the acoustic with the spatial data, one is faced with a puzzle. On the one hand, the subject demonstrates the ability to produce accurate acoustic imitations, with respect to the first two formant frequencies, among various phonation types. On the other hand, the MR images reveal markedly, and in some respects systematically, dissimilar vocal tract articulations among the phonation types. These facts seem sufficient evidence to falsify the hypothesis that the accurate imitations are achieved by maintaining or repeating a specific posture of the vocal tract.

The accurate reproduction of the first two formant frequencies of the vowels appears to have an aural basis. In order to understand the control mechanism involved, it is helpful to distinguish between three types of tuning which the singer engages in. The first, which we designate ‘pitch matching,’ refers to what is meant in saying that someone sings in (or out of) tune: his pitch matches that of accompanying instruments or, more fundamentally, stands in a correct relationship to pitches previously sung. This tuning is achieved by adjusting the repeat pattern (F0) of the voice source.

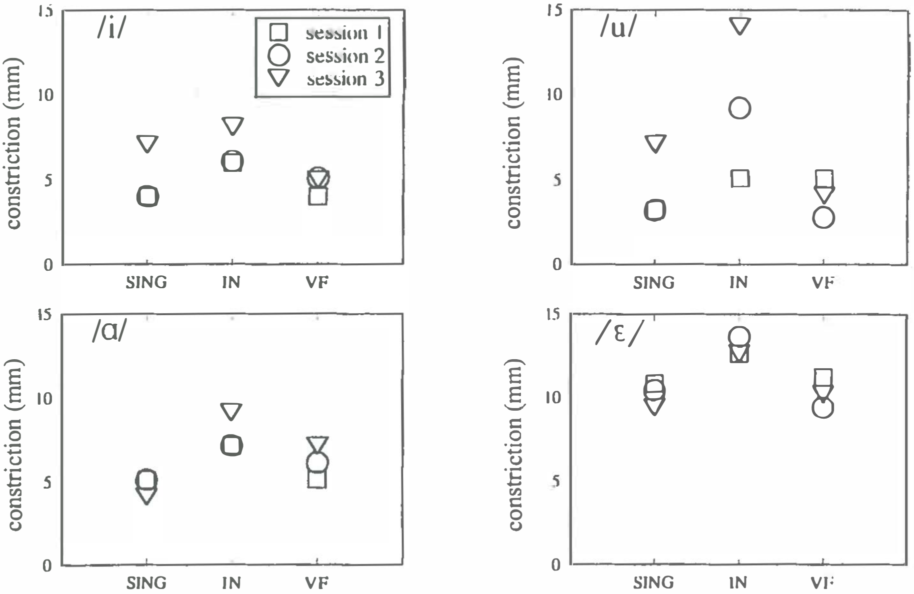


**Figure 3**
Values for the 'tongue constriction' (see text), measured from the three sessions of MR images, arranged by vowel and phonation type (singing, ingressive, and vocal fry). The systematically higher magnitude of this dimension in ingressive phonation is evident.

The second is an adjustment of the vocal tract for the purpose of obtaining the desired 'resonance' from a given vowel and harmonic spacing. Since the harmonic spacing is determined by the prescribed F0, the adjustment is made in the formants of the vocal tract. Within the constraints presented by the requirement that a vowel be recognizable (or, say, have a particular 'color'), the singer 'modifies' the vowel to achieve greater loudness, 'vowel equalization' (a singers' term referring to a desired degree of similarity among the various vowels) or the like. We shall call this adjustment 'resonance tuning,' a term which would include the narrower sense of 'formant tuning,' or adjusting the center frequency of a formant to coincide with that of a harmonic.
The third is also an adjustment of the vocal tract to (re)produce, in detail, the nuances of a given vowel (including those adopted in 'resonance tuning'), but now without respect to harmonic structure of the voice source. We give this the name 'vowel tuning.' In the case where a subject imitates his own sounds at different F0's or with non-periodic phonation, we would expect to find F1 and F2 (the 'vowel formants') as close as possible, in the 'imitation,' to their frequencies in the original sound.

If it is indeed the aural control of such 'vowel tuning' which accurately reproduced the formant frequencies in the imitations of the sung tones, then what can account for the apparent lapse of such control in the case of F1 of the vowel /i/? A speculative, but plausible answer to this question is that the 'vowel tuning' of / i/ is focused on F2 and thus tolerates imprecision in F1. It is conceivable that discrepancies in F2 (especially vocal fry) in the vowel /u/ result from a similar, but opposite, focus of the vowel tuning on the first formant. The data on /u/, however, are clouded by the problem of repeatability mentioned earlier.

The hypothesis offered regarding aural, rather than postural, control of 'vowel tuning' requires for confirmation a broader study than the present one with only one subject. There is one more point in the present study, however, which supports the hypothesis of aural control. In the present study, each time MR images of sung, ingressive, and vocal fry phonations were made, a third imitative type was added: a mute holding of the position of the sung phonation. The intention was to use this as the MRI equivalent of an acoustic recording employing a vibrator applied externally to the neck wall as a 'voice source.' (The strong magnetic field of the apparatus precluded using the vibrator during MR image acquisition.) The 12 images thus acquired often showed vocal tract postures far removed, in various directions, from those made while sustaining vocal sound. In the absence of aural feedback our subject apparently had very little awareness of the posture his vocal tract had assumed.

## Final remarks and conclusions

Using the first two formant frequencies as criteria, the experiment demonstrated a high degree of accuracy (with the exceptions noted) in one singer's repetitions and (non-periodic) imitations of sung tones. The generalizations that we can draw from this, however, are limited. Non-periodic imitations can never give a definitive determination of formant frequencies during singing, but only a more or less accurate indication of where they *might* be. The present measurements only show that, at least under certain conditions, it is possible for a skilled subject to make accurate imitations. Extrapolation of these results to other situations depends upon two basic factors: the skill of the subject and the difficulty of the given imitation. Elements of the skill include control of the voice source and aural sensitivity in the three types of tuning, areas where trained singers might be expected to excel. Regarding the degree of difficulty of the imitation, the vocal tract may assume extreme positions in some parts of the singer's range, and these are not always easily maintained in combination with non-periodic phonation. Non-extreme postures of the vocal tract were used in the present study because of the 30s duration requirement of MRI and in order to provide sufficient density of harmonics for a clear first approximation of the formants in the sung phonation.

With all the caveats, however, non-periodic phonation in conjunction with real time spectrum analysis remains a potentially powerful practical tool in exploring the resonances of the vocal tract, particularly as these relate to varying fundamental frequency. In other words, it may prove to be a significant control for what we have called resonance tuning. Its advantages include the simplicity and quickness of immediately comparing a sung tone with its non-periodic imitation. It is self-correcting, in that the sung spectrum can always be used to check the plausibility of the imitation; thus skill in its use may develop rapidly.

A more thorough assessment of the value of non-periodic phonation in exploring the vocal tract resonances during singing will have to come from a broader study than this. Here we shall restrict ourselves to two conclusions: (a) that it is possible for a skilled practitioner to match, with non-periodic phonation, the first two formants of his own sung phonation with a high degree of accuracy; (b) that the imitative process in such matching of formants does not appear to be based primarily on proprioception of the posture of the vocal tract. And we present an hypothesis: that the control mechanism for adjustment of the vocal tract in this process is primarily aural.

## A pedagogical note on locating formant frequencies

The initial hypothesis – that the similarity of formant configurations between sung tones and their imitations with a non-periodic source implies a comparable similarity in the postures of the vocal tract – was falsified by the experiment. This falsification, however, makes the high level of precision in matching the frequencies of the first two formants seem all the more remarkable. The authors submit that the information obtained in this practice of locating formants can be of considerable importance in singing pedagogy.

In the writings of leading teachers of singing one can find a strong emphasis on the role of the singer's 'ear.' What is meant, of course, is not only the ability to sing in tune ('pitch matching'), but far more, the ability to perceive and to imagine desirable nuances of sound. This constitutes a major part of both the 'talent' of the singer – as aural/vocal imagination – and the task of the teacher, as anticipating and guiding a long-term development toward a sound which suits the individual organ (see, for example, Franziska Martienssen-Lohman, 1981). The acquisition and refined application of this aural imagination occupy the singer's efforts to an extent that would hardly occur to the average layman.

Vowel 'color' and voice 'placement' are concepts from singers' terminology which are frequently used in reference to the sounds singers produce. Skilled singers become adept at reproducing specific combinations of pitch, vowel and intensity with what they consider a proper placement and desired color. We have referred to this process above as 'resonance tuning.' Although the concept of placement can be regarded as an illusion which serves as the singer's control – the voice cannot literally be placed – to the extent that such 'placement' can be heard, it is not illusory, but must necessarily be present in the acoustic signal. Locating the formants is the first step in an objective description of resonance tuning, which, as defined here, includes the singers' 'placement.'

Locating the formants by imitating sung sounds with a non-periodic voice source is not part of the normal practice of singers. In order to accomplish it, singers must increase their aural skills by adding 'vowel tuning' to the 'resonance tuning' which is already part of their practice. Acquiring accuracy and consistency in this method need not be a haphazard process, since it can be guided by spectral comparison of the sung sound and its non-periodic imitation. Implausible results are rejected, and problematical ones are explored further, preferably in the company of others who have already gained experience. In this respect it has much in common with established practices of voice training.

As the use of feedback from spectrum analysis becomes more commonplace, it is to be expected that singers and their teachers will increasingly supplement their basic aural information on resonance tuning with objective measurements of the acoustic signal. When there is a more widespread realization of the critical importance of formant frequencies in, say, the transitions between what singers recognize as registers (Schutte and Miller, 1993; Miller and Schutte, 1993 and 1994), it seems likely that this powerful tool will get more attention, whether through the practice of locating formants by non-periodic phonation or by some other practical application.

## Acknowledgement

This work was supported in part by a grant from the Netherlands Organization for Scientific Research (NWO). We appreciate very much the assistance of Meindert Goslinga for his work on the figures.

## References

Baer, T., Gore, J.C., Gracco, L.C., Nye, P.W. Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels. *Journal of the Acoustic Society of America* 1991:90(2),Part 1, 799-828.

Fant, C.G.M.. *Acoustic Theory of Speech Production.* The Hague/Paris: Mouton & Co. 1970.

Flanagan, J.L.. A difference limen for vowel formant frequency. *Journal of the Acoustic Society of America*, 1955;27:613-7.

Martienssen-Lohman, F. *Der wissende Sänger. Gesangslexikon in Skizzen.* pp 175-8. Zürich: Atlantis Musikbuch-Verlag AG 1981.

Miller, D.G., Schutte, H.K. Formant Tuning in a Professional Baritone. *Journal of Voice* 1990a;4:231-7.

Miller, D.G., Schutte, H.K. Feedback from Spectrum Analysis Applied to the Singing Voice. *Journal of Voice* 1990b;4:329-34.

Miller, D.G., Schutte, H.K. Physical Definition of the 'Flageolet Register'. *Journal of Voice* 1993;7:206-12.

Miller, D.G., Schutte, H.K. Toward a Definition of Male 'Head' Register, 'Passaggio' and 'Cover' in Western Operatic Singing. *Folia Phoniatrica et Logopaedica* 1994;46: 157-70.

Monsen, R.B., Engebretson, A.M. The accuracy of formant frequency measurements: a comparison of spectrographic analysis and linear prediction. *Journal of Speech and Hearing Research*, 1983;26:89-97.

Schutte, H.K., Miller, D.G. The Effect of f0/f1 Coincidence in Soprano High Notes on Pressure at the Glottis. *Journal of Phonetics* 1986;14:385-92.

Schutte, H.K., Miller, D.G. Belting and Pop, Nonclassical Approaches to the Female Middle Voice: Some Preliminary Considerations. *Journal of Voice* 1993;7:142-50.

Sulter, A.M., Miller, D.G., Wolf, R.F., Schutte, H.K., Wit, H.P., Mooyaart, E.L. On the relation between the dimensions and resonance characteristics of the vocal tract: A study with MRI. *Magnetic Resonance Imaging* 1992;10:365-73.

Sundberg, J. The acoustics of the singing voice. *Scientific American* 1977;236:82-92.

# Summary and Conclusions

The ideal in the 'classical' singing voice includes an extended pitch range with a 'perfect scale,' or the ability to move through a series of pitches with no audible discontinuity in voice quality. For a period of history that extends at least back into the 16th century, when virtuosic solo singing first blossomed in Italy, the collision of this ideal with the reality of the evident discontinuity between chest and head, or falsetto, registers has presented one of the greatest challenges to the skill of the artistic singer. At first the mechanism of voice itself was poorly understood, and the registers were referred to the regions of the body where they appeared to resonate: the chest and head. 'Head voice' was also called 'falsetto,' because of its contrast with the 'natural' male speaking voice.

The name falsetto also points to the reason why the registers are regarded as a problem in singing. The *identity* of a voice is an essential aspect of what the listener hears, and a discontinuity that calls into question the integrity of that identity is disruptive. Moreover, the intrusion of falsetto into the identity presented by the male chest voice has an especially disturbing effect as a threat to the maleness of the voice.

The goal of unifying the two registers was famously stated in the 18th century by Tosi and, later, Mancini, two castrato singers from the era when the virtuoso castrato was recognized as the very model of the accomplished singer. In the 19th century the theory of the singing voice underwent a growth spurt. Already in mid-century one of the leading figures of this development, Manuel Garcia, invented the laryngoscope and opened a new era in which the mechanisms of laryngeal vibration that distinguished the registers could be at least partially observed.

The history of the 'problem of registers' did not end, however, with the visualization of the distinct vibratory patterns of chest and falsetto. Other discontinuities that accompanied changes of pitch and intensity remained to compromise the ideal of the perfectly integrated voice. Singers, who did their best to smooth all the bumps for the listener, were more sensitive than their auditors to differences in production between various regions of pitch or intensity. They tended to apply the concept of register to a number of these regions, as well as to certain distinct types of production – for example, '*voix mixte*' – that purported to combine aspects of both chest and falsetto.

It was not until the second half of the 20th century that the technological and theoretical basis was laid for an objective description of what might be called secondary registers: those whose distinguishing characteristics go beyond the dualistic division based on chest and falsetto vibratory patterns of the vocal folds. Such secondary registers as a 'middle register' in the female voice had long been recognized by singers; however, the combination of poor accessibility of the organs, vague and misleading singer jargon, variability of structure and strategies among singers, and the rarity of superb singers made it extremely difficult to establish general, objective truths about the properties of these registers.

The advent of spectrum analysis, together with the closely related acoustic theory of speech, made possible a realistic and detailed description of the role of the resonances of the vocal tract in voice quality. The study of the use of formants in relatively low-F0 speech, however, did not reveal the precise adjustments of these resonances to the more widely separated harmonics of the voice source at the higher F0s of singing. With the discovery of the natural formant-tuning of the soprano upper voice in the 1970's, the stage was finally set for further specification of secondary registers characterized by acoustic, rather than laryngeal, features. Around this same time, another important technological advance for the description of registers arrived in the development of electroglottography, a non-invasive means for tracking vocal-fold vibratory patterns.

There remained two missing pieces of the puzzle. One was the perspective of adepts of singing who could relate the concepts of acoustics and physiology to the extraordinary sounds made by accomplished professionals. The other was a practical method for determining the formant frequencies of the voice in the act of singing.

The 'voice science' that arose in the wake of the acoustic theory of speech had great explanatory power, but one serious limitation with regard to the theory of the singing voice. Voice science had its origins in the efforts at the Bell Labs and elsewhere to compress information from speech for efficient transmission as electronic signals. Within this scheme the vowel 'formants,' identical to the two or three most powerful resonances of the vocal tract, were part of the linguistic information to be compressed. To the singer, on the other hand, the formants are the basis of the power and quality of the voice itself. The linguistic information that they carry is not irrelevant, but often of decidedly secondary importance when compared with the emotional impact of vocal sound.

The other problematic feature of the 'engineers' model' has to do with extrapolation from what is known about the speaking voice. It seems an obvious thing to begin with the organ from which comes the speaking voice, raise the variable fundamental frequency and subglottal pressure, and extend the vowels into steady-state time segments. With the addition of an element of vibrato the model contains the essentials of the singing voice. The functional limits of this mechanism, extended as they are, can be regarded under the aspect of biomechanics, physiology, and vocal hygiene. The problem with this model is that it does not predict the specific features – for example, closed quotients exceeding 70% and second-formant tuning – that turn up repeatedly in the strategies of some highly accomplished singers, while other singers take quite different measures to satisfy an aesthetic ideal. The complex habitual patterns of vocal behavior found in singers of high quality cannot be extrapolated from the speech model.

In addition to these more theoretical problems, a practical obstacle presents itself in the difficulty of obtaining useful information regarding formant frequencies during the act of singing. The importance of this item in the specification of 'acoustic' registers explains the presence of Chapters 3 and 10 in this volume. These chapters are not directly concerned with registers as such, but with a practical means for determining formant frequencies at the higher fundamental frequencies of singing, where automated methods are insufficient. Briefly, it consists of using a non-periodic voice source while 'imitating' the vocal tract of a particular sung sound. The method offered is not foolproof, but it is self-correcting, and the experienced user can apply it with a high degree of confidence that he is getting real information.

Central in this work are Chapters 4 through 9, a series of studies that specify determinative measurable properties of registers that are commonly recognized perceptually by singers. The use of the term register is somewhat arbitrary: it is applied here to a series of similar tones that are distinguished from some other series by a recognizable, although ideally disguised, discontinuity in the voice. The definition is not in itself unusual, but it puts the emphasis on the singer's perception of potential discontinuity (that he attempts to hide from the listener), rather than on the listener's perception.

The discontinuities and registers treated in these individual studies are presented systematically in Chapter 2, which gives an overview of the four most prominent registers in female, and then in male voices. Both sets are divided into a rough structure, according to whether the voice source is 'chest' or ‘falsetto’ (a categorical division that ordinarily occurs in untrained voices of both sexes at the primary register transition, ca. 300-400 Hz) and a finer structure based largely on major changes in the use of resonance. The female registers are the following: ‘chest,’ the low end of the frequency range, characterized by a ‘chest’ voice source; ‘middle,’ the basic singing voice with a ‘falsetto’ voice source; ‘upper,’ a higher extension where the first formant typically matches the fundamental frequency; and ‘flageolet,’ the highest segment, where the fundamental frequency exceeds the first formant. The male registers are these: ‘chest,’ from the lowest notes up to the primary register transition; ‘full head,’ an upward extension maintaining the ‘chest’ voice source; ‘falsetto,’ which mimics the female middle and upper registers with a ‘falsetto’ voice source; and ‘*mezza voce*,’ used here somewhat idiosyncratically as a soft voice with incomplete glottal closure.

Identifying these categories of singing voice production by clear objective features has important advantages for the theory of the singing voice. For one thing, it removes the discussion from the often contentious realm of terminology to that of demonstrable fact: if the phenomena themselves are understood, the names applied to them are less important. Secondly, if elements such as voice source and vocal tract can be separately described, it becomes possible to make specific statements about more complex phenomena: the variety of individuals and types, in terms of both structure and singing strategies, can be dealt with in a differentiated way, rather than assuming, as do many teachers of singing, that all good voice production is essentially similar. The most important advantage, however, is that in this way knowledge concerning the singing voice becomes cumulative. Such objective assertions as are made here can be falsified and corrected, with the result that both theory and practice are advanced.

Wide accessibility to signals from noninvasive methods such as spectrum analysis and electroglottography will soon make the sort of information found in this volume obvious and commonplace, at least among those who combine the practice of singing with the study of the theory of the singing voice. Singing will not necessarily improve as a result, since there is no shortcut through the great amount of learning and growth that goes into forming an excellent singer. We can at least hope, however, that illusion and error will decline as a result of their being easier to detect and expose, and especially that the teaching of singing will be better informed, since the sharing of information is facilitated by the signals.

Fifteen years have passed since the earliest of these articles was written, and it has been more than a decade since the practical method for determining the formants in the singing voice was presented publicly. The VoceVista program, which greatly facilitated the use of this method, has been on hand for the last three years. The accumulation of information that has occurred in the meantime has largely confirmed conclusions that were drawn earlier. One area that I now find ripe for further work, including a certain amount of revision, is the treatment of male passaggio and the related female belting. Not that I have concluded that the assertions presented in those chapters are false, but I have been learning, with the help of colleagues from the singing world, how the practical use of the ‘brightness’ of the second formant resonance can facilitate the move into the upper extension of the range of the 'chest' vibratory pattern. That, however, is material for a future study. Having touched most of the main points of the register schema in the studies included here, this map of the registers is drawn and submitted to see if it will prove useful in understanding the singing voice. The time has come to give it to the practitioners and see what seems to work and what does not, when put into practice by the singers themselves.

D.G.M.
21 February, 2000

# Samenvatting en Conclusies

Het ideaal in het 'klassieke' zingen houdt onder meer in dat er een groot stembereik is met een perfecte toonladder of de mogelijkheid om een reeks tonen te zingen zonder een hoorbare onderbreking of verandering in de stemklank (stemkwaliteit). Gedurende een historisch tijdvak dat op zijn minst teruggaat tot in de 16e eeuw, toen de virtuoze solozang voor het eerst opbloeide in Italië, is de botsing van dit ideaal met de realiteit van een duidelijke breuk tussen borststem en kopstem of falsetstem, één van de grootste uitdagingen geweest voor het kunnen van de professionele zanger. Allereerst werd de werking van de stem zelf slecht begrepen en de registers werden aangeduid met die gebieden in het lichaam waar ze leken te weerklinken: de borst en het hoofd (kop). 'Kopstem' werd ook falsetstem genoemd, vanwege het contrast met de 'natuurlijke' mannelijke spreekstem.

De naam falsetstem geeft ook de reden aan waarom registers als een probleem worden gezien bij het zingen. De *identiteit* van een stem is een essentieel aspect van wat de luisteraar hoort. Een stembreuk die de integriteit van die identiteit ter discussie stelt is storend. Bovendien maakt de falsetstem een inbreuk op de identiteit van de borststem bij de man, wat een bijzonder storend effect heeft door de bedreiging van het mannelijke van de stem.

Het doel om de twee registers tot een eenheid te vormen werd op welbekende wijze geponeerd door Tosi in de 18e eeuw en later door Mancini, twee castraatzangers uit de tijd waarin het virtuoze zingen door castraten gezien werd als het model voor een volleerd zanger. In de 19e eeuw onderging de theorie van de zangstem een groeispurt. Reeds in het midden van deze eeuw vond één van de leidende figuren van deze ontwikkeling, Manuel Garcia, de keelspiegel uit en startte een nieuw tijdperk waarin op zijn minst voor een deel het mechanisme van de laryngeale trillingen voor de verschillen in de registers kon worden geobserveerd.

De geschiedenis van het 'registerprobleem' eindigde echter niet met het zichtbaar maken van de verschillen in trillingspatronen van borststem and falsetstemregisters. Andere onderbrekingen in de klankkwaliteit bij veranderingen in toonhoogte en stemsterte bleven het ideaal van de perfect geïntegreerde stem compromiteren. Zangers, die hun best deden om al de hobbels voor de luisteraars te egaliseren, waren gevoeliger dan hun luisteraars voor verschillen in de stemvorming in verschillende toonhoogte- en geluidssterktegebieden. Ze neigden ertoe het registerconcept toe te passen op een aantal van deze gebieden, of door bepaalde vormen van stemproductie te onderscheiden

– bijvoorbeeld, '*voix mixte*' – door te beweren dat daarin aspecten van zowel borststem, als falsetstem gecombineerd werden.

Het was niet voor de tweede helft van de 20e eeuw dat de technologische en theoretische basis was gelegd voor een objectieve beschrijving van wat secondaire registers genoemd zouden kunnen worden: die waarvan de kenmerkende verschillen verder gaan dan de dualische verdeling gebaseerd op de borststem- en falsetstemtrillingspatronen van de stemplooien. Dergelijke secondaire registers werden al geruime tijd herkend door zangers, maar de combinatie van de moeilijke bereikbaarheid van de stemorganen, onduidelijk en misleidend zangersjargon, variatie in structuur en gebruikte strategieën tussen zangers en het zelden voorkomen van excellente proefpersonen, maakten het extreem moeilijk om algemene objectieve waarheden vast te stellen over de eigenschappen van registers.

De opkomst van spectraalanalyse, samen met de nauw verwante akoestische theorie met betrekking tot spraak, maakten een realistische en gedetailleerde beschrijving mogelijk van de rol van de resonanties in het aanzetstuk op de stemkwaliteit. Het nauwkeurige, hoewel nog intuïtief, afstemmen van de resonanties door de zanger op de harmonischen van de stembron werd niet meteen ontdekt. Dit kwam doordat het relatief laagfrequente spreken werd bestudeerd, echter, na het ontdekken van de natuurlijke formantafstemming bij een sopraanstem in de jaren '70, werd een verdergaande specificatie van de secondaire registers mogelijk. Hierdoor werden ze gekarakteriseerd op akoestische kenmerken en niet alleen op laryngeale kenmerken. Ongeveer in dezelfde tijd kwam een andere belangrijke technologische vooruitgang beschikbaar door de ontwikkeling van electroglottografie, een middel om non-invasief het stemplooitrillingspatroon te volgen.

Twee stukjes van de puzzel bleken nog te ontbreken. Eén was het perspectief van zangadepten, die de akoestische concepten en de fysiologie in verband konden brengen met de buitengewone geluiden, zoals die gemaakt worden door volleerde professionele zangers. De andere was een praktische methode om de formantfrequenties te bepalen van de stem tijdens het werkelijke zingen.

De 'voice science' die ontsprong in het kielzog van de akoestische theorie van spraak had een grote uitleggend vermogen, maar had een ernstige beperking met betrekking tot de theorie van de zangstem. Voice science had zijn oorsprong in de inspanningen in, onder andere, het Amerikaanse Bell Labs om de informatie uit spraak te comprimeren voor een efficiënte overdracht als electronische signalen. In dit schema waren de klinker-'formanten', als de twee of drie meest krachtige resonanties van het aanzetstuk, onderdeel van de linguistische informatie die samengevat moest worden. Voor de zanger, aan de andere kant, zijn de formanten de basis voor de kracht en de kwaliteit van de stem zelf. De linguistische informatie die ze dragen is niet irrelevant, maar vaak van beslissend secondair belang vergelijken met de emotionele invloed van stemgeluid.

Het andere problematische kenmerk van het 'ingenieurs-model' heeft te maken met de extrapolatie van wat bekend is over de spreekstem. Het lijkt een duidelijke zaak te beginnen met het orgaan waar de spreekstem vandaan komt, verhoog de variabele grondtoonhoogte en de subglottische druk, en verleng de klinkers tot langere stabiele tijdssegmenten. Met het toevoegen van een element vibrato bevat het model de essentiele onderdelen van de zangstem. De functionele grenzen van dit mechanisme, die hierdoor verlegd zijn, kunnen beschouwd worden vanuit een biomechanisch, fysiologisch en stemhygiënisch standpunt. Het probleem met dit model is dat het niet in staat is specifieke verschijnselen te voorspellen - bijvoorbeeld gesloten quotiënten hoger dan 70% en afstemming met de tweede formant - verschijnselen die herhaaldelijk optreden in de strategieën van sommige zeer bekwame zangers, terwijl andere zangers heel andere maatregelen nemen om aan een esthetisch ideaal te voldoen. Het complexe aangeleerde patroon van stemgedrag dat wordt vastgesteld bij bijzonder goede zangers kan niet geextrapoleerd worden van het spraakmodel.

In aanvulling op deze meer theoretische problemen is er een praktisch obstakel om bruikbare informatie te verkrijgen over formantfrequenties tijdens het zingen. Het belang hiervan voor het specificeren van 'akoestische' registers verklaart de aanwezigheid van de hoofdstukken 3 en 10 in deze dissertatie. Deze hoofdstukken houden zich niet direkt bezig met de registers als zodanig, maar met een praktische methode om formantfrequenties te bepalen bij de hogere grondtoonhoogten bij het zingen, waar automatische methoden falen. Kort gezegd bestaat het uit het gebruiken van een niet-periodieke stembron, terwijl de stand van het aanzetstuk wordt nagebootst van een bepaalde gezongen toon. De gepresenteerde methode is niet onfeilbaar, maar is zelfcorrigerend, en de ervaren gebruiker kan hem toepassen in een hoge mate van vertrouwen dat hij echte informatie verkrijgt.

Centraal in het hier gepresenteerde werk staan de hoofdstukken 4 tot en met 9, een reeks studies die bepaalde meetbare eigenschappen van registers beschrijven zoals die gewoonlijk perceptueel herkend worden door zangers. Het gebruik van de term registers is wat arbitrair: hier wordt hij gebruikt als een reeks van op elkaar gelijkende tonen die van andere reeksen verschilt door een herkenbare, maar gewoonlijk vermomde breuk in de stem. Deze definitie is op zichzelf niet ongewoon, maar het legt de nadruk op de perceptie van een mogelijke stembreuk door de zanger zelf, in plaats van op de perceptie door de luisteraar, een breuk die hij probeert te verbergen voor de luisteraar.

De onderbrekingen en de registers zoals die in de aparte studies worden behandeld, worden systematisch gepresenteerd in hoofdstuk 2, dat een overzicht geeft van de vier meest in het oog springende registers bij de vrouwenstem en daarna bij de mannenstem. Beide sets worden verdeeld in een grove structuur, samenhangend met het feit of de stembron 'borststem' of 'falsetstem' is (een categorische verdeling die gewoonlijk optreedt bij ongetrainde stemmen bij beide sexen op de eerste registerovergang, ongeveer bij 300-400 Hz), en een fijnere structuur die grotendeels gebaseerd is op grote veranderingen in het gebruik van resonanties.
De vrouwenstemregisters zijn de volgende: 'borststem', het lage eind van het frequentiebereik, gekarakteriseerd door een 'borststem'-stembron; 'middenstem', de gewone zangstem met een 'falsetstem'-stembron; 'hoge register', een hogere uitbreiding waarbij een typisch kenmerk is dat de eerste formant samenvalt met de grondtoon, en 'flageolet', het hoogste segment, waar de grondtoon hoger is dan de eerste formant.
De mannenstemregisters zijn: 'borststem', van de laagste tonen naar boven tot de eerste registerovergang; 'gedekt modaal', een uitbreiding naar boven waarbij de 'borststem'-stembron behouden blijft, 'falsetstem', wat lijkt op het vrouwelijke midden- en hoge register met een 'falsetstem'-stembron, en *mezza voce*, hier idiosyncratisch gebruikt als een zachte stem met een incomplete glottissluiting.

Het identificeren van deze categorieën van de zangstemvorming met duidelijk objectieve kenmerken heeft belangrijke voordelen voor de theorie van de zangstem. Enerzijds, het verplaatst de discussie van het vaak twistrijke gebied van terminologie naar dat van aan te tonen feiten; als de fenomenen zelf zijn begrepen, zijn de namen die er aan gegeven worden van minder belang. Anderzijds, als elementen zoals de stembron en stemweg (aanzetstuk) apart kunnen worden beschreven, wordt het mogelijk om specifieke uitspraken te doen over meer complexe fenomenen: met de variatie in individuen en soorten, in termen van zowel structuur als zangstrategieën, kan worden omgegaan op een gedifferentieerde manier in plaats van aan te nemen, zoals veel zangpedagogen doen, dat alle goede stemvorming in essentie hetzelfde is. Het meeste belangrijke voordeel is echter dat op deze manier kennis over de zangstem cumulatief wordt. Objectieve vaststellingen, zoals hier worden gedaan, kunnen worden gefalsifieerd en gecorrigeerd, met als resultaat dat zowel de theorie, als de praktijk vooruit gaan.

Brede beschikbaarheid van signalen van niet-invasieve methoden zoals spectraalanalyse en electroglottografie zal spoedig het soort informatie als in dit werk beschreven overgaan in duidelijke en algemeen bekende zaken, tenminste door hen die de praktijk van het zingen combineren met het bestuderen van de theorie ervan. Het zingen zelf zal als resultaat niet noodzakelijkerwijs verbeteren, want er is geen snelle route voor de grote hoeveelheid van leren en groeien die nodig is voor het vormen van een uitstekend zanger. Echter, we kunnen op zijn minst hopen dat illusie en fouten zullen afnemen als een gevolg van het feit dat ze gemakkelijker ontdekt en beschreven kunnen worden, en vooral dat de zangleraren beter geinformeerd zijn, omdat het uitwisselen van informatie vergemakkelijkt wordt door de signalen.

Vijftien jaar zijn voorbij gegaan sinds het eerste van deze artikelen werd geschreven, en het is meer dan een decade sinds de praktische methode voor het bepalen van de formanten bij de zangstem werd gepubliceerd. Het programma VoceVista dat het gebruik van deze methode enorm vergemakkelijkt is sinds drie jaar beschikbaar. Het verzamelen van informatie dat ondertussen werd gedaan, heeft grotendeels de conclusies bevestigd die eerder werden getrokken. Een gebied dat ik nu rijp vind voor verdere studie, inclusief een bepaalde hoeveelheid revisie, is de behandeling van de *passaggio* bij de man en het verwante vrouwelijke *belting*. Niet dat ik heb vastgesteld dat de beweringen die gepresenteerd werden in de hoofdstukken onjuist waren, maar ik heb bijgeleerd, met behulp van collega's uit de zangwereld, hoe het praktisch gebruik van 'helderheid' door de resonanties van de tweede formant het overgaan naar de bovenste uitbreiding van het 'borststem'-trillingspatroon kan vergemakkelijken. Dat is echter materiaal voor toekomstig onderzoek. Nu de meeste van de belangrijke aspecten van de registerschema's besproken zijn in de hier gegeven studies, is de kaart van registers getekend en hier gepresenteerd om te zien of het zijn bruikbaarheid kan bewijzen voor het begrijpen van de zangstem. De tijd is gekomen om het aan de mensen van de praktijk te geven en te zien wat werkt en wat niet als het in praktijk wordt gebracht door de zangers zelf .

D.G. Miller,
21 februari 2000

# Appendix

# Glossary of Terms and Basic Concepts

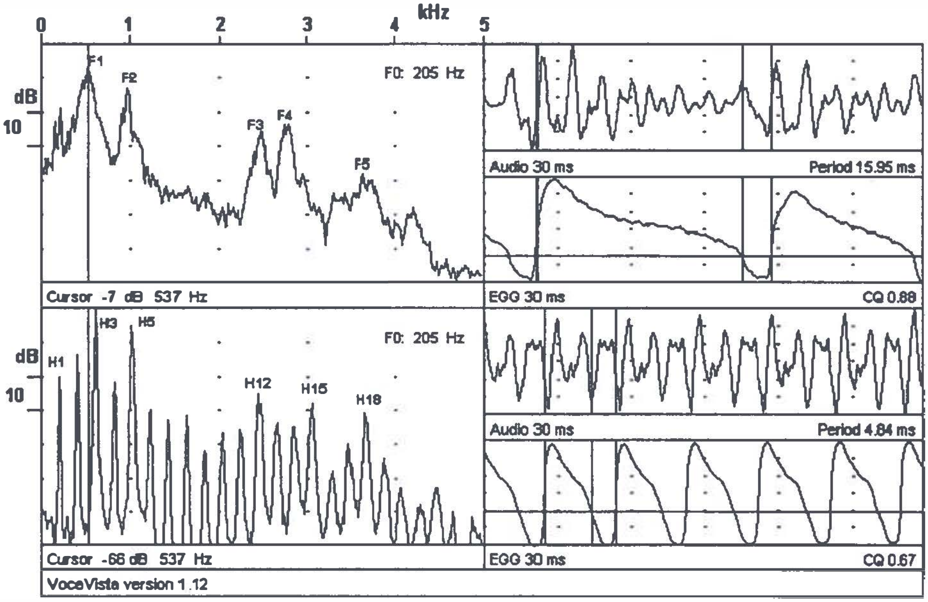
1 Basic figure of VoceVista

2 Acoustic phonetics

3 Glossary of terms

## Appendix 1 – Basic figure of VoceVista

The figure shows a full display of two separate phonations of the vowel /ɑ/ by a bass, arranged for direct comparison. Considering first the left panels, **frequency-domain** displays showing spectra from zero to 5000 Hz, there is a sung *A3-flat* (nominal F0, 208 Hz) below and its 'imitation' in vocal fry above. The (**periodic**) sung phonation results in a **line spectrum**, with substantial sound only at the **harmonics** (all whole-number multiples of the **fundamental frequency**, 205 Hz, as displayed). The harmonics form the familiar **harmonic series** (octave, perfect fifth, perfect fourth, etc.), here in a **linear** display. Vocal fry, a non-periodic **voice source**, produces a **continuous spectrum**, revealing the **formant** (resonance) frequencies more precisely (here labeled F1, F2, etc.) of the vocal tract, independently of the harmonics of the voice source. The vertical cursor is placed at F1 in the vocal-fry spectrum and extends through both spectra, revealing that the first formant, measured at 537 Hz by the cursor, falls approximately 70 Hz below the third harmonic of the sung tone. It is also apparent that the high levels of harmonics H5, H12, H15 and H18 are due to the formants F2, F3, F4 and F5, respectively. The dotted vertical lines are at 1000 Hz intervals, while the dotted horizontal lines are at intervals of 10 dB. The dB reading of the cursor is given for the point of its intersection with the spectrum signal.



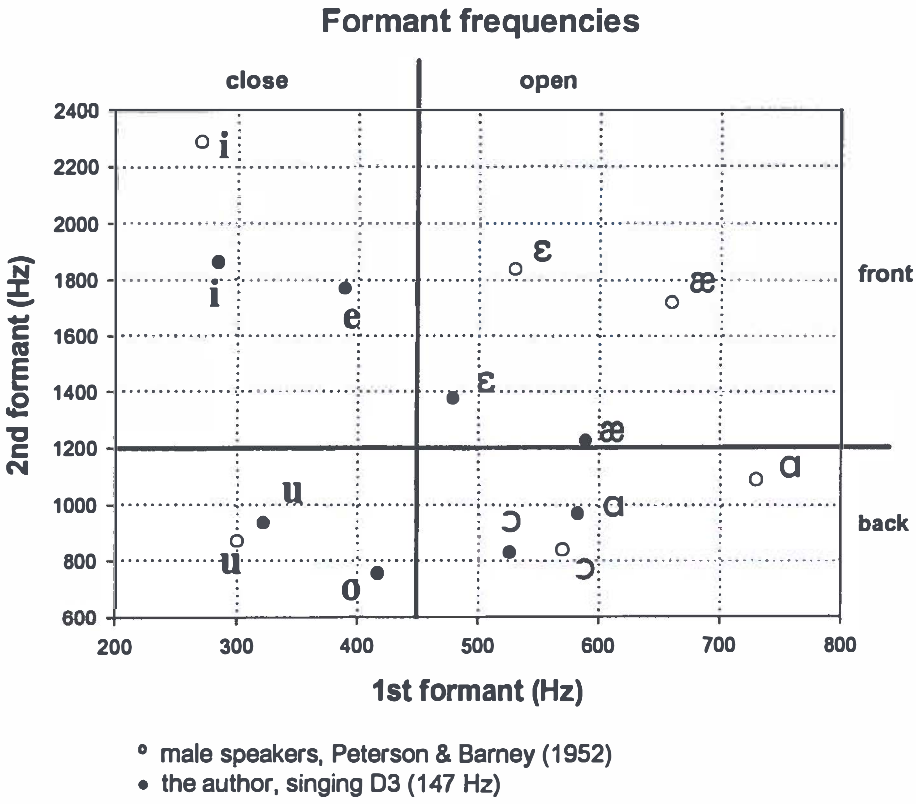
The panels to the right contain signals in the **time domain** (here 30 ms are displayed) captured simultaneously with the spectra. Each pair includes an audio (microphone) signal above and an **electroglottograph** signal below. In contrast to the vocal-fry phonation, the signals of the sung phonation show nearly perfect periodicity. The EGG signal follows **vocal-fold contact area**, and each periodic closing of the glottis results in an abrupt rise in the signal. The horizontal cursor of the EGG signal can be adjusted to the estimated point of **glottal opening** (as a first approximation: the point of maximal falling slope). The three vertical cursors then fall at the points where the horizontal cursor intersects the EGG signal, with the outer cursors lining up with **glottal closing**. From these points the **glottal period** and the **closed quotient** (CQ, the percentage of the period in which the glottis is closed) are calculated and displayed. These are respectively 4.84 ms and 67%. The audio signal has been corrected for delay so that closing and opening are indicated in this signal as well.

**Vocal fry** is the popping sound of small quantities of air passing a mostly-closed glottis. (In this case the air passes in the inward direction, producing a sudden rise in acoustic pressure, in contrast to the sudden fall on the closing of the glottis in the sung phonation.) The periods are long and irregular, producing a **non-periodic** voice source. The closed quotient, 88% for this particular period, is higher than what can ordinarily be found in normal phonation. (For more on the use of vocal fry to ascertain the formants in singing, see Chapters 3 and 10.)

## Appendix 2 – Acoustic phonetics

In order to understand the process of formant tuning in singing (adjusting the resonances of the vocal tract with respect to available harmonics), it is useful to study a chart plotting the first two formants of the various vowels. The figure gives the mean values for adult male speakers of F1 and F2 of six vowels from the frequently-cited study of Peterson and Barney, as well as the author's values for eight vowels in a comfortable singing range. Note the basic vowel dimensions: open/close, corresponding to the value of F1; and front/back, corresponding to F2.

The vocal tract can be roughly modeled as two cavities, front and back, divided by the 'tongue constriction' (where the tongue surface approximates the opposite wall of the vocal tract). The back cavity is affiliated with F1, and a smaller back cavity results in a higher first formant. This explains the natural tendency to elevate the larynx (reducing the size of the back cavity) in



order to raise F1 in the 'register violation.' The front cavity is affiliated with F2, and as the tongue constriction moves forward, the size of the front cavity is reduced, raising F2. Rounding the lips lowers all the formants, but especially F2, since the front cavity is affected most directly.

Formant frequencies of the vowels show considerable variation, often differing markedly from speech values under the constraints of singing. For example, a soprano will routinely raise F1 of even the close vowels to values approaching 1000 Hz in order to match F0, and many tenors raise F2 of back vowels by more than 50% in order to sing their top notes. Even without the distortion occasioned by high notes, the articulation of singing can result in unusual values when compared with those of speech. The set of singing values in the figure is characterized by 'centralization,' corresponding to what singers sometimes call equalization of the vowels.

## Appendix 3 – Glossary of Terms

**AC**

'Alternating current,' used to refer to peak-to-peak variations within another signal. Compare DC, 'direct current.'

**acoustic signal**

Acoustic waves within the audio-frequency-level range, which we hear as sound, as these are picked up by the microphone.

**adduction** (of vocal folds)

Bringing together. Opposite of abduction.

**amplitude**

The maximum excursion from equilibrium of a vibration or waveform.

**back vowel**

A vowel in which the body of the tongue is retracted. Back vowels have relatively low second formants. See Appendix 2, Acoustic Phonetics.

**'chest'**

Used here to designate the glottal vibratory pattern characterized by contact of the inferior portions of the vocal folds and vertical phase difference (see Chapter 2, Fig. 1).

**closed quotient (CQ)**

The percentage of the glottal cycle during which the glottis is closed, preventing the flow of air. This percentage, which varies in singing from zero (no full closure) to >70%, can usually be estimated from the EGG signal.

**close vowel**

A vowel with a low first formant frequency. See Appendix 2, Acoustic Phonetics.

**cursor**

In an electronic display, a movable point or line whose position can be accurately determined, used to mark or measure some feature of the display.

**electroglottograph (EGG)**

A non-invasive electronic device that produces a signal indicating the degree of contact between the vocal folds. From this signal it is generally possible to identify the moment of closure of the glottis and, with somewhat less precision, the moment of opening.

**damping**

Dissipation of the energy of an oscillation, causing attenuation.

**F0**

See **fundamental frequency**.

**'falsetto'**

Used here to designate the glottal vibratory pattern characterized by oscillation limited to the upper margins of the vocal folds (see Chapter 2, Fig. 1).

**formant**

A resonance of the vocal tract, having a variable frequency, depending on the dimensions and posture of the vocal tract. The formants are designated **F1**, **F2**, etc., beginning with the lowest in frequency, and the first five make important contributions to the singing voice. F2, and especially F1, are the strongest single formants, and their frequencies determine the perceived vowel. The **singer's formant**, produced by a clustering among F3, F4, and F5, is, in some male voices, frequently strong enough to lift a higher harmonic (in the range 2.2-3.4 kHz) to dominance.

**formant tuning**

In singing, the modifying of formant frequencies of the vocal tract in order to enhance certain harmonics of the voice source. See **resonance strategy**.

**front vowel**

A vowel in which the body of the tongue is advanced. Front vowels have relatively high second formants. See Appendix 2, Acoustic Phonetics.

**fundamental frequency (F0)**

The repetition rate, in cycles per second, or hertz (Hz), of the oscillating vocal folds, subjectively perceived as pitch. The F0 of G2 (bottom line of the bass staff) is 98 Hz, of C6 (two ledger-lines above the treble staff), 1046 Hz.

**glottal cycle**

A single opening and closing movement of the oscillating vocal folds.

**harmonic**

In the context of voice, sound at the fundamental frequency (designated **H1**) and multiples of that frequency (designated **H2**, **H3**, etc), together forming the familiar **harmonic series** (octave, 5th, 4th, major 3rd, etc. – see **overtone**). The harmonics, produced at the voice source in a series that diminishes in amplitude with increasing frequency, appear in the spectrum of the acoustic signal at varying degrees of strength, depending largely on whether they are close enough in frequency to formants to be enhanced by the vocal tract.

**harmonic sound**

Any sound produced by a source with a regular (**periodic**) repetition rate, usually resulting in the harmonic series. In **non-harmonic sound** (noise), all frequencies are present within a given range.

**ingressive phonation**

Vocal sound produced by drawing air inward across a narrowed glottis. The usage here refers to the relatively loud sound of 'reverse **vocal fry**.'

**loudness**

Magnitude of sound perceived by a listener, a psychoacoustic term.

***messa di voce***

Italian term for a crescendo-diminuendo on a single tone.

***mezza voce***

Italian for 'half voice,' a kind of soft voice production.

**open vowel**

A vowel with a high first formant frequency. See Appendix 2, Acoustic Phonetics.

**overtone**

One of the harmonics above the fundamental frequency. Thus the first overtone is H2, an octave higher than the fundamental frequency. In '**overtone singing**' single upper harmonics are skillfully resonated (by F2) so that they are perceived as separate pitches, distinct from the F0.

***passaggio***

In singing, a transition between registers. It is commonly used to designate the several pitches just below the upper register being approached, and is completed when the upper register is fully established.

**period**

The time interval between repeating events.

**pitch**

The perceived 'height' of a harmonic sound, the psychoacoustic correlate of fundamental frequency. In the scientific literature, what musicians call 'pitches' are given their musical letternames, followed by a number designating the octave (from *C* to *B*) on the piano. Thus 'middle C' (F0=262 Hz) is *C4*, and the *A* just below it (F0=220 Hz) is *A3*.

**primary register transition (PRT)**

In an ascending scale, the point where the vibratory pattern of the vocal folds shifts (or tends to shift) from 'chest' to 'falsetto.'

**$p_{sub}$**
Subglottal pressure. Air pressure below the glottis.

**$p_{supra}$**
Supraglottal pressure. Air pressure in the pharynx, above the glottis.

**register**
A region of a particular vocal quality, perceptually distinct to the listener or the singer as pitch or loudness is changed. Registers are in the first instance subjectively perceived, while the studies presented here attempt to identify the objective physical properties that give rise to the perception of registers. The 'natural registers' result from the distinct vibratory patterns of the vocal folds, '**chest**' and '**falsetto**'.

**registration**
An attribute of the singing voice reflecting the use of the various registers and the transitions between them or, more abstractly, the degree to which the voice production is perceived as 'heavy' or 'light.'

**registration event**
A move from one pitch to another, or from one level of loudness to another, characterized by the perception of a change in the manner of voice production.

**resonance**
Reinforced natural oscillation.

**resonance strategy**
The deliberate or habitual use of the several resonances of the vocal tract to enhance, in ease and/or amplitude, the production of a given vowel and pitch. See **formant tuning**.

**sound pressure level (SPL)**
A logarithmic measure of the physical strength of sound, in decibels (dB). The measure is the log of the ratio between the measured sound pressure and a standard reference (20 microPascal).

**source spectrum**
The (theoretical) spectrum of a vocal sound that emerges from the glottis, without the influence of the vocal tract.

**spectral slope**
A measure of how rapidly energy changes with frequency change in the spectrum.

**spectrum analysis**

The differentiation of a complex sound into its component parts, arranged according to frequency and relative level (amplitude in decibels). Harmonic sound produces a **line spectrum**, with sound only at the harmonics, while non-harmonic sound (noise) yields a **continuous spectrum**. A **real-time spectrum analyzer** gives a nearly instantaneous display of the (changing) acoustic signal.

**subharmonic**

A component of a waveform whose frequency is an integer fraction (1/2, 1/3, etc.) of the fundamental.

**vocal fry**

The relatively soft popping sound, with distinguishable individual impulses, produced when air 'bubbles through' a lightly closed glottis. Also called pulse register. The frequency of the pulses is very low and usually somewhat irregular, producing a virtually continuous spectrum.

**vocal tract**

The partially enclosed air space between the glottis and the openings at the lips and nostrils. This space forms the complex (having more than one resonant frequency) resonator of the voice, reinforcing those harmonics occurring at frequencies close to those of the formants. Changing the position of the movable parts (tongue, larynx, lips, etc.) of the vocal tract changes the formant frequencies.

**VoceVista**

A software program of real-time spectrum analysis and EGG intended for use as feedback for instruction in singing. (Available from the author.)

**voice source**

1. The glottal volume velocity waveform (the intermittent airflow through the glottis).
2. The oscillating vocal folds, driven by airflow through the glottis.

# Biography

Donald Miller was born on February 21, 1933, in Englewood, NJ, USA. He attended the public schools of Mountain Lakes, NJ, finishing high school in 1951 as valedictorian of his class. He earned the degree of Bachelor of Arts, summa cum laude, from Yale University in 1955, with a major in philosophy. He studied theology on a Fulbright Scholarship in Hamburg, Germany, in 1955-56. After obtaining the Master of Music degree from the Yale University School of Music in 1960, he spent the next four years in Milan, Berlin, and Vienna in further study and professional singing (bass-baritone), notably at the Wiener Kammeroper. In 1964 he joined the faculty of the Syracuse University School of Music, where he remained until 1987, having reached the rank of professor. During the years in Syracuse he continued to have professional singing engagements in opera (Tri-Cities Opera, Portland Opera, Syracuse Opera Theater, Artpark, etc.) and concert (Syracuse Symphony Orchestra, Rochester Philharmonic, Baldwin-Wallace Bach Festival, etc.). In 1984 he spent a semester in Groningen doing research on the acoustics and physiology of the singing voice. In 1987 he resigned his position at Syracuse University and settled in Groningen, where he is research associate at the Groningen Voice Research Lab and a private voice consultant. Since 1996 he has been identified with VoceVista, a software program of visual feedback for instruction in singing.

# List of publications not appearing in this volume

D.G. Miller and H.K. Schutte. Characteristic patterns of sub- and supra-glottal pressure variations within the glottal cycle. In: Lawrence, Van L, ed. Transcripts XIII Symposi-um Care of the Professional Voice, New York 1984. New York NY: The Voice Foundation, 1985:70-75.

H.K. Schutte and D.G. Miller. Transglottal pressures in professional singing. Acta Otorhinolaryngo-l Belg 1986;40: 395--404.

H.K. Schutte en D.G. Miller. Het spel van resonanties bij de professionele zangstem. Tijdschr Log Phon 1986;58:273-276.

M. Rothenberg, D.G. Miller, R. Molitor, D. Leffingwell. The control of airflow during loud soprano singing. Journal of Voice 1987;1:262-268.

H.K. Schutte en D.G. Miller. Transglottische drukwisselingen bij de sopraanstem. Tijdschr Log Fon 1987;59:34-38.

H.K. Schutte, D.G. Miller. Resonanzspiele der Gesangstimme in ihren Beziehungen zu supra- und sub-glottalen Druckverläufen: Konsequenzen für die Stimmbildungstheorie. Folia Phoniatr 1988;40:65-73.

M. Rothenberg, D.G. Miller, R. Molitor. Aerodynamic investigation of the sources of vibrato. Folia Phoniatrica 1988;40: 244-260.

D.G. Miller, H.K. Schutte. Formant Tuning in a Professional Baritone. J Voice 1990;4:231-237.

A.M. Sulter, R.F. Wolf, D.G. Miller. E.L. Mooyaard, H.K. Schutte, N.J. Freling. Analysis of parameters of vocal tract dimensions by using MRI. In: Andrew ER, ed. Abstractbook Society of Magnetic Resonance in Medicine. New York. Berkely CA: COS of Magnetic Resonance in Medicine, 1990;18-24 (Abstract).

H.K. Schutte, D.G. Miller. Acoustic Details of Vi-brato Cycle in Tenor High Notes. J Voice 1991;5:217-223.

D.G. Miller, H.K. Schutte. Effects of Downstream Occlusions on Pressures Near the Glottis in Singing. In: Gauffin J, Hammarberg D, eds. Vocal Fold Physiology: Acoustic, Perceptual and Physiological Aspects of Voice Mechanisms. San Diego: Singular Publishing Group Inc., 1991:91-98.

A.M. Sulter, D.G. Miller, R.F. Wolf, H.K. Schutte, H.P. Wit, E.L. Mooyaart. On the relation between the dimensions and resonance characteristics of the vocal tract: A study with MRI. Magnetic Resonance Imaging 1992;10:365-373.

H.K. Schutte, A.M. Sulter, D.G. Miller. Stroboscopy: Categories of vibrational patterns for pre- and postoperative evaluation assessment. In: Mahieu HF, ed. Proceedings 2nd International Symposium Phonosurgery. Amsterdam, May 16th-18th 1992:52 (Abstract).

A.M. Sulter, H.K. Schutte and D.G. Miller. Standardized Laryngeal Videostroboscopic Rating: Differences Between Untrained and Trained Male and Female Subjects, and Effects of Varying Sound Intensity, Fundamental Frequency, and Age. Journal of Voice, 1996;10(2):175-189.

J.G. Švec, H.K. Schutte and D.G. Miller. A Subharmonic Vibratory Pattern in Normal Vocal Folds. Journal of Speech and Hearing Research, 1996;39:135-143.

D.G. Miller and H.K. Schutte. Feedback for instruction in singing. In: Schutte HK, Dejonckere P, Leezenberg H, Mondelaers B, Peters HF (Eds.): 24th IALP Con-gress Amsterdam, the Netherlands, 23-27 August 1998. Program and Abstract Book. Nijmegen University Press. 1998:151 (Abstract)

Donald Miller and James Doing. Male Passaggio and the Upper Extension in the Light of

Visual Feedback. Journal of Singing, 1998;54:4,3-14.

Miller, D. G. & Schutte, H. K. The Use of Spectrum Analysis in the Voice Studio. In: G. Nair; Voice Tradition and Technology. A State-of-the-Art Studio 1999:189-210. San Diego: Singular Publishing Group.

Miller, D. G. & Schutte, H. K. The Use of the Electroglottograph (EGG) in the Voice Studio. In: G. Nair; Voice Tradition and Technology. A State-of-the-Art Studio 1999:211-225. San Diego: Singular Publishing Group.

D.G.Miller en H.K. Schutte. Registers van de Zangstem. In: H.F.M. Peters, P.H.O. Dejonckere et al. (Eds) Handboek Stem-, Spraak- en Taalpathologie Bohn, Scheltema en Holkema. Utrecht. 1999:10;1-16.

J.G. Š` vec, H.K. Schutte and D.G. Miller. A Preliminary Look at Phenomena Accompanying Abrupt Chest – Falsetto Transitions. Journal of the Acoustic Society of America. 1999;106(3):1523-1531.